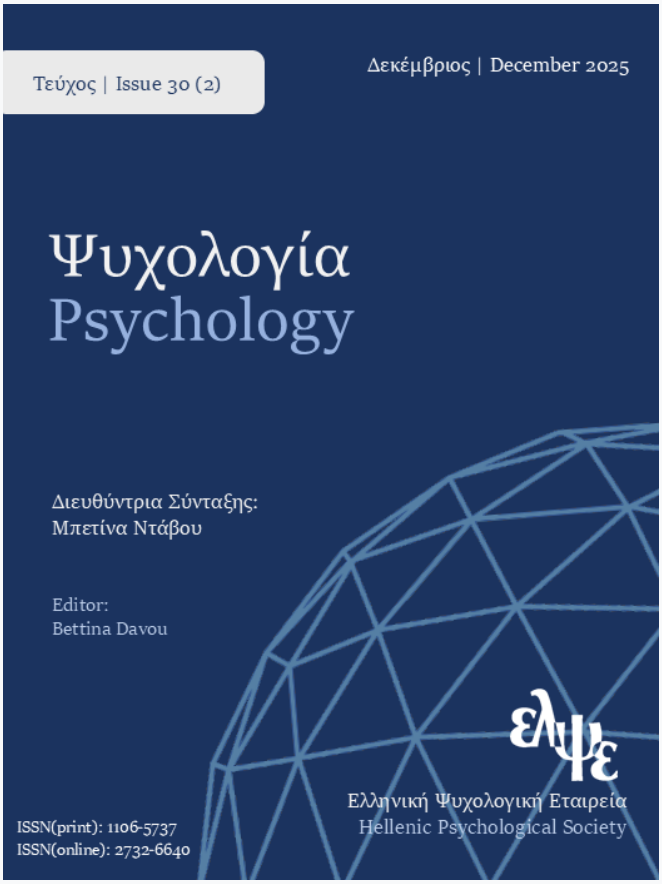
Psychology: the Journal of the Hellenic Psychological Society

Vol 30, No 2 (2025)

Special Section: Individuals, relationships and community in the digital era



A structural approach of the AI social representations. Robot humanizing technology?

Patricia Gerakopoulou, Nikolas Christakis

doi: 10.12681/psy_hps.43979

To cite this article:

Gerakopoulou, P., & Christakis, N. (2025). A structural approach of the AI social representations. Robot humanizing technology?. *Psychology: The Journal of the Hellenic Psychological Society*, *30*(2), 469–487. https://doi.org/10.12681/psy_hps.43979

ΨΥΧΟΛΟΓΙΑ | PSYCHOLOGY 30(2), 469-487

Ελληνική Ψυχολογική Εταιρεία
Hellenic Psychological Society

ΕΜΠΕΙΡΙΚΗ ΕΡΓΑΣΙΑ | RESEARCH PAPER

A structural approach of the AI social representations. Robot humanizing technology?

Patricia GERAKOPOULOU[1], Nikolas CHRISTAKIS[2]

[1] Department of Digital Media and Communication/School of Information Science & Informatics, Ionian University

[2] Department of Communication and Media Studies/School of Economics and Political Sciences, National and Kapodistrian University of Athens

KEYWORDS

Artificial Intelligence
Social representations
Social construction of technology
Scientific imaginary
Robots

CORRESPONDENCE

Patricia Gerakopoulou
Ionian University
Department of Digital Media and Communication
Antonis Tritsis Ave
GR-28100, Argostoli, Greece
patgeraki@ionio.gr

ABSTRACT

Strongly related to a "technological mythology" referred to Promethean promises and threats, Artificial Intelligence draws increasingly impressive attention from the media, as well as from the lay public. The purpose of this research paper is to investigate the content and structure of the AI social representations of undergraduate students from all six Media and Communication University departments in Greece (N=249) relying on original qualitative data collected using a "free association" questionnaire consisting of two open questions. In question one, participants were asked to write three to five words that first come to mind when they think of the term "Artificial Intelligence". In the second question, they were asked to describe further each of their previous answers. The data generated fell under six major thematic categories: Technology, Future, Threats, Uses, Robot, and Human (characteristics). These categories were further analyzed according to frequency and rank to produce the "square of the AI social representation", which is consisted of the central system (Technology & Future), the peripheral system (Threats & Uses) and the "grey area" (Robot & Human). The interpretation and discussion of the results lead to the main conclusion that the representation element of the Robot represents the ideal blurring of the boundaries between Human and AI (the latter far superior in "intelligence", e.g. data processing), attributing more familiar "human" characteristics to the, otherwise, vague and ambiguously perceived (Threats and Uses) technological object.

Modern Artificial Intelligence (AI) originates from Alan Turing's test of machine intelligence, who initially proposed the question "Can machines think?" (1950). Today, the term refers to a wide variety of diverse technologies (algorithms, robotics, machine learning, deep learning, autonomous systems, chatGPT etc.) that span a range of very vague services and applications within the spectrum of computer science (Seaver, 2017). Terminology regarding AI is perceived as highly extensive, referring to sets of widely defined devices of varying degrees of complexity, employing a range of technology and applications under the umbrella of a term that raises ontological questions about the nature of intelligence and the status of artificial life (Katz, 2017).

Unlike other technologies, AI innovations draw impressive attention from the media, as well as from the lay public. One of the main reasons for this appeal is the question of "life" creation (Mosco, 2005). AI strongly relates to a special case of "technological mythology" that has established the idea of humanity overcoming itself, freeing individuals from present constraints, and paving the way toward an apocalyptic future (Natale and Ballatore, 2020). Indeed, imagination is always active and inherent in the systems of representations (and attitudes) that are formed within thinking societies, as long as individuals and groups communicate (act symbolically) and shape this "... partly natural, partly imaginary reality, which is social reality" (Moscovici, 1984a, p.7). So, science and technology share with their traditional and irrational "sisters" (magic and religion) a whole field of social thought, communication and cultural sensibility, which extends from the Promethean promises (and apocalyptic

https://doi.org/10.12681/psy_hps.43979

Psychology: The journal of the Hellenic Psychological Society
https://ejournals.epublishing.ekt.gr/index.php/psychology

threats) of biology and techno-medicine (creation of life, victory over aging, achievement of immortality), to the complete digitization of humanity (unlimited information, management and social participation, and perhaps a "mind without body" existence and dematerialization of the world).

On the other hand, the media representation of novel technologies such as AI has a large impact on the way relevant issues are framed to the public (Chuan et al., 2019), while lack of vivid engagement in the public discussion by technology experts and scientists may lead to a polarization effect, where "hype and hope" and "gloom and doom" perspectives distort the public discussion on the subject (Dubljević, 2012, p.69). Narratives surrounding AI are strongly connected to common visions and apocalyptic scenarios that heavily influence how the public approaches the rapidly advancing technology (Cave & Dihal, 2019).

The general public's emerging representation of AI, however, is under constant investigation as it evolves together with both technological and social change (Gerlich, 2023; Ghiringheli, 2020; Schepman & Rodway, 2022;). The imaginary associations with technologies inevitably involve all kinds of social actors referred to its production and promotion, including mass media, communication experts, and market agents, all participating in a dynamic construction of ever-evolving social reality. In Marcuse's (1967) view, reliability, efficiency, and whether technological cognition is "valid" is judged by society itself and specifically by the extent to which it has integrated it. After all, as Cooley (1980) pointed out, "when we design technological systems, in practice we design grids of social relations." Therefore, it cannot be ignored by science and technology studies (Jasanoff, 2004), and especially by social scientists that an evolving critical discourse emerges highlighting the interconnection of AI technologies with media representations, the public imaginary, and everyday discussions and other symbolic actions that participate in the social construction of the world.

Following the above initial reflections, in this paper, we research the AI social representations of undergraduate students from the six Media & Communication University Departments in Greece. This choice was guided by the assumption that individuals within this demographic, by their educational background and professional orientation, are likely to have at least rudimentary, if not advanced, interest and knowledge of AI technologies, and they therefore contribute significantly to the promotion, diffuse, construction, and regulation of the wider social image of the AI and its meaning in the Greek society. Given that these individuals are often the vanguard of the adoption of anything new, we deemed their representations to be particularly focal for the establishment of AI's social meaning, imaginary, and practices.

To explore the intermediate area where public perception meets expertise and ideology in the making, our research study focuses on the structural approach of the AI social representations of media students in Greece.

Theoretical context and research questions: from AI perceptions to social representations.

Epistemological context on the social perception of new technologies. A technological determinism view implies that techno-scientific change, otherwise material forces and especially the properties of available technologies, constitute an autonomous and omnipotent process that directs the totality of human relations. This is what Ellul (1990) suggested saying that technology has become absolute domination (see also Mumford 1967, 1970).

Social constructivism, on the other hand, suggests that technologies do not evolve outside social processes (Jasanoff, 2004). They are the result of a mixture of social activities, forces, and negotiations (cultural, economic, political, organizational practices, norms, values, ideas, institutions, etc.) in various locations so the social construction of technology emphasizes its multidimensional character (Bijker, 1995).

Another approach, that of the theory of co-production of technology and society is oriented towards their joint formation, coexistence, as they affect each other bilaterally, and simultaneously (Harbers, 2005). Isolated from the influences of technological and social determinism, the theory argues that developments in science and technology cannot be explained solely by the social and political environment in which they develop, producing

results and vice versa, thus leading to a symmetrical approach. In this context, the social and cultural implications of technology and how society and culture are reshaped by its material and digital manifestations are examined in parallel.

The “actor-network” theory (Callon & Latour, 1992) was formulated in an attempt to investigate from a sociological point of view "The truths" of science. This theory proposes the principle of “generalized symmetry”, implying that sociological analysis must play an explanatory role to both human and non-human entities and techno-physical objects (physical forces, properties, technical constructions). For the “actor-network” theory, the scientific, technological, natural, and social objects are formed-structured networks of heterogeneous materials containing symbols, human entities, money, signs, machines, documents, etc (Latour, 2005).

Finally, the theory of transhumanism (Bostrom, 2003) is defined as an interdisciplinary process that aims to understand and evaluate the prospects of human enhancement with the help of technological development (genetic engineering, AI, overdeveloped technology, etc.). In contrast to humanism, which focuses on the value, interest, and dignity of the human being, transhumanism as an ideology holds that the present image of man is not his final one, but that human nature is under a perpetual evolution, which technology can help. The theory is concerned with overcoming basic physical limitations, eradicating illness, and increasing physical, mental, and emotional powers, toward a holistic approach to developing cultural, psychological, economic, and social skills and institutional projects. The transhuman is not understood as a non-human entity, but following Heidegger's prediction that “technology is a mode of revelation”, transhumanism, explores technology as a mode of revelation in an environment that overemphasizes technological achievements (Kurzweil, 2005). Transhuman beliefs are based on a representation of a kind of intelligence that attributes every human performance and therefore could give man superpowers. However, intelligence is also socially constructed, since its epistemological foundation is the information model of cybernetics (“theory of control and communication in animal and machine” as in the sub-title of the famous Wiener (1948) book), which ultimately reduces intelligence to a mere information processing, a simplified and mechanistic form of the human spirit that can “be transmitted” to the machine (Zacklad, 2018).

On AI public perception. AI holds close correlation with an imaginary that refers to the more or less autonomous existence of artificial entities, an imaginary fueled by science fiction literature and cinema, but also by the wider dialogue in traditional and social media. Furthermore, the public AI perception inevitably falls under the spectrum of research traditions and methodologies of several disciplines, which combined lead to questions regarding the ontological status of intelligence and artificial life, especially in their connection with the “human”. In addition, the discussion on AI crosses the social field in a variety of ways, e.g. policy decisions, expert advice and plans, ethical considerations, which also fertilize (positive or negative, soteriological or apocalyptic) public expectations. This is why the variables influencing perceptions of AI gain the interest of social science researchers (Schepman & Rodway, 2022). However, literature review concludes that AI perceptions include multiple elements (Gerlich, 2023). Above all, it is worth noting that although AI concerns every area of everyday life, it is common for people to not recognize its applications and uses, since most of them appear disassociated from the perception of reality, being viewed as futuristic, science fiction technologies (Tai, 2020). However, a survey involving 2000 Americans (Zhang & Dafoe, 2019) found that the 41% of participants where pro AI while 22% were against to it. The Special Eurobarometer (2017) researched the lay public attitudes about AI and robots to find that 61% of Europeans were pro about AI technology and 30% were against, while 68% of research participants thought that AI and robots are helpful for society.

The research literature indicates multiple factors influencing public attitudes towards AI, with risks and trust, as well as uses, to be the most prominent. Primarily, trust appears as the most important factor in shaping public perceptions (Vance et al., 2008), in accordance with research proof that trust supports generally the adoption of new technologies (Venkatesh & Davis 2000; Lewicki et al., 1998). Risks and uses have also a

prominent impact on AI adoption or usage; extensive research in 142 nations and 154,195 participants indicates that people are most worried about the threats of AI uses (Neudert et al., 2020). Relevant to this, a study including 6054 individuals in the US, Australia, Canada, Germany, and the UK highlighted trust as a crucial factor for AI acceptance (Gillespie et al. 2021). However, there is empirical evidence (Rogers, 2003) that misconceptions about AI may influence negatively its adoption (see also the Technology Acceptance Model, Davis, 1989). Furthermore, trust, use, and risks appear to form a positive or negative AI perception (Gerlich, 2023).

A large-scale study including 27,901 participants (European Commission & Directorate-General for Communications Networks, Content & Technology, 2017) revealed a favorable attitude of most Europeans towards robots and AI. Another finding was that attitudes are dependent on knowledge, since more educated individuals, as well as those with heavy Internet use, appeared more favorable towards AI, as well as people of younger age and male sex.

Other studies (Neudert et al., 2020; Park & Woo, 2022, Stein et al., 2024) found that AI adoption is linked to personality characteristics, such as psychological, as well as technological factors (e.g. perceived practicality and ease of use). Additionally, positive attitudes about AI are linked with optimism and trust in science technology, and government. Further research highlights that cultural factors such as norms, and fears related to work replacement also influence participants' attitudes of AI technologies (Kaya et al., 2022).

On the media representation of AI. Crepel & Cardon (2022) show how the new wave of AI mobilizes two contrasting critical representations in the media. The authors studied about 30,000 related articles published (in English) between 2015 and 2019. They found that 7.1% of them were critical as regards ethical issues and social consequences. They carried out systematic corpus analysis using automatic language processing tools, aimed at identifying how the media frame the problems associated with the deployment of artificial intelligence. The semantic mapping of this corpus of press articles on AI revealed two different types of entity, two differentiated "worlds", that of robots and that of algorithms. First, robots are depicted as autonomous, embodied, and independent entities, associated with a prophetic discourse warning about our ability to control these artificial agents that are jeopardizing humans' physical and cognitive capabilities and societal model. Second, the algorithms that shape our everyday computational environments are associated with a critical discourse on bias, discrimination, surveillance, censorship, and the amplified dissemination of inappropriate content. Robots, in particular, are characterized by a high degree of intelligence and AI autonomy of, as well as their ability to incarnate in machines with a kind of their own "personality", with some kind of initiative, ability decision and action. Under this category we find robots, machines, computers, vehicles, weapons, drones or human-like dolls. Robots threaten humans and provoke fear, undermining the very survival of humanity in the distant future, while humans have to resort to institutions such as the military and human rights organizations. On the other hand, algorithms are not anthropomorphic and autonomous but constitute parts or components of computer systems, distributed in digital environments, and empowering or limiting areas of human action actors. This category includes algorithms, devices, programs, tools, and applications such as facial recognition, deepfake, social networks, chatbots, and criminal justice algorithms. Here, references to people are less abstract, and concern social groups with specific positions and roles in society, while questions of justice and law are raised rather closely: discrimination, privacy, human rights, censorship, fraud, and crimes, requiring political and legal regulations.

In conclusion, there is a continuity between people's loss of autonomy in the algorithmic world and the independence of machines in the world of robots in the form of communicating vessels. These two worlds are interdependent, with the media often ready to mobilize an imaginary of total war between humans and robots.

Research purpose and questions: Towards a structural approach to the AI social representation. In the context of our study, the abovementioned literature findings regarding the contradictions hidden in the AI public and media perceptions highlight the role of the representational systems that serve communication, practices, social identity, ideology, and all kinds of symbolic and social presence and interaction in everyday life. In our paper, we explore how social subjects perceive AI through the lenses of the theory of social representations. In this theoretical context, there is not the subject on the one hand, and the object of his/her attitude, on the other. A third element is highlighted, that of social representations, which consists simultaneously of a set of concepts, ideas, and images, but also a mechanism of continuous construction and reconstruction (or building and reprocessing) of meaning (Moscovici, 1984a). Social representations therefore exist both in the minds of (individual and, above all, collective) subjects, as well as in circulation and processing in social communication, that is, in what people do and say (symbols, acts, speeches, media texts, etc.). A cognitive polyphasia (Jovchelovitch, 2002) approach can contribute to the understanding of the significance and meaning that AI acquires in the sociocultural world, when social subjects are trying to capture the essence of a new object/idea/theme, incorporating several forms of knowledge, traditions, paradigms and rationalities originating from science, technology, mass culture, as well as myths, religion, myths, and magic.

In conclusion, the research purpose and the research questions that the study poses are: a) to reveal the internal organization and the main elements of the social representations that Greek media students have on the subject of AI, b) to examine the socio-cultural content and the symbolic meaning that the object of AI receives for social subjects, and c) to explore any links of the AI representation to the collective identity of the group and the system of norms to which it refers (Abric, 1993).

The structural model of social representations

The theory of social representations. There is a broad consensus that social representations are defined as a system of values, concepts and practices related to objects, aspects, or dimensions of the social environment (Moscovici, 2008). This system allows the stabilization of the life framework of individuals and groups (making familiar everything new and strange) and is a tool for orienting the perception of social situations, elaborating attitudes towards the social environment, and giving meaning to action within it. It is a way of interpreting and thinking about everyday reality, a form of social cognition that we often, sometimes pejoratively, call "common sense" or "natural thinking." This knowledge is socially processed, that is, it is constituted through everyday experiences, as well as through information, knowledge, and ways of thinking that we receive and transmit through tradition, education, social communication and of course mass communications. Social representations characterize the modern world, which is dominated by information overload and the consequent demand for opinion or adoption of attitudes, and behaviors by individuals and groups. They spread, circulate and evolve on a large or small scale through communication, and serve to understand the world, to interact with others, and to insure the cognitive, symbolic and practical "control" of the social environment (Jodelet, 1989). Social representations constitute and construct the real: they are performative, like language or symbol, while also selecting and connecting persons and objects according to the terms of society or group, i.e. a collective subject that communicates and acts based on common images and concepts (Jovchelovitch, 2007). In conclusion, social representations constitute an organized set of cognitive elements (cognitions) that a homogeneous population shares on a subject. These elements are varied, they may have a cognitive, informational, ideological, normative character, beliefs, values, attitudes, images, etc., being descriptive, evaluative, or prescriptive.

Social representations and common sense. As a science of "phenomena concerning ideology and communication" (Moscovici, 1984a, p.6), social psychology refers to the study of the perpetual interdependence and interpenetration of science and common sense. It also concerns the modalities of diffusion of science in society and its transformation into common and "mundane" knowledge, as a simple part of our cultural heritage

and (formal and informal) way of thinking, living, and interpreting our everyday experiences. Today's predominant scientific knowledge not only does not abolish common sense and popular knowledge but fertilizes and enriches them (Moscovici, 2001, 2019). Where science and common-sense meet, it is not necessarily the latter who changes the most; common sense is the basis of every cognitive process and spontaneously resists any attempt at reification, while science and technique constantly borrow concepts, images, analogies, constructions, using it to identify, name, describe and communicate their elements. Science is thus becoming a new common sense, the "vernacular or popular metascience", with its own industry, with its own mentality, criteria, and representatives. In this society of constant change, where the continuum and abstruse of technological/scientific change threatens individuals and groups with loss of meaning and of the continuity and the comfort of tradition, it is of major importance the quasi-magical "(…) instantaneous comprehension that their association has on the profane mind" (Kalampalikis, 2014, p. 756).

The unfamiliar becomes familiar through two basic complementary mechanisms: (a) anchoring (transfer the unfamiliar to our sphere where we can compare and interpret it, classify, label and name, reduce a strange, foreign, and disturbing idea to an ordinary suitable category and image), and (b) objectivation (turn something abstract into something obvious and concrete, transfer what is in the mind to something existing in the physical world and accessible by selecting and decontextualizing).

Central and peripheral systems of elements. According to Abric (1993), social representations present two contradictory but also dynamic characteristics: a) they are both rigid and flexible, stable and moving, and b) they are consensual but at the same time marked by strong inter-individual differences. Understanding their function obliges us to consider its content and structure. It is an organized whole: not only in terms of the hierarchy of elements, but in terms of the existence of two structural components, which include a) a central core "whose elements are directly linked to the historical, sociological and ideological conditions", and b) the peripheral system of elements, "which constitutes the indispensable complement of the central system which it depends on" (Abric, 1993, p. 75-76).

Moscovici (2008) was already referring to the notion of the "*noyau figuratif*" (the core element) which results from individuals selectively holding a part of the information circulating in society in relation to the object, organizing/synthesizing this knowledge in a particular way (objectivation) and decontextualizing it, which gives them greater autonomy and therefore increases their possibilities of use. The central core is reality itself, and is therefore simple, concrete, schematic and coherent. At the same time, it corresponds to the value system to which the subsection refers, sealed by the surrounding culture and social norms. As the core is obvious and stable, it presents the framework for selecting, categorizing, and interpreting new information that comes to the attention of the subordinate, directing behaviors and giving meaning to events. Therefore, the central system of elements consists of one or some element(s) that give the representation its meaning. It is determined, on the one hand, by the nature of the represented antique, on the other hand, by the system of values and rules that represent the ideological environment of the moment and the group. As far as we are concerned, the central system of elements has a weaker functional dimension and a stronger normative dimension (i.e. its elements are related to rules, stereotypes, attitudes and concerns, i.e. situations that directly intervene in socio-emotional and ideological dimensions).

After exposing the pertinence of the structural approach about the understanding of social representations, we are going to present the Method and Findings of the research held to investigate the research questions formulated in section 2/paragraph 2.4. (after the bibliographical review).

Method

Participants, research design and procedure

To study the AI social representations of undergraduate students from Media and Communication University departments in Greece (N=249), we used a simple "free association" questionnaire consisting of two open questions.

In question one, our participants were asked to write three to five words that first come to mind when they think of the term "Artificial Intelligence". In a second question, we asked them to describe each of their previous answers with a few more words. The spontaneous character and "projective" dimension of this method of free association allows relatively easy and quick access to the fundamental elements that constitute the semantic universe of representation (see e.g. Abric, 1994; De Rosa, 1988).

Employing convenience sampling, the procedure took place in each one of the six Media and Communication University departments (two located in Athens, one in Thessaloniki, one in Kastoria, one in Argostoli, and one in Mytilini) before or after classes. The researchers or their associates (in some cases) informed participants on the research topic and research ethics, asking for their kind contribution in terms of free will, and ensuring the anonymity of their involvement in the procedure. The study was conducted in accordance with the Declaration of Helsinki principles, and the type of data collected did not raise ethical considerations

Data analysis

Out of 249 questionnaires, a total of 1022 words were collected. To ensure research validity the analysis of the collected data followed three steps. First, the two researchers/authors of the article conducted individually and independently an initial classification of the collected words, taking (when necessary) into account the specifications provided in the second question. At a second stage, the two researchers discussed and merged their separate categorization systems (already significantly converging), concluding on one unified and inclusive classification. Finally, at a third step, an independent researcher was asked to attribute a sample of the collected words into the specified categories, checking the validity of our codification. This process confirmed the initial researchers' classification to a percentage near 90%. The six wide categories detected reflect both the deeper meaning that participants attach to words as they spontaneously retrieve them, and the theoretical/bibliographic framework that guides us in understanding AI. This categorization method resembles classical content analysis of qualitative research material (see e.g. Clark, 1985).

Apart from this quantitative criterion, the ranking of words was attributed according to the order in which each word was reported. The rank of each category was calculated by the formula (Abric, 1996): (Absolute frequency of words falling into the category and denoted in 1st position * 5) + (Absolute frequency of words falling within the category and declared in 2nd position * 4) + and so on) / Total category words.

Findings

The six wide categories coming from the analysis are the following (in descending order of frequency occurrence of words):

Technology. This category refers to anything regarding technological and scientific notions, objects, services, processes, and practices. The most common words included in this category were: technology, machine, computer, software, programming, computing, algorithm, automation, internet, chatGPT, robotics, virtual reality, 3D, animation, Apple, Turing. Indicative specifications that were given in the context of Question 2 were: *"AI is an extension of technology"*, *"with AI technologies we try to automatize everything"*. In total 338 words fall into this category (33.07% of the total 1022 words).

Future. This category covers words and meanings, which represent the existential projection in the future, a quasi-inexorable evolution of our civilization and humanity. It is the embodiment of the "new", representing a forward movement encrypted in humanity's "nature" and fate. Indicative words in this category were: progress, innovation, novelty, development, evolution, change, avant-garde, perspective, topical, unprecedented, change, the unknown. Indicative specifications, which were given in the context of the Question 2 were: *"new possibilities"*, *"perpetual quest for innovation"*, *"AI is the field that represents the emerging evolution", "transition from a state of things to another, a change similar to industrial revolution", "we are in the future about which we talk all these years", "new job positions, technological possibilities we haven't imagine so far"*. A total of 190 words fall into this category (18.59% of the total).

Uses. The category is about the AI beneficial applications and services, suggesting an amelioration in terms of both technical progress, and its positive effects on everyday life. Indicative words in this category: speed, facilitation, adaptability, efficiency, detail, objectivity, clarity, convenience, help, production, answers, problem solving, services, money, research, education, medicine, pharmaceutic industry, design, chess, entertainment, communication. Indicative specifications, which were given in the context of the Question 2 were: *"speed of processes and time-saving"*, *"our lives become easier, both in our everyday routine, as well as in our field of work"*, *"mass production increases and costs diminish", "progress in the field of medicine and biomedical sciences", "AI provides attention to the slightest detail and this makes it a lot more efficient than the human"*. A total of 178 words falls into this category (17.42% of the total).

Robot. This category covers descriptions about technological devices and machines taking the appearance and some behavioral aspects of human beings. It appears that the Science Fiction imaginary and cultural universe feed the description of the Robot within this context. The most common word appeared in this category is "Robot". However, words such as "humanoids", as well as references to science fiction ("SF") and science fiction literature, films, and characters appear often. However, the specifications, which were given in the context of the Question 2, reveal that Robots may look like humans simulating human characteristics and capabilities although lacking the essence of humanity (see category 6). Indicatively: *"Robots look like humans"*, *"technology taking the form of human"*, *"digital humans, machines, speech"*, *"the idea of the Robot embodies the "physical" dimension of AI"*, *"something like a super-human with special capabilities"*, *"metal that takes "life", "Robots in the future will work as humans due to AI", "a kind of a device with an artificial mind"*, *"Robot is a kind of simulation", "I have associated AI with Robots that can do whatever people can do", "responses close to the human being"*. In total 141 words fall into this category (13.8% of the total).

Threats. The category includes various perceived dangers, pessimistic projections, and fears concerning the future of work, democracy, social order, culture, the structure of society, and even the survival of humanity. Indicative words included in this category: fear, terror, control, domination, enforcement, abuse, job loss, unemployment, human replacement, cloning, uncertainty, misinformation, isolation, poverty, war, terrorism, death, hyper-reality, and dystopia. Indicative specifications, which were given in the context of the Question 2 were: *"Will AI be used for the Common Good or in a deceitful way?"*, *"unpredictable, cause we don't know how it will end up"*, *"replacement of the labor hands by automatized machines", "artificial, fake, unhuman"*, *"machine-controlled societies, massification, non-specificity, everyone the same, control, globalization*", *"AI contributes to the production of misinformation and deep fake news", "diffusion of the screen to screen way of life and regress of the face to face"*. A total of 113 words fall into this category (11.06% of the total).

Human. This category describes essential and profound qualities attributing to human intellect and psyche. It refers to essential and distinguishing "human" characteristics that so far remained imaginarily untouchable by anything artificial (nothing coming from the artificial and technological world could simulate or mimic them). However, this category indicates that the ongoing emergence of AI technologies may conquer this nuclear

domain, so far exclusively occupied by humans. Words that appear in this category include mind, brain, intellect, emotions, insight, critical ability, critical thinking, imagination, ideas, neural networks, psychology, imagination, philosophy. Some specifications, which were given in the context of the Question 2 were: “simulation of the human brain neurons”, “via AI machines become capable of reproducing cognitive functions of the human”, “the human as an entity that will manage to create the relationship between the man and the machine”, “the relationship of the man with technology as an unmediated interaction”, “they say that Robots will be capable to feel even if it is impossible to imagine”. A total of 62 words fall into this category (6.07% of the total).

As regards the process of how "quickly" or "automatically/unmediated" the words came to the mind of participants, the average rank is 3.27. The ranks are presented below in descending order:

1. Robot. Medium Category Rank (MCR): 3.94
2. Technology. MCR: 3.64
3. Future. MCR: 3.27
4. Human. MCR: 3.06
5. Uses. MCR: 2.94
6. Threats. MCR: 2.75

By cross-referencing these two criteria, the "square of social representation" emerges. (For this method of analysis of social representation, see Vergès (1994).

Graph 1 below presents the identified categories. In brackets under the name of each category, the frequency of occurrence of the words and their ranking position within each category are reported. At the intersection of the two axes, the average frequency and the average rank are displayed (170.3 and 3.27, respectively).

In the upper right quadrant appear the categories whose words have both a high incidence (above average) and a high rank (also above average): Technology and Future (the rank of this category coincides with the average rank). They are therefore the categories in which the words appear most frequently and quickly. These categories are part of the central system of the social representation.

In the lower left quadrant are the categories whose words have both low frequency and low rank (peripheral elements of the representation): Threats and Human.

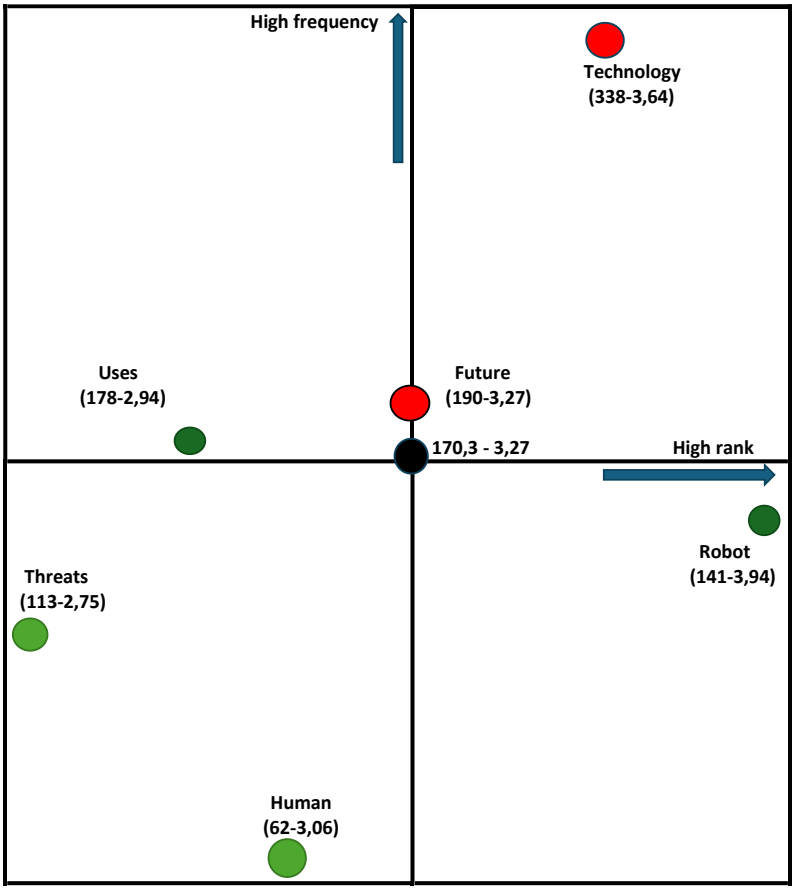
The remaining two quadrants (upper left and lower right) include categories that score high on one criterion but low on the other. These quadrants can be considered ambiguous, as they represent categories with uneven contributions across the two dimensions. The categories *Uses* and *Robot* fall within these “grey” areas of representation, which are not to be interpreted as less important.

Discussion and interpretation of findings

Central system of the AI social representation

Technology & future. We will begin discussing the findings, starting with the central system elements, in particular the category described as Technology. Since the earliest science fiction novels, technology has stood in for the future, the radically new or different, and an obsession for modernity. We define our societies by the technologies used (from the definition of the “Stone Age” to that of the “digital era”) and fit those definitions into linear, progressive models of technological advancements. Summarizing in simple terms, AI is a computational program with specific mechanisms and functions that its conceptors call "intelligence" (based on the -socially-constructed meaning of intelligence itself), and which are related to research on artificial neuron networks and the enormous processing power, Deep Learning algorithms that allow (statistical) predictions extracted from large amounts of raw data, as well as from the internet and the so-called “big data”. By default, technology appears at the heart of the AI social representation, especially by students whose academic background and professional future are situated in the digital technology and communication field.

Graph 1. *Square of AI social representation*



As mentioned above, the central system, composed as the core of the social representation, is directly linked and determined by the surrounding historical, ideological, and sociological conditions, strongly marked by the system of norms to which it refers. It constitutes a collectively shared basis, through a consensual function defining the homogeneity of the social group (Abric, 1993). The nature of the object and the subject's connection to it is technological, while both are situated at the heart of the so-called “digital era”.

This era, following modernity’s command (Giddens, 1990), is clearly “forward-looking”, a condition indicated by the second element of the central system, namely the Future. It is a category that refers to the adventurous path to the unprecedented and unknown, that is progress. In conjunction with the “grey” representational element of Uses (upper left quadrant of the AI representation square), it refers both to technical advances in performance (speed, efficiency, etc.), and to its beneficial areas of application (medicine, education, research etc.).

It is important here to add that in the theoretical and methodological framework of the structural approach, social representation is knowledge in which the possibilities of action in relation to the represented objects are also important. It seems here that research participants, as students introduced to new means, forms, and phenomena of digital technology and communication, feel that they are part of something emerging from the digital era, and are watching something happening, something that surpasses them, surpasses us all: technological research and application at the AI level is a social and cultural one-way forward motion. Under the technological deterministic approach “in a world dominated by the frenzy of scientific and technical progress, as soon as something can be done, it is accepted that it must be done” (Lewontin, 1987, p.55).

It seems that the question that emerges through participants’ AI representation encloses the ambiguity of the use of AI technology in the Future; What are we going to do with it in the future? What is coming next? How will what is going to come affect our mere existence? This is where the central system of Technology and Future

generate conjunctions with the peripheral system of Threats (down left quadrant of the AI representation square) and with the "grey" area of Uses.

All these appear in line with Baudrillard's ideas on technology, in his book "System of objects" (1997/1968). He discussed how technological experiences are inherent in the present and everyday life, mapping the transition from modern to postmodern society. In Baudrillard's terms, the technological object does not embody the grand progress narratives; instead, it is restricted by the interpretation of fantasy and desire. In this sense, technology holds back from "true" development and is limited in its application and pre-programmed stereotypes. Referring to science fiction's "robot" archetype, which is both a technological object of preference and an artificial slave, the ideal robot could perform anything a human being does. However, by being a perfect simulation, it would allow space for confusion (Breton, 1995).

Next, we are going to examine in more detail the elements of the AI representation's peripheral system (Threats and Human), further reflecting on their inter-connections with the central element of Technology, as well as with the AI representation's "grey area" (Uses and Robot).

Peripheral system of the AI social representation

As mentioned, contrasting with the central system, the peripheral system is far more sensitive and determined by the immediate context characteristics constituting the interface between concrete reality and the core of the representation (Abric, 1993). The peripheral system elements indispensably complement the central system on which they depend. In the following presentation of the peripheral system elements of the AI representation (Threats and Human), we see how the representation is rooted in the reality of the moment, with reference to the two basic mechanisms of anchoring and objectivation, through which the unfamiliar (AI) becomes familiar (Moscovici, 1984b), something that falls under the wider psychosocial need for conventionalization. Novelty, and what opposes the natural flow of things, which cannot be conceived or interpreted with the available mental tools – as is often the case with the (most difficult to handle with, cutting edge) products of science and technique – triggers the search for meaning and explanation. In this context, the past prevails over the present, irrational structures over current intellectual or perceptive structures, memory over deduction and reason, response over stimuli, and images over "reality".

Threats and/vs uses Threats appear inherent in technology, even more so in the case of AI because it is cutting-edge technology, always in the making, in statu nascendi, with a permanent opening to the future, and therefore to the unknown, the destabilizing and insecure, and perhaps the unfamiliar, giving rise to negative associations and dark sentiments, from individual emotional insecurity and socio-political dystopia to the threat of Armageddon. Thus, as Sontag (2009) notices, a central science fiction concern is the aesthetics of destruction and the peculiar imaginary of wreaking havoc, making a mess, and the threat of the destruction of humanity itself.

The peripheral element of Threats raises connotations similar to Giddens' Juggernaut of Modernity (1990) which can be directed to some extent but threatens to run out of control; a runaway with great increases over prior systems in the place, scope, and profoundness of change, leading to Beck's (among others) discussion on "risk society" (1992). Risk becomes global in intensity and in the expansion of contingent events that affect large numbers of individuals in the mass society. Awareness of these risks generates the sense of insecurity implied in the Juggernaut archetype.

The interconnectedness of (post)modern society, where every occurrence has a cause (Butler, 1999) is in line with the deterministic approach that takes the notion of security to the limit, a kind of oversaturation of secure systems that leads to its opposite. Baudrillard (1990, p. 37) theorizes three stages in the pursuit of increasing security, leading to its reversal: first, "a relative loose, diffuse and extensive state of the system produces liberty"; second, "a different state of the system (denser) produces security (self-regulation, control,

feedback, etc.)"; third, "a further state of the system, that of proliferation and saturation, produces panic and terror". An example of a similar transition may be incorporated in the AI interconnected representations of Uses and Threats, both also in rapport to the central element of Future: on the one pole, we find gaining advanced security and freedom which takes various functionalistic forms via AI applications as accounted under the category of Uses (in Medicine, Science, Research, etc.). On the other side of the spectrum, we find a heightened sensitivity to possible negative AI uses that may threaten to become incredibly frightening realities e.g. replacement of man by machines, total control, the ravage of mankind…) and their sentimental equivalent (fear, terror, uncertainty, lack of confidence and trust…). The territory of AI threat is represented as no longer located at the margins of society (as used by a "distant" and "aloof" group of expert people), at a fracturing or insertion point, but as an inherent part of our "natural", everyday life condition, everywhere present, however invisible.

Technology and the digital era scientists are perceived as able to release forces that, if not controlled for good, could destroy humans themselves, same as one of the oldest images of the scientist, Shakeshpear's Prospero, the overdetached scholar forcibly retired from society to a desert island, only partly in control of the magic forces in which he dabbles. Equally, classic is the figure of the scientist as satanist (from Doctor Faustus to Doctor No, and stories of Poe and Hawthorne). Technology, same as science, is perceived in the imaginary as magic (Sontag, 2009; see also Caillois' approach of the "sacred", 2001), and man has always known that there is "black" magic (Threats), as well as "white" (Uses). It seems that the emerged AI representations (as reflected in the research participants' accounts, as well as in the Science Fiction's concept on Technology since ever) remain ambivalent, since Technology and its "priests" (digital scientists, and famous representatives such as Bill Gates, Elon Mask etc.) are treated as both black magicians and saviors.

The "grey area": Human – robot. We can further examine and better understand this contradictory dynamic of Uses (represented as hopeful promises) and Threats in the light of the Human and Robot categories. Since, according to Moscovici (1984b), when studying a social representation, it is important to discover the unfamiliar element that motivated it, and which was absorbed by it – but we need to know how this element developed once it emerged in the public sphere.

The various terms used by research participants (e.g. "cognition", "wisdom") for the notion of intelligence as an (uber ales) human characteristic represent the ideal blurring of the boundaries between human and AI characteristics. A kind of "second-order simulation" (Baudrillard, 1880/1983), where reality and representation blur together (same as Disneyland, where fake castles look more real than the actual ones because they embody our childlike notions of what a castle should ideally look like).

Humans are self-conscious beings, so they need to "process" themselves cognitively and mentally through images they create for themselves. Technique, as an extension of it, offers a privileged path, so humans have always been interested in their artificial models, which have led to different kinds of creations depending on the epistemological and social context. We can mention here in the order in which they appear in mythology and literature: Talos, Pygmalion, Golem, Frankenstein, robots, androids… which refer, among other things, to the super-powers attributed to them (projection of human technical ability), to the always ambivalent relationship between creator and creation, to the potential loss of control of the former over the latter, to the confusion between natural and artificial. All of this may boil down to the (ambivalent!) concern that it will replace humans; from Space Odyssey's HAL and Terminator's machines to the common existential fear of extinguished professions and get individuals replaced by the AI, leading to the elimination of humanity itself, with (or without) its replacement by purely mechanical social beings (see research participants' verbatims in section 5).

It is also clear that these technological human creations do not refer solely to the technological imaginary in combination with the era and its cultural "sensibilities" and styles (enlightenment, romanticism, etc.), but to the broader question "what is human, what characterizes the human and what are its limits" (Breton, 1995). As well as perhaps how the "human" and the cultural-social define or influence one another.

Undoubtedly, AI is, for modern common sense, the artificial being of our time, containing a very important human trait in today's cultural context, intelligence (as perceived in the current Zeitgeist as computing power and efficiency). Intelligence, however embodied in everyday life through a variety of expressions among individuals, and through its biological dimension (remember Barthe's (1972) "Einstein's brain") and its ubiquitous achievements, requires another degree of domestication and adaptation to the "human", and to the most tangible, a more "humanized" image of it, which must be at the same time sufficiently technological - this is obviously the case of the Robot. The Robot, with its huge career in mass culture and social imaginary of the 20th century, is indeed much more anthropometrically and anthropomorphically adapted (than the algorithms) to anchor the representation, mainly in terms of its virtual part, of AI.

In terms of establishing a social representation, the Robot probably obeys both the anchoring (reducing a strange, distant, and vague object to an ordinary and accessible category), as well as the objectivation (turning something abstract and unfamiliar into something familiar, obvious, almost concrete). As regards anchoring, AI appears incarnated in the Robot (an object strangely resembling a human), while as regards objectivation, the abstract and elevated idea of intelligence becomes tangible in a machine capable of infinite processing. A crucial note at this point: in Mugny & Carugati's (1989) study on the social representation of intelligence, the central system consists of the everyday experience of intelligence differences that exist between individuals. However, in the case of AI, we have an intelligence (artificial), perceived as overwhelmingly superior (in performance) to that of humans.

The unfamiliar (actuality of something absent, the "not quite rightness", the similarity of the difference, the accessibility of the inaccessible) attracts and intrigues individuals and communities, while, at the same time, it alarms them, compels them to make explicit the implicit assumptions that are basic to consensus; "...worries and threatens, as when a robot that behaves exactly like a living creature, although it lacks life itself, suddenly becomes the Frankenstein monster, something both fascinating and terrifying" (Moscovici, 1984b, p. 25), entering the area of the Freudian "uncanny" (1919), where an inanimate object is coming alive.

Furthermore, robots are closer to humans as they appear to have a structure resembling a body, made of something tangible, not abstract such as merely digits. They look like they are mortal (we see them dissolve or dismantle in movies, they can fall downstairs and break, explode, get an electric stroke, etc.); you see that they are perishable and may be eliminated, as well as human bodies do. Even if they are not bodies, they are metonymic bodies (e.g. the arm of the car assembly chain). In contrast, algorithms never die; they appear to live eternally growing endlessly fed by information and data.

However, of all standard motifs in science fiction imaginary, the theme of dehumanization is one of the most threatening, although there is an ambiguity in it, too. On the one hand, there is the profound existential threat of "replacement", deplored in the peripheral system of the research participants' representation as the ultimate horror. On the other hand, certain characteristics of a more or less dehumanized model for individual and social existence, such as the ascendancy of reason over feelings, the idealization of digital efficiency, the consensus-creating activities of science, a marked degree of moral simplification among others, appear precisely as traits of AI, the savior, the ideal model of intelligence to which human intelligence and social being should imitate. It is not about an AI created "in the image and likeness" of human intelligence, but about human beings that shortly will be upgraded into "humanoids" having all the properties of a "real" digitalized human. Furthermore, this attainment may lead to intense anxiety and feeling of threat; the kind of anxiety described by Philip K. Dick "Do Androids Dream of Electric Sheep?" and its film adaptation "Blade Runner" (Baudrillard, 1997). The Robot is seen as the ultimate technological perfection, however not as good as a human being in his/her subjectivity. In terms of its evolution, the robot is thus a dead end, as all objects in our consumer world.

Robot vs Frankenstein

This research paper aimed at investigating the social representations of Media and Communication undergraduate students in Greece towards AI, relying on original data collected using a free association questionnaire. The main conclusion is that the representation element of the Robot (involving the two basic social representations' mechanisms of anchoring and objectivation) represents the ideal blurring of the boundaries between Human and AI, attributing more familiar characteristics to the, otherwise, vague and ambiguously perceived (Threats and Uses) technological object.

Social thinking shares with technical thinking the concern for efficiency and instrumentality, that is, for control of the environment, so the rightfulness of thought is determined only by social action (Flament & Rouquette, 2003). Knowledge and technique are not developed in any direction, but are guided by opinions, symbols, rituals, as is the case with every aspect of coexistence. The view here is that in our society, science and common sense coexist dynamically, influencing each other interdependently. However, in our social and symbolic life (and representational joint process) not only communication but also personal processing plays a crucial role: much of our waking (even if not necessarily conscious) life is spent telling ourselves stories about the world, recreating the past and inventing the future (Bruner, 2003), a concern shared by scientists' theoretical narratives (Latour & Woolgar, 1987).

Considering the above, the notion of science and technology as a social activity, interlocking with social and political agents is unacknowledged behind the split antithesis between AI perceived as threat, and AI represented through its beneficial and practical uses. There is no room for social criticism in the Greek Media and Communication students' representation regarding the conditions of our society that create the impersonality and dehumanization displaced onto the influence of an alien "It", such as AI. The depersonalization (being "taken over" or "replaced") allegory disguised in the AI representation as "Threats" and "Uses", and as "Robot" and "Technology" with "Human Characteristics" perhaps expresses the depersonalizing conditions (and their equivalent defenses) of modern urban life and the digital era, offering an area of new myths about the perennial human anxiety of death, same as the myths of ghosts, vampires and, even, trans-humans. Ours is, indeed, an age of extreme need for consolation and compromise of "two equally fearful, but seemingly opposed, destinies: unremitting banality and inconceivable terror" (Sontag, 1996, p.224). "Frankenstein" creations such those of the common mind may offer a Darwinian adaptation to rapidly emerging survival challenges.

An approach in terms of social psychology is "about the analysis of the internalization and of the externalization of the social on the individual's level, but also about the understanding of the intervention of the individual on the social level" (Moscovici, 1970, p. 35). As regards the group identity of our research participants, it refers to the psychological group ("undergraduate students"), but it is situated also in an everyday common context: students interact with each other in class, at the campus, at their school… They are involved in a continuous processing of the social and representational reality regarding AI, feeling and experiencing their interdependence in constructing the meaning of this new object and its place in the world (in what extent AI is going to change everyday life? Its effects in social reality will be beneficial or destructive?). After graduating, each unique student will find him/herself in a different environment and they will become members of new groups, something that will change further their representation of the object and at the same time will bring changes to the object itself, in a dynamic interaction with other images of other groups of the social field. We may hypothesize that this group is on hold; aware of the multiple aspects of the AI object experiencing cognitively and emotionally its ambiguity and conflictuality, worried and reserved.

This study's aim was to contribute to the investigation of the AI representation's structure and content. However, it seems that it puts also under further investigation the future of knowledge processing, acquisition and transmission per se. As one of our participants put it: "AI will replace university professors".

References

Abric, J.-Cl. (1993). Central system, peripheral system: their functions and roles in the dynamics of social representations. *Papers on Social Representations, 2*(2), 75-78.

Abric, J.C. (1994). Méthodologie de recueil des représentations sociales [Methodology for collecting social representations]. In Abric J.C. (Ed). *Pratiques sociales et représentations* (pp. 59–82). Presses Universitaires de France.

Abric, J.C. (2003). L'analyse structurale des representations [Structural analysis of social representations]. In S. Moscovici (Ed.), *Méthodologie des sciences sociales*. Presses Universitaires de France.

Barthes, R. (1972). *Mythologies.* Paladin.

Baudrillard, J. (1990/1983). *Fatal strategies.* Pluto.

Baudrillard, J. (1997/1968). *The system of objects*. Verso.

Beck, U. (1992). *Risk Society*. Sage.

Butler, J. (1999). Revisiting Bodies and Pleasures. *Theory, Culture & Society*, *16*(2), 11-20. https://doi.org/10.1177/02632769922050520.

Bijker, W. (1995). *Of Bicycles, Bakelites, and Bulbs: Toward a Theory of Sociotechnical Change.* MIT Press.

Bostrom, N. (2003). *Transhumanist Values, Ethical Issues for the 21st Century*. Philosophical Documentation Center Press. https://doi.org/10.5840/jpr_2005_26

Breton, P. (1995). *À l'image de l'Homme : Du Golem aux créatures virtuelles* [In the image of humankind: From the Golem to virtual creatures]. Seuil.

Bruner, J. (2003). *Making stories*. Harvard University Press.

Caillois, R. (2001). *Man and the sacred*. University of Illinois Press.

Callon, M., & Latour, B. (1992). Don't throw the baby out with the bathwater! A reply to Collins and Yearley. In A. Pickering (Ed.), *Science as practice and culture* (pp. 343–368). University of Chicago Press.

Cave, S., Dihal, K., Dillon, S. (Eds). (2020). *AI narratives: a history of imaginative thinking about intelligent machines*. Oxford University Press.

Chuan, C.-H., Tsai, W.-H., & Cho, S. (2019). Framing artificial intelligence in American newspapers. In *Proceedings of the AAAI Workshop on AI, Ethics, and Society*. AAAI Press. https://doi.org/10.1145/3306618.3314285

Clark, H. H. (1985). Language use and language users. In G. Lindzey & E. Aronson (Eds.), *Handbook of social psychology* (Vol. 2, pp. 179–231). McGraw-Hill.

Cooley, M. (1980). *Architect or bee? The human–technology relationship*. Transnational Co-operative Ltd.

Crépel, M., & Cardon, M. (2022). *Robots vs algorithmes: Prophétie et critique dans la représentation médiatique des controverses de l'IA* [Robots vs algorithms: Prophecy and critique in the media representation of AI controversies]. *Réseaux: Communication, technologie, société*, (232–233), 129–167. https://doi.org/10.3917/res.232.0129

Davis, F. D., Bagozzi, R. P. & Warshaw. P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, *35*, 982–1003. https://doi.org/10.1287/mnsc.35.8.982

De Rosa, A. S. (1988). *Sur l'usage des associations libres dans l'étude des représentations sociales de la maladie mentale* [On the use of free associations in the study of social representations of mental illness]. *Connexions*, 51, 27–50.

Dubljević, V. (2012). Principles of justice as the basis for public policy on psychopharmacological cognitive enhancement. *Law Innovation and Technology*, *4*(1), 67–83. https://doi.org/10.5235/175799612800650617

Eurobarometer. (2017). *Attitudes towards the impact of digitisation and automation on daily life* (Special Eurobarometer 460). European Commission. https://europa.eu/eurobarometer/surveys/detail/2160

European Commission. (2017). *Attitudes towards the impact of digitisation and automation on daily life* (Special Eurobarometer 460; Directorate-General for Communications Networks, Content and Technology). https://europa.eu/eurobarometer/surveys/detail/2160

Flament, C., & Rouquette, M. L. (2003). *Anatomie des idées ordinaires : Comment étudier les représentations sociales* [Anatomy of ordinary ideas: How to study social representations]. Armand Colin.

Freud, S. (1919). The uncanny. In J. Strachey (Ed. & Trans.), *The standard edition of the complete psychological works of Sigmund Freud* (Vol. 17, pp. 217–256). Hogarth Press.

Gerlich, M. (2023). Perceptions and Acceptance of Artificial Intelligence: A Multi-Dimensional Study. *Social Sciences, 12*, 502. https://doi.org/10.3390/socsci12090502.

Ghiringhelli, A. (2020). *Intelligence artificielle : Impacts des représentations sociales de la notion d'« intelligence » sur le secteur de la santé* [Artificial intelligence: Impacts of social representations of the notion of "intelligence" on the health sector]. *Médecine et Philosophie*, 2, 34–40.

Giddens, A. (1990). *The consequences of modernity*. Stanford University Press.

Gillespie, N., Lockey, S. & Curtis, C. (2021). *Trust in artificial intelligence: A five country study*. The University of Queensland and KPMG.

Harbers, H. (2005). *Inside the politics of technology: Agency and normativity in the co-production of technology and society*. Amsterdam University Press.

Jasanoff, S. (2004). Ordering knowledge, ordering society. In S. Jasanoff (Ed.), *States of knowledge: The co-production of science and the social order* (pp. 13–45). Routledge.

Jodelet, D. (1989). *Les représentations sociales: Un domaine en expansion* [Social representations: An expanding field]. In D. Jodelet (Ed.), *Les représentations sociales* [Social representations] (pp. 31–61). Presses Universitaires de France.

Jovchelovitch, S. (2002). Re-thinking the diversity of knowledge: Cognitive polyphasia, belief and representation. Culture & Psychology, 8(4), 389–414. https://doi.org/10.1177/1354067X0200800400 2

Jovchelovitch, S. (2007). *Knowledge in context: Representations, community and culture*. Routledge.

Kalampalikis, N. (2014). The magic of social thought. *Public Understanding of Science, 23*(7), 755–758. https://doi.org/10.1177/0963662514527494

Katz, Y. (2017). Manufacturing an artificial intelligence revolution [Preprint]. SSRN. https://doi.org/10.2139/ssrn.3078224

Kaya, F., Aydin, F., Schepman, A., Rodway, P., Yetişensoy, O., & Demir Kaya, M. (2022). The roles of personality traits, AI anxiety, and demographic factors in attitudes toward artificial intelligence. *International Journal of Human–Computer Interaction, 40*(2), 497–514. https://doi.org/10.1080/10447318.2022.2151730

Kurzweil, R. (2005). *The Singularity is Near*. Viking Books.

Latour, B. (2005). *Reassembling the social. An introduction to Actor-Network Theory*. Oxford University Press.

Latour, B. & Woolgar, S. (1987). *Laboratory life. The social construction of scientific facts*. Princeton University Press.

Lewicki, R. J., McAllister, D. J. & Bies. R. J. (1998). Trust and distrust: New relationships and realities. *Academy of Management Review*, *23,* 438–58. https://doi.org/10.5465/amr.1998.926620

Lewontin, R. C. (1987). *L'évolution du vivant: Enjeux idéologiques* [The evolution of life: Ideological stakes]. In A. Jacquard (Ed.), *Les scientifiques parlent* [Scientists speak] (pp. 54–73). Hachette.

Marcuse, H. (1964). *One-dimensional man. Studies in the ideology of advanced industrial society*. Beacon Press.

Mosco, V. (2005). *The Digital Sublime: Myth, Power and Cyberspace*. The MIT Press.

Moscovici, S. (1970). Préface [Preface]. In D. Jodelet, J. Viet, & Ph. Besnard (Eds.), *La psychologie sociale: Une discipline en mouvement* [Social psychology: A discipline in motion] (pp. 9–64). Mouton.

Moscovici, S. (1984a). Introduction: Le domaine de la psychologie sociale [Introduction: The domain of social psychology]. In S. Moscovici (Ed.), *La psychologie sociale* (pp. 5–22). Presses Universitaires de France.

Moscovici, S. (1984b). The Phenomenon of Social Representations. In: R. M. Farr, & S. Moscovici (Eds.), *Social Representations* (pp. 3-69). Cambridge University.

Moscovici, S. (2001). Why a theory of SR?. In K. Deaux, & G. Philogène, (Eds.). *Representations of the social: Bridging theoretical traditions* (pp. 18-61). Blackwell Publishing.

Moscovici, S. (2008). *Psychoanalysis: Its image and its public*. Polity Press.

Moscovici, S. (2019). Sens commun : représentations sociales ou idéologie ? [Common sense: Social representations or ideology?]. In N. Kalampalikis (Ed.), *Serge Moscovici : Psychologie des représentations sociales. Textes rares et inédits* (pp. 17–29). Éditions des Archives Contemporaines.

Mugny, G., & Carugati, F. (1989). *Social representations of intelligence*. Cambridge University Press.

Mumford, L. (1967). *The Myth of the Machine: Technics and Human Development*. Harcourt Brace Jovanovich.

Mumford, L. (1970). *The Myth of the Machine: The Pentagon of Power*. Harcourt Brace Jovanovich.

Natale, S., Ballatore, A. (2020). Imagining the Thinking Machine: Technological Myths and the Rise of Artificial Intelligence. *Convergence: The International Journal of Research into New Media Technologies*, 26(1), 3-18. https://doi.org/10.1177/1354856517715164

Neudert, L.-M., Knuutila, A., Howard Ph. (2020). *Global Attitudes Towards AI, Machine Learning & Automated Decision Making Implications for Involving Artificial Intelligence in Public Service and Good Governance*. Oxford Commission on AI and Good Governance. https://www.oii.ox.ac.uk/research/projects/public-attitudes-to-ai/

Park, J. & Woo, S. W. (2022). Who likes artificial intelligence? Personality predictors of attitudes toward artificial intelligence. *The Journal of Psychology,* 156, 68–94. https://doi.org/10.1080/00223980.2021.2012109

Rogers, E. M. (2003). *Diffusion of Innovations*. Free Press.

Seaver, N. (2017). Algorithms as culture: Some tactics for the ethnography of algorithmic systems. *Big Data & Society*. https://doi.org/10.1177/2053951717738104

Schepman, A., & Rodway, P. (2022). The General Attitudes towards Artificial Intelligence Scale (GAAIS): Confirmatory validation and associations with personality, corporate distrust, and general trust. *International Journal of Human–Computer Interaction, 39*(16), 2724–2741. https://doi.org/10.1080/10447318.2021.2012740

Sontag, S. (2009). *Against interpretation and other essays*. Penguin Books.

Stein, J. P., Messingschlager, T., Gnambs, T., Hutmacher, F., & Appel, M. (2024). Attitudes towards AI: Measurement and associations with personality. *Scientific Reports, 14*, Article 53335. https://doi.org/10.1038/s41598-024-53335-2

Tai, M. C. T. (2020). The impact of artificial intelligence on human society and bioethics. *Tzu Chi Medical Journal, 32*, 339–43. https://doi.org/10.4103/tcmj.tcmj_62_20

Turing, A. M. (1950). Computing machinery and intelligence. *Mind, 59*(236), 433–460.

Vance, A., Elie-Dit-Cosaque, C. & Straub, W. (2008). Examining trust in information technology artifacts: The effects of system quality and culture. *Journal of Management Information Systems, 24*(4), 73–100. https://doi.org/10.2753/MIS0742-1222240403

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186–204. https://doi.org/10.1287/mnsc.46.2.186.11926

Vergès, P. (1994). Approche du noyau central : propriétés quantitatives et structurales [Central core approach: Quantitative and structural properties]. In C. Guimelli (Ed.), *Structures et transformations des représentations sociales* (pp. 233–253). Delachaux et Niestlé.

Wiener, N. (1948). *Cybernetics or Control and Communication in the animal and the machine*. MIT Press.

Zacklad, M. (2018). *Intelligence artificielle : Représentations et impacts sociétaux* [Artificial intelligence: Social representations and societal impacts] [Research report]. Conservatoire national des arts et métiers. https://halshs.archives-ouvertes.fr/halshs-02937255

Zhang, B. & Dafoe, A. (2019). *Artificial Intelligence: American Attitudes and Trends*. University of Oxford. https://doi.org/10.2139/ssrn.3312874

ΨΥΧΟΛΟΓΙΑ | PSYCHOLOGY 30(2), 469-487

Ελληνική Ψυχολογική Εταιρεία
Hellenic Psychological Society

ΕΜΠΕΙΡΙΚΗ ΕΡΓΑΣΙΑ | RESEARCH PAPER

Μια δομική προσέγγιση για τις κοινωνικές αναπαραστάσεις για την Τεχνητή Νοημοσύνη. Εξανθρωπίζουν τα ρομπότ την τεχνολογία;

Πατρίτσια ΓΕΡΑΚΟΠΟΥΛΟΥ[1], Νικόλας ΧΡΗΣΤΑΚΗΣ[2]

[1] Τμήμα Ψηφιακών Μέσων & Επικοινωνίας/Σχολή Επιστήμης της Πληροφορίας & της Πληροφορικής, Ιόνιο Πανεπιστήμιο
[2] Τμήμα Επικοινωνίας & ΜΜΕ/Σχολή Οικονομικών & Πολιτικών Επιστημών, Εθνικό & Καποδιστριακό Πανεπιστήμιο Αθηνών

KEYWORDS IN GREEK

Τεχνητή Νοημοσύνη
Κοινωνικές Αναπαραστάσεις
Κοινωνική κατασκευή της τεχνολογίας
Επιστημονικό φαντασιακό
Ρομπότ

CORRESPONDENCE

Πατρίτσια Γερακοπούλου
Ιόνιο Πανεπιστήμιο
Τμήμα Ψηφιακών Μέσων και Επικοινωνίας
Λεωφόρου Αντώνη Τρίτση 1
Αργοστόλι, Κεφαλονιά, ΤΚ 28100
patgeraki@ionio.gr

ABSTRACT IN GREEK

Άρρηκτα συνδεδεμένη με μια «τεχνολογική μυθολογία» που σχετίζεται με προμηθεϊκές υποσχέσεις και απειλές, η Τεχνητή Νοημοσύνη προσελκύει το αυξανόμενο ενδιαφέρον τόσο των μέσων ενημέρωσης, όσο και του κοινού. Σκοπός της παρούσας ερευνητικής εργασίας είναι η διερεύνηση του περιεχομένου και της δομής των κοινωνικών αναπαραστάσεων της τεχνητής νοημοσύνης που έχουν οι φοιτητές των έξι τμημάτων Επικοινωνίας και ΜΜΕ στην Ελλάδα (N=249) μέσα από τη συλλογή πρωτογενών ποιοτικών δεδομένων που παράχθηκαν με τη χρήση ενός ερωτηματολογίου που περιείχε μόνο δύο ανοιχτές ερωτήσεις ελεύθερου συνειρμού. Στην πρώτη ερώτηση, οι συμμετέχοντες κλήθηκαν να γράψουν τρεις έως πέντε λέξεις που έρχονται άμεσα στο νου αναφορικά με τον όρο «Τεχνητή Νοημοσύνη». Σε μια δεύτερη ερώτηση τους ζητήθηκε να σχολιάσουν περαιτέρω αυτές τις λέξεις. Τα δεδομένα που συλλέχθηκαν ταξινομήθηκαν σε έξι μεγάλες θεματικές κατηγορίες, και δη Τεχνολογία, Μέλλον, Απειλές, Χρήσεις, Ρομπότ και Άνθρωπος, πριν αναλυθούν περαιτέρω στη βάση των κριτηρίων της συχνότητας και της δυναμικότητας, παράγοντας το «τετράγωνο της κοινωνικής αναπαράστασης της ΤΝ», που αποτελείται από τον κεντρικό πυρήνα (Τεχνολογία & Μέλλον), τα περιφερειακά στοιχεία (Απειλές & Χρήσεις) και μια «γκρίζα» περιοχή (Ρομπότ & Άνθρωπος). Η ερμηνεία και η συζήτηση των αποτελεσμάτων οδηγεί στο κύριο συμπέρασμα ότι το στοιχείο αναπαράστασης του Ρομπότ λειτουργεί ως ενδιάμεσος μεταξύ Ανθρώπου και Τεχνητής Νοημοσύνης, αποδίδοντας πιο οικεία χαρακτηριστικά στο κατά τα άλλα ασαφές και διφορούμενα αντιληπτό (Απειλές και Χρήσεις) τεχνολογικό αντικείμενο.

https://doi.org/10.12681/psy_hps.43979

Ψυχολογία: Το περιοδικό της Ελληνικής Ψυχολογικής Εταιρείας
https://ejournals.epublishing.ekt.gr/index.php/psychology