Rational Animals

Seven essays on reason's fields of application

edited by

Giorgio Grimaldi
Marialuisa Parise

Isonomia *Epistemologica*

Isonomia – Epistemologica

Volume 10

RATIONAL ANIMALS

SEVEN ESSAYS ON REASON'S FIELDS OF APPLICATION

Volume 1
Il realismo scientifico di Evandro Agazzi
Mario Alai (ed.)

Volume 2
Complessità e riduzionismo
Vincenzo Fano, Enrico Giannetto, Giulia Giannini, Pierluigi Graziani (eds.)

Volume 3
Oltre la fisica normale
Isabella Tassani (ed.)

Volume 4
Mettere a fuoco il mondo
Elena Casetta, Valeria Giardino (eds.)

Volume 5
Metaphor and Argumentation
Francesca Ervas, Massimo Sangoi (eds.)

Volume 6
Forecasting the Future
Stefano Bordoni, Sara Matera (eds.)

Volume 7
Teaching and Learning Mathematics
Laura Branchetti (ed.)

Volume 8
Animali razionali
Pierluigi Graziani, Giorgio Grimaldi, Massimo Sangoi (eds.)

Volume 9
Reasoning, Metaphor and Science
Flavia Marcacci, Maria Grazia Rossi (eds.)

Volume 10
Rational animals
Giorgio Grimaldi, Marialuisa Parise (eds.)

ISONOMIA - Epistemologica Series Editor
Gino Tarozzi gino.tarozzi@uniurb.it

RATIONAL ANIMALS

SEVEN ESSAYS ON REASON'S FIELDS OF APPLICATION

Edited by

Giorgio Grimaldi
Marialuisa Parise

ISSN 2037-4348

Scientific Director: Gino Tarozzi
Managing Director: Pierluigi Graziani
Department of Foundation of Sciences
P.za della Repubblica, 13 – 61029 Urbino (PU)

http://isonomia.uniurb.it/

Design by massimosangoi@gmail.com

Cover image by Luca Di Battista.

Table of contents

Introduction

Giorgio Grimaldi
University of Urbino Carlo Bo
giorgio.grimaldi@yahoo.it

Marialuisa Parise
University of Cassino and Southern Lazio
marialuisa.parise@unicas.it

The volume you are reading – *Rational animals. Seven essays on Reason's fields of application* – is the ideal and concrete sequel of another special issue of "Isonomia": *Animali razionali. Studi sui confini e sulle possibilità della razionalità*, edited by Pierluigi Graziani, Giorgio Grimaldi and Massimo Sangoi. The aim – then and today – is to overcome the barriers among different ways of thinking the *status*, the role, the function and the tools of philosophy. These barriers first of all preclude a chance and a goal: free philosophical research. Therefore, we have to remove them. This does not mean that everyone is invited to leave his/her own way of conceiving philosophical thought in view of *one* way of thought. These volumes intend to let everyone think in his/her own way *with* the ways of others. It is a collective process in which everyone carries forward his/her own approach and way of thought. He/she carries it *to* the others, for a free, open and serene debate, in which the first rule is to show – not hide – the differences. We believe that this is the only way to build a community of free thought – *one* community, but with *many* differences. A community that defends the right to freely think and confront one other.

The theme of this issue is – once again – *Reason*, one of the most important things we have in common, but that is at the same time so enigmatic, complicated and problematic to think about. The authors of these seven

essays reflect on some of Reason's fields of application, that is, history of thought, economics, artificial intelligence, theory of argumentation, architecture, neuroscience and aesthetics.

Marialuisa Parise illustrates the Italian reception of Rousseau's works, in particular his *Discours sur le sciences et sur les arts* (1750). Abbot Ferdinando Orlandi (1795-1860) wrote *Apologia delle scienze e delle arti in confutazione del Discorso di Giangiacomo Rousseau* (1831) and *Elogio delle principali scoperte ossia trattato dell'eccellenza dell'umano ingegno* (1831) to oppose Rousseau's argument against science and the role of reason. Orlandi, through an apologetic form, defends reason and faith, that – according to him – are internally connected. So, as the title of Parise's essay declares, *Reason* is *vindicated by religion*.

Gian Italo Bischi, **Maria Gabriella Pediconi** and **Fabio Tramontana**'s paper aims to discuss the issue of rationality within economics, showing some meaningful steps marking a paradigm shift in economic modelling. More specifically, regarding the role of the real individuals and the economic agents, this paradigm shift is characterised by an interdisciplinary approach involving methodologies which include experimental economics, adaptive and evolutionary processes, complexity theory and psychology.

The new paradigm had (and has) a relevant role in the field of Artificial Intelligence. **Antonio Lieto**'s paper aptly enlightens this issue, analysing the impact of the notion of bounded rationality in AI. In particular, he argues in favour of the indirect role of this notion in the field of automated decision making, showing how it has determined the *heuristic approach* to problem solving in AI.

Despite this fact, it should be asked whether and how this new paradigm relates to the issue of the universality of reason.

Paolo Labinaz questions the universal range of rationality. He examines whether and how rationality can be said to be universal. Labinaz shows how the approach called "argumentative rationality" is a way for conceiving rationality that can help to understand its universal character and to overcome the difficulties encountered in this respect by evolutionary (or ecological) and dual-process theories.

Giorgio Grimaldi's essay is also connected with the question of universality of reason. He analyses Mies van der Rohe's writings, underlining the centrality of reason and rationality in the thought and in the architectural projects and works of Mies. Through an in-depth analysis, Mies' thought reveals a hidden path: the thought of Hegel, and the tracks of his philosophy of history.

Following the silver thread of working of depicted objects, **Gabriele Ferretti** and **Silvano Zipoli Caiani**'s paper examines action possibility connected to them. They aim to show that depicted objects can recall action possibilities and how this phenomenon is perfectly coherent with the rationality of human beings. According to classical view, the action possibility depends on the fact that you can really act on the object that represents the action possibility as having. Hence an agent can not intend to act on the target that he/she believes are depicted objects, without being guilty of irrationality. The authors show how empirical evidence form visual neuroscience can endorse a different point of view.

But, rationality has a crucial role not only with regard to sight, but also with hearing. So, in closing, **Alessandro Giovannucci** shows the complex relationship among sound, reason and body, as sound interacts not only with reason, but also with the materiality of the body. Giovannucci discusses the works of Derrida, Nancy and Erlmann to aptly approach the question.

Reason vindicated by Religion
The Writings of Abbot Ferdinando Orlandi against Jean Jacques Rousseau

Marialuisa Parise
University of Cassino and Southern Lazio
marialuisa.parise@unicas.it

1. Introduction

According to Ernst Cassirer, the essence of the Enlightenment resides in the power of Reason: "Reason becomes a unitary and central point: it becomes the expression of all that it longs for, craves, wants and does".[1]

In his first work, *Discours sur les sciences et les arts*[2], Jean Jacques Rousseau opposed a central theme of the Enlightenment, the concept of the progress of man as the progress of his reason; for the Genevan philosopher science and the arts were instead the fruit of vanity and pride, the cause of man's departure from a primitive state of happy ignorance, and of his physical and moral corruption. The process of civilization was no less than an estrangement from the state of nature, and a headlong rush into vice.

The genesis of the *Discours* is well-known – it was conceived following the so-called "crisis" or "illumination" of Vincennes. In October 1749, while traveling along the road from Paris to the castle of Vincennes where his friend Diderot was imprisoned, Rousseau was leafing through the copy of the *Mercure de France* when his gaze fell on the announcement for next year's essay prize at the Academy of Dijon[3]; thus the *Discours sur les sciences et les arts* was born. According to Paolo Casini this was the debut of his entire political "system".[4] This event is famously described as a

[1] Cassirer (1932: 21).
[2] Rousseau (1750).
[3] See «Mercure de France», October 1749, pp. 153-155. See Rousseau (1964b: 1237-1238).
[4] See Casini (1974: 15).

Marialuisa Parise, "Reason vindicated by Religion. The Writings of Abbot Ferdinando Orlandi against Jean Jacques Rousseau", in Giorgio Grimaldi and Marialuisa Parise, Rational *Animals. Seven essays on Reason's fields of application*, pp. 11-30

fateful and irreversible turning point in his life and thought, which Rousseau mentions in several of his autobiographical writings, in the *Lettre a Malesherbes*[5], in Book VIII of the *Confessions*[6] and in other writings[7]: "A l'instant de cette lecture je vis un autre univers et je devins un autre homme"[8]. Composed in less than six months (October 1749-March 1750) at the sacrifice of many sleepless nights, drafting the *Discours* cost him dearly[9]. The awarding of the prize and the work's subsequent publication in 1750[10] provoked "a kind of revolution in Paris, establishing the fame of an unknown writer"[11]. As George Havens[12] pointed out, all this occurred at the same time that the *Prospectus* of the *Encyclopédie* emerged, which instead marked a glorious future for knowledge and progress.

The *Discours* immediately provoked a tremendous series of objections and refutations. Much of public opinion had been shocked by the type of response it received[13], although acknowledgments were not lacking: among them by the *Mercure de France* (December 1750), which saluted it as "one of the most beautiful discourses to have been rewarded in the academies"[14]. Responses to the refutations were written by Rousseau in less than a year (June 1751-April 1752); they were important for clarification of his theoretical positions.

2. A 19-century reaction to the *Discours* in Italy

In Italy the debate on Rousseau's first *Discours* arose only after the first Italian translation was published in the *Raccolta di opuscoli scientifici* in 1760[15]; Rousseau became "the mirror of all the ambiguous play of the various forces operating in Italy"[16].

---

[5] See Letter to Chrétien-Guillaume de Lamoignon de Malesherbes, Montmorency 12 January 1762, Rousseau (1969: 24-29).
[6] See Rousseau (1959: 350-352).
[7] The twelfth *Dialogue* and the third *Promenade*.
[8] Rousseau (1959: 351).
[9] Rousseau (1959: 352); see Rousseau (1988: 946).
[10] The frontispiece: Genève, Barillot et fils 1750. In reality the edition was printed in Paris by the bookseller Pissot.
[11] Rossi (1972: XLIV).
[12] See Rousseau (1946: 1).
[13] See Rousseau (1997: 6).
[14] Quoted by: Rousseau (1970: 118).
[15] See Rousseau (1760); Rota Ghibaudi (1961: 47, 317).
[16] Garin (1962: 283).

In the translation of the *Discours* printed in Catania in 1812[17], the positions held by his annotator Agatino Longo (1791-1889)[18] were significant. In his preliminary remarks[19] Longo compared the *Prospectus* of the *Encyclopédie* to the *Discours* by the Genevan, in order to make it easier to see the difference and by their mutual contrast, "make the truth stand out"[20].

> One who reads the preliminary Discourse of the Encyclopedia remains dazzled, and sweetly enchanted by the brilliant prospect before him, inflamed with the desire to travel such a happy country [...] but on the other hand it is right that he be warned of the rocks that surround him, the pitfalls that are hidden there, the lewd and fallacious passages that exist within. It is true that Science, no less than the Arts, is in itself very precious, and deserves our attention, but it is no less true that it can create stumbling blocks for many who do not give themselves to it with the proper provisions of mind and of heart.[21]

Agatino Longo declared that he had decided to reproduce Rousseau's *Discorso sopra le scienze* for young people, who were more exposed to the danger of "stumbling", and despite the fact that he did not entirely approve of the Geneva philosopher's assertions nor fully agree with his positions, he found the work useful "to rouse young scholars applied to learning them through the fiery and touching eloquence of its author, to avoid the dangers that go inseparably united with them"[22]. In fact, it was necessary for young people to be cautious when approaching "the sanctuary of Science, so as not to be invaded by a vain enthusiasm, or by a profane fire"[23]. He thus shared with Rousseau the idea that science could be extremely dangerous, considering it the cause of leading astray simple folk "with a light and unwise spirit", so much so that "the study of Science is quite often a waste of time, how many risks that the common man does not heed, should be feared from the reading of Books! What dire consequences cannot be dragged from a false principle adopted for true [...]"[24].

[17] See Rousseau (1812). See Rota Ghibaudi (1961: 317).
[18] Longo, teacher of experimental physics and philosophy at the University of Catania from 1815 to 1876 and extremely prolific man of letters; see Campione (2006:12-13, 21).
[19] The preliminary remarks can be read in their entirety in Schiff (1908: 10-11).
[20] Schiff (1908: 10).
[21] Schiff (1908: 10).
[22] Schiff (1908: 11).
[23] Schiff (1908: 11).
[24] Schiff (1908: 10).

A different opinion was expressed by Abbot Ferdinando Orlandi (1795-1860)[25], author of the *Apologia delle scienze e belle arti in confutazione del Discorso di G. Giacomo Rousseau*[26] and of *Elogio delle principali scoperte ossia trattato dell'eccellenza dell'umano ingegno*[27]printed in Florence in 1831 at the Magheri typography. He too was convinced of the great damage to society caused by the publication of ungodly and obscene books, and that the authors of these works had ruined youth, but nevertheless believed that one should not confuse "abuse with science", using the metaphor of the bee and the snake taken from Metastasio -- both animals often sip from the same flower but with different consequences: becoming poison in the snake, and honey in the bee[28].

Rousseau's paradoxical principle to be demolished was the idea that progress in the arts and science went along with the corruption of morals in a cause-effect relationship[29]. It is therefore an error to proscribe the sciences because in this way reason is prevented "from fortifying itself against the preponderance of the passions"[30].

According to Garin, in the works of Abbot Orlandi the polemic against Rousseau, and in general against the entire Encyclopedia and D'Alembert, slanderer of knowledge, "is indissolubly connected with the exaltation of culture and its revival, while paradoxically, the Age of Illumination becomes the age of detractors of science"[31].

The *Apologia delle scienze e belle arti* opens with a letter dedicated to Monsignor Ugolino Carlini (1773-1847)[32], bishop of Cortona; seeking his approval and support, Orlandi sent him this refutation of Rousseau's sophisms against Science, calling him "an archimandrite of unbelievers"[33]. The monsignor was in fact an avowed defender of the "beautiful daughter of Zion", religion; he was famed for his learned homilies with which "he

[25] Ferdinando Orlandi, abbot and professor, simple priest without the care of souls, man of vast culture, member of numerous scientific Academies, Italian and foreign; president of the Accademia Valdarnese of Poggio from 1836 to 1839; see Bacci (1956: 155, 455-458); Tartaro (2014: 45); Tartaro (2018).
[26] Orlandi (1831a). See Felice (1987: 63).
[27] See Orlandi (1831b).
[28] Orlandi (1831a: 150).
[29] Orlandi (1831a: 152).
[30] Orlandi (1831a: 212).
[31] Garin (1962: 283).
[32] See Ritzler & Sefrin (1968: 165).
[33] Orlandi (1831a: 4).

frightened vice and fulminated irreligion", held in the Florentine and Fiesole dioceses[34].

The following is a preface entitled *Ai cerberi della letteratura*, in which Ferdinando Orlandi shows he is aware that this work would attract much criticism. He expects the accusation of being a presumptuous "subaltern writer" for having dared to compete with the great Genevan author; however, in his opinion the rightness of the cause should prevail over the inferiority of the speaker. As for the criticism of having resorted to too many quotations and literary plagiarism, he admits that he "took few steps without crutches"[35] and preferred to be a "good collector rather than a bad author"[36]; on the other hand, the readers would do better to welcome "the old truths, rather than [...] new paralogisms"[37].

He vindicates the importance of this work, which is part of the more general task of Catholic apologists -- defending the truth of religion and morals from a vast host of enemies. "If writers multiply who would debaseman to the class of brutes, it will be a good thing if those multiply who proclaim their excellence"[38]. The purpose of the work, very different from that of Rousseau, consists in "insinuating respect for the truths of the Faith"[39].

Instead, for Orlandi, Rousseau was motivated by his hunger for celebrity. The paradoxes he offered to readers were aimed at glorifying himself and his abilities: "every sacrifice seemed to him sweet in order to obtain it, dazing the world with the singularity of his opinions and the talent to maintain them"[40]. There were various refutations to his *Discours*; however, Orlandi maintained that no one had been able to "erase the impression of that famous discourse"[41] and that the apologies hitherto produced had been incomplete and defective. To make his thesis of the usefulness of science prevail, the abbot followed the method of opposing "to the sophisms of Rousseau, metaphysical reasoning and rigorous demonstrations"[42].

To refute Rousseau's thesis, in which science is "a dangerous weapon placed in the hands of furious men, who will perpetually abuse it, brandish-

[34] Orlandi (1831a: 6).
[35] Orlandi (1831a: 4).
[36] Orlandi (1831a: 4).
[37] Orlandi (1831a: 4).
[38] Orlandi (1831a: 14).
[39] Orlandi (1831a: 13).
[40] Orlandi (1831a: 20).
[41] Orlandi (1831a: 20).
[42] Orlandi (1831a: 21-22).

ing it against their fellow men, or turning it against themselves"[43], science that is a daughter of vice and the cause of luxury, of the depravation of customs, of hypocrisy, the ruin of kingdoms and empires[44], Orlandi begins by demonstrating the advantages of science and then responds to these accusations by countering the slanderers of the enlightened with "the picture of ignorance". One example (among many) is the development of jurisprudence thanks to the enlightened, moving away from the legislation of barbarians who sought "the criterion of truth in the weaving of nerves"[45]. When ignorance predominated, superstition and fanaticism led them to make horrendous sacrifices to the gods, whereas history has shown how in civilized times even war, always cruel, sets limits on ferocity.

Finally, Orlandi offers an example of those Asian regions "subject to the barbarous yoke of Mohammad", where the sciences are forbidden; in fact, the Mohammedan religion "makes tyranny sacred by placing the throne above the altar [...], it favors universal stupidity by prohibiting the culture of the intellect"[46]. For him it was a matter of unfortunate peoples, of whom it is certainly impossible to envy their ignorance, as they cannot harbor any hope of "great revolutions", because "slavery is established forever"[47]. The author concludes by stating that the arts are necessary for the body as they allow it to escape from a state of nudity and savagery, thus achieving satisfaction of its needs, whereas science is necessary for the spirit, offering it the means to fulfill its duties, without which it would be "a slave to softness and vice".

Philosophy is defined as "a solid science, drawn from the pure sources of reason and experience"[48], and is praised as a doctrine that allows man to know himself and the world around him and, above all, to "go back to the author of his existence"[49]. After describing the process of philosophy, he refers to the classic distinction between healthy and unhealthy philosophy; he defines healthy philosophy as that science which "dispels all doubts, fixes all uncertainties, and provides for all human needs"[50].

---

[43] Orlandi (1831a: 26).
[44] Orlandi (1831a: 26-27).
[45] The reference is to Cesare Beccaria. Beccaria (1984: 62): "quasi che il criterio di essa risieda nei muscoli e nelle fibre di un miserabile (as if its criterion lies in the muscles and fibers of a poor wretch)".
[46] Orlandi (1831a: 33).
[47] Orlandi (1831a: 33).
[48] Orlandi (1831a: 34).
[49] Orlandi (1831a: 34).
[50] Orlandi (1831a: 35).

For Orlandi, science has the merit of "illuminating man in his duties, of perfecting his reason, of elevating him to the knowledge of God"[51]; for the author, the universe is a manifestation of its Creator, so all things are the work of his hand and bear the imprint of the divine architect. Following a path already trodden by other exponents of Catholic apologetics, he maintains that "a stead fast philosopher is never an irreligious spirit"[52].

The cornerstone of a certain apologetic "tradition" of the 18th century is in fact the concept that in-depth study does not distance one but instead draws one nearer to God. The thesis had found an effective synthesis in Bacon's maxim: "a little philosophy inclineth man's mind to atheism; but depth in philosophy bringeth men's minds about to religion"[53], quoted by all these apologists in their works[54], including Orlandi who writes: "Great studies, says Bacon, lead to the docility of faith; impiety is the vice of weak and limited spirits"[55]. Elsewhere he invokes it again, "Let us seek Bacon's maxim, that is, that superficial science leads us away from God, while a more extensive science leads us back to the humility of faith"[56]. In this work he also cites another motto of the English philosopher, reflecting his great esteem: "So rightly then, Bacon wrote that natural philosophy is the most effective remedy against superstition, and a sincere nourishment of faith"[57].

Among the prominent exponents of this apologetic was Giacinto Sigismondo Gerdil (1718-1802)[58]. He maintained that the enemies of religion are those who because of their incapacity are also enemies of science and its advancement: "Enemies are those untalented who, because of their failure, are also enemies of science and of the arts"[59].

In the *Apologia* Orlandi then draws up a long inventory of the arts and sciences (astronomy, physics, mathematics, medicine, theology, etc.) showing how each of these, if studied in depth, coincides with Christian truth and

---

[51] Orlandi (1831a: 36).

[52] Orlandi (1831a: 36). Another author, representative of this apologetic is monsignor Francesco Colangelo (1769-1836), among his works on the subject: see Colangelo (1804).

[53] The expression, made famous in the first book of *De augmentis scientiarum* in 1623, is present in Francis Bacon's works from 1597 and is also found in *Advancement of learning* (1605) and in the essay *Of Atheisme* (1612).

[54] On the subject: see Parise (2012: 433-474).

[55] Orlandi (1831a: 160). In a note he quotes the original version in Latin: "leves gustus in philosophia movere fortasse ad atheismum, sed pleniores haustus ad religionem reducere".

[56] Orlandi (1831a: 215).

[57] Orlandi (1831a: 69).

[58] On Gerdil: see Borghero (2001: 31-61); Borghero (2011a: 23-41) and Borghero (2011b: 238-243).

[59] See Gerdil (1755: 51).

with God, providing evidence of his existence. For Orlandi the Christian religion is not founded on ignorance but in the principles of reason: "man must be religious by principles, not by weakness, nor by a blind veneration"[60].

In *Elogio delle principali scoperte*, conceived by the author as an appendix to the *Apologia delle scienze e delle arti*, there is the celebration of man, considered the noblest creature of all, the only immortal:

> a vapor, a single atom kills him: but even if the whole universe buries him by its weight under its mass, he is nobler than that which kills him; while he knows that he perishes, the universe that survives him knows nothing. He dies but the physical and the moral do not perish together and the most beautiful part of him escapes from her arrow.[61]

And he continues maintaining that man's greatest privilege is the possession of reason:

> Thus man is great, but his greatness consists principally in reason. Alone among all animals does he enjoy the gift of words: he dresses his ideas in arbitrary signs, and with this admirable prerogative he establishes a bond between them, which forms a priceless treasure of knowledge in his imagination. For he communicates his thoughts, he perfects all his faculties, he takes possession of all the arts, and of all the sciences, and in this way all of nature is subjected to him.[62]

The aim of the work is to vindicate reason, humiliated by the materialists and skeptics who came to assign man to the class of brutes; it was then for Orlandi that "reason was therefore called a path for getting lost, the intellectual faculties became attributes of matter; man was called a plant, a self-propelled machine".

> The purpose of this treatise is certainly not to stir up the pride and presumption of man, whose legacy according to divine language is lies and sin, but to vindicate reason debased by the materialists, whose forerunners are Lucretius, Hobbes, Tolland, Spinoza, La Metrie, as well as all the skeptics, among whom Sesto Empirico, Montaigne, and Bayle stand out, and finally by all those whose declamations about the weakness of the human spirit came to destroy the foundation of all certainty in human judgment, until they became the lawyers of paradox [...].[63]

---

[60] Orlandi (1831a: 40).
[61] Orlandi (1831b: 9).
[62] Orlandi (1831b: 9-10).
[63] Orlandi (1831b:102 n4).

The work was intended to show the greatness of the human intellect "since the most beautiful praise of genius is its most sublime productions"[64]. Praise for the dignity of the human intellect was in fact an indirect response to the bleakest and most influential maxim of false philosophy[65].

3. Selected reviews of Orlandi's works:

3.1. *Antologia* - Giuseppe Montanelli's first review

Orlandi's works did not escape the attention of the newspapers of the time; a review was published by Giuseppe Montanelli (1813-1862)[66], in the *Antologia* of Giovan Pietro Vieusseux (1779-1863)[67]. Montanelli had met Vieusseux in September 1831 and "undertook from the beginning to assault him with an interminable series of proposals for work. By the end of the year he had already had his review of Orlandi accepted"[68] and it was his very first article accepted for publication. Vieusseux, as editor of the journal, tested the skills of the recruits through the review "test bed" and then promoted them, if they deserved it, to the rank of columnists; Montanelli, barely eighteen and just graduated with a Law degree from the University of Pisa, also underwent this process[69].

The review began:

> At a time when the healing effects of the sciences and the arts are celebrated and recognized everywhere, and there is no doubt about the excellence of the human intellect, a book called the *Apologia delle scienze e delle arti, -Elogio delle principali scoperte ec*.- cannot help but excite the curiosity of philosophers.

Although he acknowledged Orlandi's merit in having written learned and erudite works in which there was a conspicuous use of classical, ancient and modern sources, thus works useful to the Republic of Letters, he reproached him for a lack of method in dealing with the proposed subject and for not having given a philosophical form to his works: "much greater would have been their merit, if to the erudition and the doctrine he had

---

[64] Orlandi (1831b: 11).
[65] See Orlandi (1831b: 59).
[66] On Giuseppe Montanelli, see Bagnoli (2011: 835-842).
[67] See Vieusseux (1979); Montanelli (1831: 29-32). The review is signed M. G. In Vieusseux (1863) this review is recorded under the name Orlandi, previously noted in Bagnoli (1995: 93n).
[68] Pitocco (1972: 137).
[69] Rosselli (1946: 126).

added a more methodical disposition of the subject, a more severe analysis in the partial investigations, an elocution in short less rhetorical and more philosophical"[70]; a concept on which he had not failed to insist: "we would have liked a more philosophical form in his book, more suited to the current needs of the sciences"[71].

As for Orlandi's first work, the *Apologia delle scienze e delle arti, nuova confutazione di Rousseau*, in an attempt to illustrate the advantages of science and the arts, after briefly discussing philosophy in general Orlandi dealt with each science and art individually:

> How much, in our opinion, would have better satisfied his assumption, if before discussing this or that science in particular, he had examined science and art in general in their offices and in their relation to perfecting the individual and society! Then he would have found that science is necessary and indispensable, because it is a means of achieving happiness.[72]

On the other hand, the young Montanelli would have argued that man's happiness depends on his conforming to the order of nature and that this is made possible only through science, which allows man to know this order. Hence science, besides being useful and advantageous, becomes the manifestation of a real need for nature. The common purpose of the sciences consists in the preservation of the human race and in the improvement of society. Thus also the arts of beauty, although differing from the sciences in not being indispensable and necessary, contribute to the perfection of man and to the diffusion of truth, thus rendering their usefulness undeniable. "After these general principles the discourse on the sciences and the arts in particular would have acquired an entirely new and interesting aspect, since it would have presented a confirmation of established truths". Although the subject addressed was not new, the work would thus be new and interesting.

Another accusation made against the abbot was that he had examined the various branches of knowledge in a confused and disordered way, without following the classification now in use in the sciences.

Regarding the second tract on *Elogio delle principali scoperte* Montanelli criticized the author for not distinguishing between discoveries made due to some natural law and discoveries made due to the invention of new methods for understanding nature better, a distinction that would also have shown the superiority of the moderns over the ancients.

---

[70] Montanelli (1831: 32).
[71] Montanelli (1831: 29).
[72] Montanelli (1831: 30).

Montanelli had quickly dashed off his harsh review of Orlandi in one morning, on the wave of negative feeling he had experienced reading those works, and proposed it to Vieusseux: "I so disliked the things discussed by this author that I could not help but write down some ideas as literary news, which I would like to have included in your accredited journal"[73].

From another letter replying to Vieusseux, we know that the editor, having accepted the text, had asked him to make some corrections:

> With great pleasure, I learned from your very kind letter of the 20th of this month that my article met with your approval [...]. I will gladly modify those some what harsh expressions that escaped me in the impetus of composition. It was a morning's work and I did not have time to reflect upon it. But it is very fair and reasonable that in criticism one always retains that dignity, which is appropriate to the current state of knowledge and is the distinctive characteristic of true wisdom. Please tell me exactly which passages you want changed in the copy to be sent.[74]

However, the issue of the journal, dated December 1831, only appeared in the middle of the following February,[75] when by now Montanelli had embraced Saint-Simonianism. The review contains some typical themes of this current of thought: the critique of Rousseau who supported the thesis of "the uselessness and damage caused by the sciences and the arts", a very common criticism at the time, but particularly deep-rooted among the Saint-Simonians, and the pragmatic conception of art as useful for the improvement of man, for the diffusion of truth and for the exaltation of science, etc[76].

However, Montanelli must have partly changed his mind, because in a letter to Vieusseux dated January 12, 1832, he informed him that in correcting them he had mitigated certain expressions in his harsh review:

---

[73] Letter from Giuseppe Montanelli to Giovan Pietro Vieusseux, Pisa, December 16, 1831, in the National Central Library of Florence [henceforth BNCF], Carte Vieusseux, cass. 71, ins. 56 quoted by: Bagnoli (1995: 95); see Montanelli (1993: 365); Luseroni (1996: 51); Rosselli (1946: 127).

[74] Letter from Giuseppe Montanelli to Giovan Pietro Vieusseux Pisa on December 21, 1831 in BNCF, Carte Vieusseux, cass. 71, ins. 57: Bagnoli (1995: 95); see Montanelli (1993: 366); Rosselli (1946: 128).

[75] See Letter from Giovan Pietro Vieusseux to Giuseppe Montanelli, February 11, 1832 in which he announced that next Monday his article on Orlandi would appear in the *Antologia*, BNCF, Carte Vieusseux, cass. 71, ins. 59; Luseroni (1996: 51n).

[76] For Montanelli's Saint-Simonian collaboration with *Antologia* and especially for the presence of Saint-Simonian themes in this review, see Pitocco (1972: 98, 142).

> Perhaps you will marvel at the many corrections in the printing of my little article, which I send back to you by mail. But these corrections had their reasons. I wanted to mitigate the pungent spirit with which I had criticized poor Orlandi, since I knew he was a young man full of good intentions and love of study. These good inclinations deserved consideration. Moreover, I have been told that he is very much persecuted by the priests of his city, who seek in every way to attack him both in his conduct and in his scientific productions. Also this reason made me treat him with more regard, without defrauding the love of truth and of science in any way.[77]

We know from the same letter that the draft of this "little article" had been given to Professor Gaetano Cioni (1760-1851)[78] to read and he said he was very satisfied with it.

However, Vieusseux did not agree with Montanelli's excessive retouching of his criticisms: "You did well to mitigate some expressions that were a bit pungent, but it would have been better, perhaps, to mitigate a little less. We will manage it better another time"[79].

In 1833, an answer from Orlandi to these criticisms was printed at Magheri's typography[80].

3.2. *La Voce della Ragione* - Monaldo Leopardi's opinion

There is a review of Orlandi's works among the *Annunzi bibliografici* of 1834 of *La voce della Ragione* by Monaldo Leopardi (1776-1847)[81]. A brief but effective excursus on the state and modifications sustained by literature and the sciences precedes his analysis of the works. Letters and the sciences, once considered useful for removing man from error and guiding him toward justice and the truth (the reason why the thesis of Rousseau "who was shouting, remain in ignorance if you want to live happily"[82] had provoked a great scandal), by now had fallen into the clutches of false philosophy.

---

[77] Letter from Giuseppe Montanelli to Giovan Pietro Vieusseux, Pisa January 12, 1832 in BNCF, Carte Vieusseux, cass. 71, ins. 59: Bagnoli (1995: 97-98); see Montanelli (1993: 367-368); Rosselli (1946: 127n).

[78] On Gaetano Cioni, humanist man of letters with a strong interest in science, friend of Vieusseux, one of the founders and editors of the *Antologia*, see Giovanardi (1981: 685-686); Montanelli (1993: 366n, 368).

[79] Letter from Giovan Pietro Vieusseux to Giuseppe Montanelli, February 11, 1832: Rosselli (1946: 127n).

[80] See Orlandi (1833).

[81] On Monaldo Leopardi, father of Giacomo: Del Corno (2005: 654-657). See Leopardi, M. (1834: 57-60); Fantoni (2004: 247-248n).

[82] Leopardi, M. (1834: 57).

Leopardi noted that they had "become a fuse to light the fire of passion, a lever to lift peoples against princes, and a thunderbolt to target the altar and the church"[83]. Hence the different warning of the learned men who once encouraged people to study and improve one's mind and now are forced to warn them, and to suggest moderation:

> It happens to the sciences as happens with remedies and medicines that when they are seen to be administered by honest and worthy professors, one says to the sick, run to medicate and heal yourself, but when they are doled out by charlatans, you cry to the idiot populace, beware of those who are only capable of poisoning you.[84]

For Leopardi the present time was dominated by false philosophers; the aims they attempt to conceal are deadly: "they intend to subjugate in a murky and stormy sea of pedanticism, of sophistry, of ideology and of Cormentalism, thrones, princes, society, religion, altars, and the church"[85]. Orlandi's intention is different:

> In his most learned and very eloquent apology of the sciences, he recommended what is good for us, without undermining with what is harmful, he has distinguished the true from the false, the use of abuse, and above all, in a time when everything is useful for philosophers for deceiving up right reason and targeting the Christian faith, he has defended the sciences with the words of a wise man, and with the heart of a true Christian.[86]

Leopardi mainly dwells on the *Risposta* to criticism of the *Antologia*, where Orlandi philosophically addresses the choice of the term *Natura*, a rhetorical rather than a philosophical word substituted for the *Autore dell'universo,* and repudiates the thesis proposed by the young Montanelli in which science is a need of nature -- if anything it is a need of the social state. For Orlandi science is not necessary in order to know the *jus costituito*, to which men must conform their actions, but it is necessary to know the *jus costituendo*, i.e., the best laws and the wisest institutions that must regulate the political and civil order of society. For Orlandi it is not necessary for everyone to devote themselves to study, but only those who direct the social order.

Thus, for Leopardi the philosophical part is present and effective and the criticisms made by the *Antologia* are to be attributed to an ideological

[83] Leopardi, M. (1834: 58).
[84] Leopardi, M. (1834: 58).
[85] Leopardi, M. (1834: 58).
[86] Leopardi, M. (1834: 58).

principle: "These clear and firm principles are enough to form an idea of the philosophical and moral part of the work, to whose praise there is the no small addition that it was rigorously censured by the Florence *Antologia*, which however was wrapped up in its false concepts; all know why it censured, why it praised"[87].

Therefore he advised his associates to buy the volumes:

> As for the literary and scientific part, we do not undertake to praise it because we would linger above every line, and the beauties of this great work cannot be portrayed, but one must see them. And because these excellent books, nobly printed, are released at a very good price, we recommend and encourage our members to purchase them to be able to judge for themselves.[88]

And backing his extremely positive judgment about the works, he referred to the review written by Pietro Bagnoli (1767-1847)[89].

3.3. *Nuovo giornale de' letterati* - Review by Pietro Bagnoli

In 1831, an exhaustive review of Orlandi's two works was published in the *Nuovo giornale de' letterati*, written by Bagnoli[90].

Rousseau is portrayed as the "Coryphaeus of the absurd"; his *Discours* condemning the arts and sciences was in reality the result of the state of perfection reached by them in his time, so it was therefore a singular contradiction. The judgment of disapproval is clear: "a contradiction and extravagance of the human spirit, when it deviates due to vanity or superficiality or malice from the straight path, toward the absurd"[91].

Thus the refutation of this work was considered a work of extreme importance and utility, "since the sciences have a direct influence on public morals and on the good of civil society" and the superior abilities shown by Orlandi in fighting against this absurdity were recognized, and against "a tremendous adversary, as everyone considers the Ginevran" due to his most skillful eloquence.

For Bagnoli, Orlandi had achieved his goal and soundly defeated the enemy; the description that he offers of Orlandi's way of proceeding had struck Leopardi who relates at the end of his announcement:

---

[87] Leopardi, M. (1834: 59).
[88] Leopardi, M. (1834: 59).
[89] On Pietro Bagnoli, see Carranza (1963: 264-266).
[90] See Bagnoli (1831: 130-136).
[91] Bagnoli (1831: 130).

> His way is not a cold demonstration that leaves the field open for retaliation, never to be concluded convincingly, but it is a trend in the vein of vehement polemic eloquence, concept by concept, maxim by maxim, doctrine, authority, erudition, which succeed each other like the waves of a rushing river, which carry the reader, and do not leave him time to think how the adverse cause can be sustained. The fruit, we see, of universal, tireless reading in this genre.[92]

As for his second book the *Elogio delle principali scoperte,* which proposed to deal with the vast subject of the greatest inventions of the human intellect and therefore of the progress of knowledge, its design seemed incomplete; although it was only dealt with here in an essay, it would have required "the expertise of a society of scientists".

3.4. *Biblioteca Italiana*

Finally, also in 1831, an anonymous review of his works came out in the *Biblioteca Italiana*[93]*;* however, it is of less interest than those described so far, consisting for the most part of a mere literary notice. The author claims that the paradox of the Genevan philosopher, according to which the restoration of science and the arts contributed to corrupting rather than improving morality, is one of his most famous paradoxes, and that many authors, even among his contemporaries and including numerous Italians, wrote works refuting this paradox. Nevertheless, it recognizes the value of Orlandi's undertaking. As a good apologist, worried about the education of young people who are easily influenced by dangerous doctrines, he had the ability to face the problem in the best way:

> But if it is true, and it is in fact very true, that it is always a good undertaking to repudiate pernicious doctrines, especially those which can harm the young, who by nature are lovers of novelty and shy from study and fatigue; we should thank the illustrious Abbot Orlandi, above all now in our times, when all sorts of absurd and extravagant doctrines have been brought back from oblivion. And all the more we must be grateful to him, as he is doing so with a wealth of eloquence and philosophy.[94]

---

[92] Bagnoli (1831: 132).
[93] See Anonimo (1831: 341-342); Felice (1987: 63).
[94] Anonimo (1831: 342).

References

Anonimo (1831), "Apologia delle scienze e belle arti e Elogio delle principali scoperte dell'abate Ferdinando Orlandi", in *Biblioteca italiana o sia giornale di letteratura, scienze ed arti*, LXIV, a. 16, pp. 341-342.

Bacci, D. (1956), *Terranuova Bracciolini nella sua storia*, Firenze, E. Ariani e «L'arte della stampa».

Bagnoli, P. (1831), "Apologia delle scienze e delle arti. - Elogio delle principali scoperte opere del sig. Ab. Ferdinando Orlandi", in *Nuovo giornale de' letterati*, XXII, n. 59, pp. 130-136.

Bagnoli, P. (1995), *La politica delle idee. Giovan Pietro Vieusseux e Giuseppe Montanelli nella Toscana preunitaria*, Firenze, Edizioni Polistampa.

Bagnoli, P. (2011), "Montanelli, Giuseppe", in *Dizionario biografico degli italiani*, LXXV, Roma, Istituto della Enciclopedia Treccani.

Beccaria, C. (1984) *Dei delitti e delle pene*, Francioni, G. (ed.), *Edizione nazionale delle opere di Cesare Beccaria*, Milano, Mediobanca, I.

Borghero, C. (2001), "Gerdil e i moderni: le strategie apologetiche di un anti-illuminista", in Lionello Sozzi *et al*. (eds), *Nuove ragioni dell'anti-illuminismo in Francia e in Italia*, Pisa/Genève, ETS and Slatkine.

Borghero, C. (2011a), "L'anti-Voltaire de Gerdil ou la stratégie apologétique d'un malebranchiste", in Borghero, *Les Cartésiens fàce à Newton. Philosophie, science et religion dans la première moitié du XVIII*e *siècle* (trans: T. Berni Canani), Turnhout, Brepols.

Borghero, C. (2011b), "Religionsapologetik", in *Die Philosophie des 18 Jahrunderts. Band 3: Italien*, ed. J. Rohbeck e W. Rother, Basel, Schwabe Verlag.

Campione, F. P. (2006), *La nascita dell'estetica in Sicilia*, Palermo, Centro internazionale studi di Estetica.

Carranza, N. (1963), "Bagnoli, Pietro" in *Dizionario biografico degli italiani*, V, Roma, Istituto della Enciclopedia Treccani.

Casini, P. (1974), *Introduzione a Rousseau,* Roma-Bari, Laterza.

Cassirer, E. (1932), *Die Philosophie der Aufklarung*, Tübingen, Mohr (*La filosofia dell'Illuminismo*, trans. Ervino Pocar, Firenze, La Nuova Italia 1935).

Colangelo, F. (1804), *L'irreligiosa libertà di pensare nemica del progresso delle scienze*, Napoli, Vincenzo Orsino.

Del Corno, N. (2005), "Monaldo, Leopardi", in *Dizionario biografico degli italiani*, LXIV, Roma, Istituto della Enciclopedia Treccani.

Fantoni, N. (2004), *La voce della Ragione di Monaldo Leopardi*, Firenze, Società Editrice Fiorentina.

Felice, D. (1987), *Jean-Jacques Rousseau in Italia. Bibliografia (1816-1986). Prefazione di Alberto Postigliola*, Bologna, CLUEB.

Garin, E. (1962),"Il secondo centenario della pubblicazione del *Contratto sociale"* in *Giornale critico della filosofia italiana*, terza serie, XVI, a. XLI, fasc. II, pp. 282-283.

Gerdil, G. S. (1755) *Introduzione allo studio della religione*, Torino, Stamperia Reale.

Giovanardi, S. (1981), "Cioni, Gaetano", in *Dizionario biografico degli italiani*, XXV, Roma, Istituto della Enciclopedia Treccani.

Leopardi, M. (1834), *La Voce della Ragione. Giornale filosofico, teologico, politico, istorico, letterario*, X, n. 55, pp. 58-60.

Luseroni, G (1996), *Giuseppe Montanelli e il Risorgimento: la formazione e l'impegno civile e politico prima del '48*, Milano, Franco Angeli.

Montanelli, G. (1831), "Apologia delle scienze e delle arti. - Elogio delle principali scoperte. Opere del sig. Ab. Ferdinando Orlandi", in *Antologia*, XLIV, n. 132, pp. 29-32.

Montanelli, G. (1993), "Lettere di Giuseppe Montanelli a Giovan Pietro Viesseux (1831-1860)", ed. Paolo Bagnoli in *Nuova Antologia*, a. 128, vol. 570, fasc. 2187.

Orlandi, F. (1831a), *Apologia delle scienze e belle arti in confutazione del discorso di G. Giacomo Rousseau*, Firenze, Magheri.

Orlandi, F. (1831b), *Elogio delle principali scoperte ossia trattato dell'eccellenza dell'umano ingegno,* Firenze, Magheri.

Orlandi, F. (1833), *Risposta dell'abate Ferdinando Orlandi alla critica delle sue suddette due opere inserita* nell'Antologia *fiorentina,* Firenze, Magheri.

Parise, M. (2012), *Moderni contro moderni. L'uso di Bacon e di Galilei nell'apologetica cattolica tra Sette e Ottocento* in *Pensare la modernità*, ed. Giorgio Grimaldi, Villasanta, Limina Mentis.

Pitocco, F. (1972), *Utopia e riforma religiosa nel Risorgimento. Il sansimonismo nella cultura toscana,* Bari, Laterza.

Ritzler, R., & Sefrin, P. (1968), *Hierarchia Catholica medii et recentioris aevi*, Padova, Il messaggero di S. Antonio, Vol. VII (1800-1846).

Rosselli, N. (1946), "Frammento della incompiuta 'Vita di Giuseppe Montanelli'", in *Saggi sul Risorgimento e altri scritti*, Torino, Giulio Einaudi.

Rossi, P. (1972), *Nota sui testi e sulla biografia*, *Discorso sulle scienze e sulle arti*, traduzione di Rodolfo Mondolfo in *Opere di Jean Jacques Rousseau*, Firenze, Sansoni.

Rota Ghibaudi, S. (1961), *La fortuna di Rousseau in Italia (1750-1815)*, Torino, Giappichelli.

Rousseau, J. J. (1750), *Discours qui a remporté le prix à l'Academie de Dijon en l'année 1750 sur cette question proposée par la même Académie: Si le rétablissement des sciences et des arts a contribué à épurer les moeurs*, Genève, Barillot et fils.

Rousseau, J. J. (1760), *Discorso che ha riportato il premio dell'Accademia di Dijon,* in *Raccolta di opuscoli scientifici tradotti da diversi celebri autori francesi ed inglesi*, t. I, Ferrara, all'insegna del Sansone.

Rousseau, J. J. (1812), *Discorso di Giangiacomo Rousseau sopra le scienze e le arti, tradotto dall'abate Giovanni Sardo ed annotato dal dottor Agatino Longo,* in *Discorsi sopra le arti e le scienze*, Catania.

Rousseau, J. J. (1946), *Discours sur les sciences et les arts*, edition critique avec une introdution et un commentaire par George R. Havens, New York, Modern Language association of America, Kraus reprint 1966

Rousseau, J. J. (1959), *Les Confessions*, VIII, in *Oeuvres complètes*, I, Paris, Gallimard.

Rousseau, J. J. (1964a), *Discours sur les sciences et les arts* in *Oeuvres complètes*, III, Paris, Gallimard.

Rousseau, J. J. (1964b), *Du Contrat social, écrits politiques*, edition publiée sous la direction de B. Gagnebin et M. Raymond in *Oeuvres complètes*, III, Paris, Gallimard.

Rousseau, J. J. (1969), *Correspondance complète de J. J. Rousseau*, édition critique établie et annotée par R. A. Leigh, t. X, Geneve-Madison, Institut et Musee Voltaire - The university of Wisconsin press.

Rousseau, J. J. (1970), *Scritti politici*, ed. Paolo Alatri, Torino, UTET.

Rousseau, J. J. (1988), *Le Confessioni*, traduzione di V. Sottile Scaduto in *Opere*, ed. Paolo Rossi, Firenze, Sansoni.

Rousseau, J. J. (1997), *Discorsi*, traduzione di Rodolfo Mondolfo, ed. Luigi Luporini, Milano, BUR, 4ª ed. 2018.

Schiff, M. (1908), *Éditions et traductions italiennes des Oeuvres de Jean-Jacques Rousseau*, Paris, Librairies H. Champion.

Tartaro, G. (2014), *L'Accademia Valdarnese del Poggio. Due secoli di storia e di ricerca*, Firenze, Aska.

Tartaro, G. (2018), *Un laboratorio culturale nella provincia toscana. L'Accademia valdarnese del Poggio*, Firenze, Aska.

Vieusseux, G. P. (1863), *Indice generale alfabetico delle materie contenute nell'*Antologia*, giornale fiorentino diretto da G. P. Pietro Vieusseux: 18212-1832*, Firenze, Antonio Cecchi.

Vieusseux, G. P. (1979), *Vieusseux e il «Vieusseux»: storia e cronaca di un istituto di cultura e del suo fondatore. Catalogo delle mostre*. Firenze, Arti grafiche C. Mori.

Heuristics, Psychological Biases and (Bounded) Rationality in Economic Behaviour

Gian Italo Bischi
University of Urbino Carlo Bo
gian.bischi@uniurb.it

Maria Gabriella Pediconi
University of Urbino Carlo Bo
maria.pediconi@uniurb.it

Fabio Tramontana
Catholic University of Milan
fabio.tramontana@unicatt.it

1. Introduction

The mainstream modelling approach of economics in the 20th century is based on the paradigm of efficient markets and the optimizing behaviour of individuals and institutions, with rational economic agents assumed to be able to correctly forecast future states of the economic system where they operate, like physicists who know the equations governing real systems they are studying. The behaviour of such individuals is summarized by a single representative agent, sometimes denoted as *Homo oeconomicus*, a fictitious being endowed with unlimited computational ability and perfect information. In classical and neoclassical economics such features are stated according to an axiomatic approach, following methods similar to the ones adopted by 18th and 19th century physics that led to the notable results in mechanics and electromagnetism.

However, classical physics has the possibility to check its axioms through laboratory experiments, and the consequences of such axioms are often revealed to be fairly good descriptions of real phenomena, even giving

Gian Italo Bischi, Maria Gabriella Pediconi and Fabio Tramontana, "Heuristics, Psychological Biases and (Bounded) Rationality in Economic Behaviour", in Giorgio Grimaldi and Marialuisa Parise, Rational *Animals. Seven essays on Reason's fields of application*, pp. 31-51

rise to important predictions. On the other hand, the basic axioms of classical economics have not been checked through laboratory experiments, and many consequences of these axioms fail to explain some important features of economic systems. Hence the classical approach of rational and representative agents has been criticized on a number of grounds[1], and economic modelling is now witnessing a paradigm shift in methodology. This paradigm shift, characterized by an interdisciplinary approach involving methodologies which include experimental economics, adaptive and evolutionary processes, complexity theory and psychology[2], has been strengthened by the recent economic and financial crisis and the evident difficulties shown by the classical approach in managing it. At the same time, a growing interest has emerged in alternative approaches to economic agents' decision making, which allow for factors such as *bounded rationality*, *heterogeneity* of agents, social *interaction, learning*, in which the behaviour of individuals is governed by simpler "rules of thumb" (or "heuristics"), "trial and error" or even "imitation" mechanisms, often driven by sentiments and biases related to psychological issues. A new approach, known as *behavioural economics*, is now becoming more and more important.

Even the founding fathers of Economic Theory were well-aware of the role played by sentiments in shaping the behavior of economic agents. Adam Smith referred to them in his *Theory of Moral Sentiments* while J.M. Keynes called them "Animal Spirits". The so-called neoclassical synthesis, still considered the dominant paradigm in economic theory, pushed the psychological traits of the economic agent into the background in favor of a theoretical system founded on the assumptions of perfect rationality and homogeneity. This despite the fact that experimental proof of the descriptive limitations of such assumptions had been provided by several studies, from the paradoxes of Allais and Ellsberg to the works of Simon, Kahneman and Tversky. Such is the origin of behavioral economics, whose basic principles are derived from experimental results concerning the means through which real people take decisions. Psychologists have found a set of well defined mental shortcuts used by individuals to simplify complicated problems in order to obtain a fast decision. By knowing about these heuristics, it is possible to foresee possible deviations (*biases*) between the outcome of this decisional mechanism and the outcome identified by neoclassical theories of decision. For example, loss aversion and the importance of a reference point for making evaluations, have been proved to be well rooted in human be-

[1] See e.g. Simon (1957), Simon (1982), Kirman (1992), Kirman (1993).
[2] See e.g. Bischi (2012), Bischi (2014), Flabbi & Pediconi (2013).

ings, regardless of their culture. Moreover, there is a link between economics behaviour and happiness[3].

2. Behavioral economics and experimental approaches to social sciences

In truth, neoclassical economists do not always see human beings as robots, lacking in sentiments and always able to select the optimal strategy given the available information. They are, in general, perfectly aware of the fact that irrational behavior can be adopted and irrational choices can sometimes be made by anyone. Nevertheless, since the birth of economic theory, the assumptions of perfect rationality have been regularly adopted on the basis of what (as in physics) is known as the "*as if*" argument. This point of view, developed among others by the "Chicago economist" Milton Friedman, allows for a justification of such strong rationality assumption on the basis of the observation that real decision makers behave *on average* as if this assumption was correct, not unlike a snooker player who makes perfect shots as if aware of the snooker geometry, the forces impressed by the shots, and computational skillness (the same analogy used by Friedman itself)[4].

A prerequisite for the validity of the "as if" argument is that any behavior which is not perfectly rational is cancelled out by an opposite behaviour. In other words, irrational behavior is considered a random noise around perfectly rational behavior, thus leaving the average behavior unmodified which becomes the behavior of the *representative agent*.

---

[3] Kahneman (1999), Kahneman *et al*. (2006).

[4] Also Freud (1927) dealt with the "as if" argument and its role within the thinking processes. He emphasized the philosophical consistency of the "as if" argument in contrast to the subjective competence in distinguishing what is rational from what isn't. Competence, he claims, is typical of human beings since their childhood. According to Freud, 'as if' "asserts that our thought-activity includes a great number of hypotheses whose groundlessness and even absurdity we fully realize. They are called 'fictions', but for a variety of practical reasons we have to behave 'as if' we believed in these fictions. […] But I think the demand by the 'As if' argument is one that only a philosopher could put forward. A man whose thinking is not influenced by the artifices of philosophy will never be able to accept it; in such a man's view, the admission that something is absurd or contrary to reason leaves no more to be said. It cannot be expected of him that precisely in treating his most important interests he shall forgot the guarantees he requires for all his ordinary activities. I am reminded of one of my children who was distinguished at an early age by a peculiarly marked matter-of-factness. When the children were being told a fairy story and were listening to it with rapt attention, he would come up and ask: 'Is that a true story?' When he was told it was not, he would turn away with a look of disdain". (Freud, 1927: 28-29).

This is still the dominant (or mainstream) paradigm in Economics and Finance, even if there is a growing and flourishing strand of literature that questions the rationality assumption. This heterodox approach to Economics can be labelled "Behavioral Economics".

Even if nowadays the spread of Behavioral Economics seems fast and unstoppable (at least two Nobel laureates in the last 15 years can be considered Behavioral Economists), the birth and development of Behavioral Economics has really been a slow process evolving through several decades. Indeed, in order to become a credible alternative to the dominant approach to rationality in economic and financial decisions, several key discoveries have been necessary. First of all, it has been crucial to understand that rational behavior is regularly disregarded by real decision makers. Probably the most important breakthrough has been the identification of regular and predictable ways through which rational behavior is violated. Finally, the occurrence of a huge economic and financial event, such as the great economic and financial crisis of 2007-2008 with its far-reaching consequences, which was not forecasted and was impossible to control, or even explained by the axioms and methods of mainstream economics. Another aspect of contemporary research within the field of social sciences very close to economics, is the current research in affective neuroscience and neuroeconomics and its exploration of potential mechanisms affecting modulation along the computational process of subjective value and decisions. Current research suggests that specific neural-mechanisms mediate the relationship between affect and choice, underlining the process of decision making in psychological terms[5].

Before discussing these in the following section, we wish to underline that these events would have failed to bring Behavioral Economics to the actual level of relevance, without casting aside the erstwhile opinion that "*economists cannot perform controlled experiments like chemists or biologists because they can't easily control other important factors*"[6].

Once economists started to recognize that indeed Economics is also an experimental science , they became open to contamination with the methods and the principles of other social sciences, such as psychology and sociology. Although this process is still not completely developed, it is now not so unusual to find papers in eminent economic journals in which the authors have performed laboratory experiments. Without these, the work of Behavioral Economists would only be limited to some articles in conventional and

[5] Phelps *et al*.(2014).
[6] Samuelson & Nordhaus (1985: 8).

niche scientific journals of no actual relevance. As we shall see in the last section, the most promising and recent development of Behavioral Economics is probably its being used by important politicians and governments to improve the behaviour of citizens.

3. Apparent paradoxes of rationality and the failures of the axioms of normative neoclassical decision theory

In this Section we show some experimental results concerning the deviations from perfectly rational behavior in three different situations: risky decisions, intertemporal choices and probability updating.

In dealing with risky[7] decisions, the *Homo oeconomicus* of neoclassical economics is assumed to adopt a behavior that maximizes the so-called *expected* utility[8]. Roughly speaking, utility in Economics is a measure of the satisfaction experienced through the consumption of some good. In a risky environment, economic agents can only calculate, and try to maximize, a weighted average of the possible utilities associated to any achievable outcome, the weight being given by the probability of realization of the outcome. In formulas:

$$EU = \sum_i^n U_i p_i \quad i = 1...\text{n}$$

Where U_i is the utility of the outcome i that has a probability p_i to occur, and EU is the expected utility.

Do people, in making decisions, really maximize their expected utility? The answer is: "not always". Probably the first and most famous example of violations of such a normative decision theory is thanks to Maurice Allais. In an article, published by *Econometrica* in 1953, he described the results of a pair of experiments he performed. In each experiment participants have been asked to choose between two mutually exclusive alternatives and at least one of them was risky. In the first experiment the choice was between:

A) 1 Million dollars, for sure;

[7] In what follows we refer to risky decisions as defined by Knight (1921), that is, cases where the probability associated to any outcome is measurable. In this sense risk, is different from uncertainty.

[8] Von Neumann & Morgenstern (1947), Savage (1954).

B) 10% probability of winning 5 Million dollars and 89% probability of winning 1 Million dollars.

The majority of participants selected option A.
In the second experiment, the options were:

C) 11% probability of winning 1 Million dollars;
D) 10% probability of winning 5 Million dollars

In this case the most popular alternative was D.

It can be demonstrated that the two choices (A in the first experiment and D in the second one) are not consistent with the maximization of the expected utility.

In order to see this, let us consider that the choice of A means that the utility of A is higher than the expected utility of B:

$$U(1\text{M}) > 0.89*U(1\text{M})+0.1*U(5\text{M}) \qquad (1)$$

where we have assumed that the utility of winning nothing is zero.

To prefer D over C instead means that:

$$0.1*U(5\text{M}) > 0.11*U(1\text{M}) \qquad (2)$$

Now, in this last disequality we can substitute $0.11*U(1\text{M})$ with

$(1–0.89)*U(1) = U(1)–0.89*U(1)$ obtaining:

$$0.1*U(5\text{M}) > U(1\text{M}) –0.89*U(1\text{M})$$

that can be rearranged as:

$$0.89*U(1\text{M})+0.1*U(5\text{M}) > U(1\text{M}). \qquad (3)$$

Clearly (3) contradicts (1).

In 1961 Daniel Ellsberg gave another experimental demonstration of the violations of the expected utility theory. Furthermore, Ellsberg asked the participant to select between pairs of alternatives, but in his case not all the probabilities were known.

Here is a version of his experiments. Suppose there is an urn containing 90 balls. You know for sure that 30 balls are red and the other are either black or yellow but you don't know in what measure.

Now you must randomly extract a ball from the urn. What would you prefer:

A) To win 100 dollars if the ball is red;
B) To win 100 dollars if the ball is black.

The majority of participants usually choose A.
By contrast, in the case of the following question:

C) To win 100 dollars if the ball is red or yellow;
D) To win 100 dollars is the ball is black or yellow.

D is the most popular option.

Also in this case it is possible to prove that the combination between A over B and D over C is a violation of expected utility.

The first choice implies that:

$$P(R)*U(100) > P(B)*U(100) \qquad (4)$$

while the second one:

$$P(B)*U(100)+P(Y)*U(100) > P(R)*U(100)+P(Y)*U(100) \qquad (5)$$

that simplifies in:

$$P(B)*U(100) > P(R)*U(100) \qquad (6)$$

which is the opposite of (4).

Even if the experiments of Allais and Ellsberg, that were immediately labelled as *paradoxes*, are extremely important, they did not lead to any paradigm shift with respect to the expected utility theory. In order to obtain such a result it is not enough to find some cases where this theory fails. One should also propose an alternative, better theory. Allais and Ellsberg tried to give some explanations for their results, but a better theory was still far from being obtained.

We now move to intertemporal choices that concern decisions whose consequence are, at least partially, cast into the future. They are particularly

important in Economics and Finance because they involve, for instance, the interests in our bank account or how much to save for our pension after retirement.

The usual way of dealing with this kind of choice consists in calculating the present value of each alternative and choosing the highest one. It is necessary to discount utilities that will be experienced in the future because it is well established that people prefer to receive rewards sooner than later, so in order to be equivalent, the reward of tomorrow must be higher than today's.

The problem is that it is not irrelevant how this problem is mathematically solved. The mainstream approach dates back to 1937 when Samuelson proposed the exponential discounting (ED) model. According to this model, any future utility must be multiplied by a discounting factor that permits the calculation of its present value. If we denote by $\delta \in [0,1]$ the discount factor of tomorrow's utility, then δ^2 is the discount factor of an utility perceived the day after tomorrow, and so on. The present value U_0 of a sequence of utilities ($u_0,u_1,..,u_n$) perceived today, tomorrow,..., n days after today, is:

$$U_0 = u_0 + \sum_i \delta^i u_i \qquad (7)$$

So, if one needs to compare the following mutually exclusive alternatives:

a) To consume a good today, obtaining a certain utility u_0^a;

b) To save money today to receive a certain amount t years in the future, to which the utility $u_t^b > u_0^a$ corresponds.

Their present values must be computed:

$$U_a = u_0^a$$
$$U_b = \delta^t u_t^b$$

Respectively and if $U_a > U_b$ the decision of consuming prevails, otherwise one decides to save.

This appears quite simple and intuitive but the implications of the exponential discounting model are too demanding to be realistic.

Suppose you re-evaluate the two alternatives two years later (at t=2). In that case you must bring both the utilities at *t=2* instead of *t=0:*

$$\tilde{U}_a = u_0^a / \delta^2$$
$$\tilde{U}_b = \delta^{(t-2)} u_t^b$$

that can be obtained from today's values by dividing both by the same factor δ^2:

$$\tilde{U}_a = U_0^a / \delta^2$$
$$\tilde{U}_b = U_0^b / \delta^2$$

Therefore, if today consuming is better than saving, it should be better also two years later and, similarly, at any time in the future. This property is called *dynamic consistency* and implies that the *Homo oeconomicus* will never regret any decisions. But what about real human beings? Indeed, thousands of experiments support the idea that people frequently regret their economic decisions and they would change them if they could turn back time[9].

However, even in this case, without any alternative explanations of how people really discount the future, these would only remain interesting counterexamples of the exponential discounting model, but not a stumbling block for it.

Finally, let us consider how the *Homo oeconomicus* updates his/her beliefs about the probability of an event in the light of new evidence. This aspect is also important in economics because sometimes people acquire important information about an uncertain event only gradually, and accordingly they must update the probability they associate to its occurrence.

In this problem we move from some apriori probability *P(H)* about certain *H* hypothesis. In some cases *P(H)* may be based on prejudices. For example, let us consider a jury member who looks at the black skin color of a person charged for murder. Before the trial he/she may remain neutral and consider the hypothesis of guilt equal to 50% (*P(H)=0.5*) or maybe that because of some racial prejudice he/she considers the person guilty at 90% (*P(H)=0.9*). Suppose now that during the trial different evidence of the innocence of the accused person is brought forward. Let us simplify the scenario by assuming that the probability that a murderer is able to bring forward one of these items of exculpatory evidence is equal to $P(E/H) = 0.1$ while an innocent person would bring it forward for sure ($P(E/\bar{H})=1$).

[9] A survey of this (and other) experimental violation of the exponential discounting utility model can be found in Frederick *et al*. (2002).

The English mathematician Thomas Bayes, in 1763, asserted that a rational decision maker should update the prior probability according to the following formula:

$$P(H/E)=\frac{P(E/H)*P(H)}{P(E/H)*P(H)+P(E/\bar{H})*(1-P(H))} \qquad (8)$$

By applying this formula (known as *Bayes rule*) the probability of guilt for our potential jury members would be equal to around 9% and 47% for the neutral one and the one with a racial prejudice, respectively. Even more interesting is the fact that after four of these items of evidence the jury becomes convinced of the innocence of the accused person (their probabilities are undistinguishable). This feature of the Bayesian updating has been called *washing out of the priors* and it implies that after enough evidence our priors (i.e. our prejudices) will become irrelevant. After some time people should come to agree almost on everything. What a beautiful world it would be! But unfortunately this rarely happens in the real world.

Prejudices are extremely important in the real world, as proved through a famous experiment by Lord *at al*. (1979) who gave two groups of people the same set of documents concerning the death penalty. The first group was formed by individuals who declared themselves in favor of the death penalty, the second group by individuals against it. The documents included both positive and negative arguments concerning the death penalty. After reading such documents, the two groups became even more divided in their opinions. They were more convinced than ever about their idea on the death penalty.

Why do people deviate from the Bayesian updating and how do they actually update their probabilities?

To answer these questions, and the other ones concerning risky and intertemporal choices, different experiments performed by psychologists and economists have been necessary.

4. Predictable effects of bounded rationality, biases, and psychological issues

In the previous section we have seen that the normative theories of the rational decision making are often violated, thus limiting their descriptive properties.

Once experiments have been accepted as part of the economists' toolkit, their role has not been limited to performing some tests to check existing theories. They are used to find regularities in the way people deviate from the behavior of the neoclassical *Homo oeconomicus*.

One of the main achievements of the famous couple of psychologists Daniel Kahneman and Amos Tversky, concerns the discovery that people often simplify complicated tasks using simple shortcuts to make even the harder decisions. Through their experiments they have found that these shortcuts are always the same and can be classified and used to provide alternative behavioral theories. These rules of thumb are known as *heuristics*.

In an attempt to explain the failures of the expected utility theory, Kahneman and Tversky discovered that an individual's attitude towards risk is not absolute, but may depend on the *framing* adopted to present the alternatives. As in the well known experiment concerning the "Asian disease"[10], participants were asked to imagine the existence of an Asian disease that, if not stopped, was going to kill 600 people. To deal with the problem, the government could choose between two mutually exclusive programs, labelled A and B. By adopting program A 200 people would be saved, while with program B there is 1/3 probability that 600 people would be saved and 2/3 probability that no one would be saved.

72% of participants opted for the A program, revealing their risk aversion. The majority of participants seemed to be particularly afraid of the risk that no one would survive. It would not be a strange result, but the interesting part of the experiment is the following: where the same "Asian disease" issue has been presented to another group of participants through a different formulation of the same question.. In this version in program A 400 people would die while with program B there is a 1/3 probability that no one would die and a 2/3 probability that 600 people would die. For a rational decision maker the problem is exactly the same as in the first formulation, so also the answer should be the same. However, everything changed in the experiment. In the second formula the majority of participants (78%) selected program B.

The explanation of this experimental result is that the participants were now risk prone. How was it possible? What is the key to understanding these contradictory results? The key is the framing. The first formulation was in terms of gains (saved people) while the second in terms of losses (people who die). Real agents are risk averse when they must decide about

[10] Tversky & Kahneman (1985).

gains but risk prone is the realm of losses. This result is extremely important and it is not present in the expected utility theory according to which people can be risk prone or (more frequently) risk averse but the concepts of losses and gains do not cause any difference. In fact, losses and gains are *relative* concepts. To define them you must have a reference. This is what changed in the two formulations of the Asian disease problem. The first formulation encouraged participants to use as a reference the whole group of people alive (how many of them would you accept to lose?), while the second favoured a reference where everyone is going to die (how many of them would you be able to save?). Neoclassical utility functions only accept absolute arguments, so are not able to capture the difference between a loss and a gain. According to the mainstream approach, your satisfaction for your new monthly salary of 1.500 dollars should be the same, either if until yesterday it was equal to 1.000 dollars or if it was 2000 dollars. Obviously the real world is not like that.

In deepening the analysis of how people evaluate losses and gain with respect to a reference value, Kahneman and Tversky also found that losses loom larger than gains of the same size. This behavioral feature, known as *loss aversion*, is at the origin of many different unexplained behaviors according the expected utility theory (from the decisions of NYC cab drivers to the most experts traders of Wall Street) and it is so deep-rooted in our genes that it has been found even in monkeys.

These experimental discoveries permitted Kahneman and Tversky to formulate an alternative theory of decision making under risk, the Prospect Theory[11]. The utility function with an absolute argument is replaced by a *value function* centered in a (subjective) reference value, with a shape denoting risk aversion in the realm of gains and risk proneness in the realm of losses. Moreover, the value function is characterized by asymmetry with respect to the reference value, such that it captures loss aversion.

This is probably the most famous example of the experiments-driven theory of decision making.

Concerning intertemporal decisions, we have seen in the previous section how the normative property of dynamic consistency (or coherence) is systematically violated by real decision makers. Experiments also suggest that people tend to prefer larger and later (LL) to smaller and sooner (SS) rewards when the two alternatives are both in the future. Usually this choice

[11] Kahneman & Tversky (1979), Tversky & Kahneman (1992). Indeed, there is another part of Prospect Theory devoted to how people really *weight* probabilities, but we prefer not to consider this issue in the present paper, even if it contributes to explaining the Allais paradox, see e.g. Kahneman & Tversky (1979).

is also considered the virtuous one (saving over consuming, healthy over junk food, etc...). As long as the SS alternative is no more in the future, but becomes available today, often people tend to switch and change their minds and fall into temptations. This phenomenon is known as *present bias* and may cause the regret of the choice when it is too late to change it.

Some researchers have tried to propose an alternative to the exponential discounting function that is not able to foresee this kind of behavior. They needed a function steeper than the exponential one in reducing the value of something whose availability passes from immediate to the near future, and almost constant in the far future (i.e. the value of a reward available 10 years or 11 years from now is not so important, while the difference between today and one year from now is a lot more important). They found than a *hyperbolic* shape permits the inclusion of dynamic inconsistency in the analysis, producing more descriptive predictions. Several hyperbolic functions have been proposed during the last decades, starting from Ainslie (1975) and then Mazur (1987) and Loewenstein & Prelec (1992). Probably the most representative is the one proposed by Laibson (1997) who adjusted the exponential discounting function (7) by only introducing a new parameter, β:

$$U_0 = u_0 + \beta \sum_i \delta^i u_i \qquad (9)$$

With $0<\beta\leq 1$ measuring self-control. When $\beta=1$ the quasi-hyperbolic function (9) becomes equal to the exponential one (7), implying a total self-control and dynamic coherence. Instead if $\beta<1$ each utility that is not immediately available is furtherly discounted by the self-control factor.

This is a clear example of the difference that a single parameter may bring to a model which was originally not descriptive. Without the experimental evidence it would have been quite difficult (if not impossible) to think of such an alternative formulation of the discount function.

Regarding the Bayesian updating, in this case it has also been necessary to perform a large number of experiments to understand why it fails to be descriptive and how it could be adjusted to take into account the cognitive limitations of decision makers. People have repeatedly been shown to have the tendency to look for, select and interpret data confirming their already existing beliefs. This bias has been named *confirmation bias* and is spread in every field of decision making[12]. Given a certain belief, human beings are

[12] See Nickerson (1998) for a survey.

more inclined to search for corroboration instead of denial. When evidence is ambiguous or neutral it is often interpreted in favour of pre-existing beliefs (which is what occurred in the death penalty experiment). Moreover, when we face evidence against our priors, we are more selective about the source of this evidence, disregarding it if the source is not completely reliable, whereas we are not so demanding about the quality of the source if the evidence is by contrast consistent with our current opinion.

In a famous experiment, Wason (1966) showed the subjects of the experiment a set of four cards placed on a table. It was common knowledge that he had an integer number on one side and a letter on the other side. Two cards displayed the number side (2 and 7, respectively), while the other two the letter side (A and K).The experiment asked how many cards must be turned over in order to test the truth of the following proposition: "*if a card shows a vowel on one face, then its opposite face shows an even number*".

About 33% of the participants in the experiment answered that it is enough to turn over only one card: the one showing the letter A. According to their reasoning, if we find an even number on the other side then the proposition was true, otherwise it was false. But this is not the right answer! It is incomplete. 46% of the subjects added a second card to be turned over: the one with the number 2 exposed. But this is unnecessary! In fact, even if we find a consonant on the other side this would not be a confutation of the proposition. Only 5% of subjects gave the correct answer. Two cards must be turned over: the one with the vowel A and the one with the odd number 7. In fact, if we find a vowel on the other side the proposition could be considered false. Why do the vast majority of participants correctly identify one card but not the other one? According to psychologists (and coherently with the confirmation bias) people only look to evidence supporting a given proposition (that becomes a prior) and hardly look for proof against it.

If this is obvious, then it is possible to propose a different rule for updating probabilities. This has been done by Rabin and Schrag (1999), who just changed the definition of one variable in the original Bayes Law. Instead of using objective evidence in favour or against a certain proposition, they introduced the concept of "perceived" evidence. In other words, they introduced a positive probability that evidence adverse to the prior is perceived as favourable to it. The higher is this probability the more is the subject affected by confirmation bias. They show that it is not unlikely that a subject continues to believe (with high probability) in false hypothesis even after having received infinitely more items of adverse information.

With the help of these three leading examples, we have shown that alternative theories of human decision making have been developed on the basis of the regularities found experimentally.

We can take into account other important extensions of studies on heuristics and their roles within the process of decision making and economic choices. Gigerenzer and Goldstein (1996) discovered the heuristic called "take-the-best" by which people estimate information and given alternatives even partially but effectively, following a specific cue validity that orders discrimination among the alternatives. So, when economic agents cannot clearly distinguish between two given alternatives, they will begin a search in order to find a cue which will provide a reason for choosing what is perceived as more appropriate[13].

Feelings linked with the decision making process have become other specific research lines about heuristics. People rely on affect when they judge risks and benefits of specific hazards, as stated by influential research on *affect heuristic*[14]. According to this, people are prone to using an 'affect heuristic' which improves judgmental efficiency by deriving both risk and benefit evaluations from a common source, the affective reactions to the given situation. Under a time-pressure condition one limits the use of analytic thought and enhances reliance on affect.

Two more studies prove the interest of heuristics for research in social science. The attempt to use the individual differences in thinking styles as predictors of heuristics choices[15] and some speculative remarks about how heuristics and biases, outside of laboratory, are illuminating many problems of law and policy. Indeed, Sunstein[16]claims that "if heuristics play a role in factual judgments, and sometimes lead people to make systematic errors, there is every reason to believe that heuristics also help produce normative judgments, both moral and legal, and sometimes produce errors there as well. If this is harder to demonstrate, it is largely because we are able to agree, in the relevant cases, about what constitutes error in the domain of facts, and often less able to agree about what constitutes error in the domain of values. […] But we should not be surprised if the ideas of attribute substitution, and of the correction of rapid, intuitive assessments by more reflective processes, have analogues in moral and legal intuitions as well".

Following Sunstein's point we can obtain the observation of two consistent and parallel processes of thinking: the heuristic making and the

[13] Balconi & Antonietti (2009).
[14] Finucane *et al*. (2000), Kahneman & Frederick (2002).
[15] Shiloh *et al*. (2001).
[16] Sunstein (2002: 36).

reflective way. They are extensions of rationality. They can work simultaneously or alternatively, consciously or unconsciously: both are resources and aptitudes of subjects engaged in managing decision processes. They present two possibilities of reasoning, both logic but based on different logics. Therefore, heuristics are not irrational, but based on a different kind of reasoning. Decisions are affected by individuals' personality, cases of life, social conditions as well as subjective feelings and thoughts which are usually excluded from experimental research, not only within economics research. We can hope for an improvement in research so that it will be able to include more features of real agents when processes of decision making are described[17].

According to Gigerenzer's critics (1991), we prefer to distinguish heuristics and bias as arguments produced by a cognitive revolution. We see heuristics as representative of an extended concept of rationality, able to consider abstract models of behaviour compared with real subjects' behaviour. The phenomena highlighted by the experiments concerning heuristics are representative of orientation lines of subjective normativity discovered by the Freudian method[18]. The *present bias* as well as the *loss aversion* are behaviours centred in a *subjective reference value*, based on *perceived evidences*. Why are they called irrational? Subjective is not synonymous of irrational. This is the challenge raised by the research on heuristics in all fields of human knowledge.

5. Some conclusions and future trends

The process of introducing psychological concepts into economics and decision making theory has been long, slow and is still incomplete. An expediting factor has been the recent financial crisis, that made evident the limitations of the traditional approach to decision making. Even the former chairman of the Federal Reserve Board, Alan Greenspan (certainly not a behavioral economist), during an expansive phase of financial markets in 1996, used the expression "*irrational exuberance*", while Paul Krugman, Nobel laureate in Economics, wrote a quite meaningful article on 6th September 2009 in the New York Times emblematically entitled "How did economists get it so wrong?". Moreover, the behavioral economists Richard Thaler and Robert J. Shiller have recently been presidents of the American

[17] See also Pediconi *et al*. (2013).

[18] A survey of Freud's discoveries implications for economics and social sciences can be found in Contri (2011).

Economic Association. If it is true that Behavioral Economics and Finance are now commonly accepted as useful instruments in understanding human behaviors, the neoclassical *Homo oeconomicus* dies hard and it is still the cornerstone of the dominant approach. It is quite meaningful that in the same year that the Nobel prize in Economics was assigned to Robert J. Shiller (one of the pioneers in the application of psychology to finance) and Eugene Fama (father of the efficient market hypothesis, based on perfect rationality) who shared the prize with him.

Nevertheless, several developments can be observed. By combining the findings of behavioural economics with other disciplines (such as the network theory), it has been possible to move from a micro to a macro level, studying the aggregation of heterogeneous and boundedly rational agents (see agent-based models) with interesting results. One of them comes from the analysis of the role of the unconscious in financial decisions, called emotional finance[19].

While the study and discovery of regularities in the boundedly rational behavior of real agents continues, the consideration is being made that the cognitive limitations of decision makers can be in some sense exploited. Once it is understood that decisions are not based on objective and innate preferences, but can be induced by varying some aspects of the context where the decisions take place, this knowledge can be a powerful tool for reaching important goals. Sometimes these goals can be positive, embraceable and of benefit to the decisors (see for instance the SmarT program to save more for retirement). However, at other times the beneficiary can be a government (the so-called "*nudges*" to regularly pay taxes), the collectivity (psychological incentives to sustainable and environmental friendly behavior) or even private subjects. It may belong in the last case to the field called *neuromarketing* in which it limits itself to the study of consumers' preferences in order to better satisfy them, but it aims at influencing and changing these preferences for some brands[20].

In 1999 Richard Thaler wrote in a survey paper that the aim of these psychological studies is to arrive at a moment when "...the term *Behavioral Finance* will be correctly view as a redundant phrase"[21]. After 18 years we have not arrived at such a stage yet, but we are surely closer.

[19] Taffler & Tuckett (2010).
[20] See e.g. Viganò *et al*. (2015).
[21] Thaler (1999).

References

Ainslie, G. (1975), "Specious Reward: A Behavioral Theory of Impulsiveness and Impulse Control", in *Psychological Bulletin*, 82, 463-496.

Allais, M. (1953), "L'extension des théories de l'équilibre économique général et du rendement social au cas du risqué", in *Econometrica*, 21, 269-290.

Balconi,M., & Antonietti, A. (2009), *Scegliere, comprare. Dinamiche di acquisto in psicologia e neuroscienze*, Springer Science & Business Media.

Bayes, T. (1763), "As Essay Solving a Problem in the Doctrine of Chances", in *Philosophical Transactions*, 53, 370-418.

Bischi, G. I. (2012), "Modelli dinamici per le scienze sociali", in *Complessità e riduzionismo*, pp. 7-19, *Isonomia, Rivista online di Filosofia* – Epistemologica – vol. 2 (2012).

Bischi, G. I.(2014), "Adaptive and Evolutionary Mathematical Models in Economics", in *Contributi del Centro Linceo Interdisciplinare "Beniamino Segre",* n.131, Scienze e Lettere, pp. 179-196.

Contri, G. B. (2011), *A chi non ha sarà tolto. Economia e psicopatologia*, Sic Edizioni, Milano (ebook).

Ellsberg, D. (1961), "Risk, Ambiguity, and the Savage Axioms", in *The Quarterly Journal of Economics*, 75, 643-669.

Finucane, M.L., Alhakami, A., Slovic, P., & Johnson, S. M. (2000), "The Affect Heuristic in Judgments of Risks and Benefits", in *Journal of Behavioral Decision Making*, 13, 1–17.

Flabbi, L., & Pediconi, M. G. (2013), "Unconscious and Game Theory", in *International Journal of Applied Psychoanalytic Studies*, 11(4): 339–359.

Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002), "Time Discounting and Time Preference: A Critical Review", in *Journal of Economic Literature*, 40, 351-401.

Freud, S. (1927), *The Future of an Illusion*, SE vol. XXI.

Gigerenzer, G. (1991), "How to Make Cognitive Illusions Disappear: Beyond 'Heuristics and Biases'", in *European Review of Social Psychology*, 2, 83–115.

Gigerenzer, G., & Goldstein D. G. (1996), "Reasoning the Fast and Frugal Way: Models of Bounded Rationality", in *Psychological Review*, 103, 650-669.

Kahneman, D. (1999), "Objective Happiness", in *Well-being: The Foundations of Hedonic Psychology*, 3, 25.

Kahneman, D., & Frederick, S. (2002), "Representativeness Revisited: Attribute Substitution in Intuitive Judgment", pp. 49–81, in Gilovich, T., Griffin, D., & Kahneman, D. (Eds.), *Heuristics of Intuitive Judgment: Extensions and Applications*, Cambridge University Press, New York.

Kahneman, D., Krueger, A. B., Schkade, D., Schwarz, N., & Stone, A. A. (2006), "Would You Be Happier If You Were Richer? A Focusing Illusion", in *Science*, 312 (5782), 1908-1910.

Kahneman, D., & Tversky, A. (1979), "Prospect Theory: An Analysis of Decision under Risk", in *Econometrica*, 47, 263-292.

Kirman, A. P. (1992), " Whom or What Does the Representative Individual Represent?", in *Journal of Economic Perspectives*, 6 (2), 117-136.

Kirman, A. P. (1993), "Ants, Rationality, and Recruitment", *The Quarterly Journal of Economics,* 108 (1993) 137–156.

Knight, F.H. (1921), *Risk, Uncertainty, and Profit*, Hart, Schaffner & Marx. Noughton Mifflin Company, Boston.

Laibson, D. (1997), "Golden Eggs and Hyperbolic Discounting", in *The Quarterly Journal of Economics*, 112, 443-478.

Loewenstein, G., & Prelec, D. (1992), "Anomalies in Intertemporal Choice: Evidence and interpretation", in *The Quarterly Journal of Economics*, 107, 573-597.

Lord, C. G., Ross, L., & Lepper, M. R. (1979), "Biased Assimilation and Attitude Polarization: The Effects of Prior Theory on Subsequently Considered Evidence", in *Journal of Personality and Social Psychology*, 37, 2098-2109.

Mazur, J. E. (1987), "An Adjustment Procedure for Studying Delayed Reinforcement", pp. 55-73, in Commons, M. L., Mazur, J. E., Nevin, J. A., & Rachlin, H. (Eds), *Quantitative Analyses of Behavior,* Eribaum, Hillsdale, New York

Nickerson, R.S. (1998), "Confirmation Bias: A Ubiquitous Phenomenon in Many Guises", in *Review of General Psychology*, 2, 175-220.

Pediconi, M. G., Genga, G. M., & Flabbi, L. (2013) "Invidia versus legame sociale. Una nuova idea di profitto", in *Teorie&Modelli*. XVIII, 2, pp. 7-23.

Phelps, E. A., Lempert, K.M., & Sokol-Hessner, P. (2014), "Emotion and Decision Making: Multiple Modulatory Neural Circuits", in *Annual Review of Neuroscience*, 37, 263-287.

Rabin, M., & Schrag, J. L. (1999), "First Impressions Matter: A Model of Confirmatory Bias", in *The Quarterly Journal of Economics*, 114, 37-82.

Samuelson, P. (1937), "Some Aspects of the Pure Theory of Capital", in *The Quarterly Journal of Economics,* 51, 469-496.

Samuelson, P., & Nordhaus, W. (1985), *Principles of Economics*, McGraw-Hill, New York.

Savage, L. J. (1954), *The Foundations of Statistics*, Wiley, New York.

Shiloh, S., Salton, E., & Sharabi, D. (2002), "Individual Differences in Rational and Intuitive Thinking Styles as Predictors of Heuristic Responses and Framing Effects", in *Personality and Individual Differences,* 32, 415-429.

Simon, H. (1957), *Models of Man, Social and Rational: Mathematical Essays on Rational Human Behavior in a Social Setting*, Wiley, New York.

Simon, H. (1982), *Models of Bounded Rationality*, Volume 3, MIT Press, New Haven.

Sunstein, C.R. (2002), "Hazardous Heuristics", in Olin, J. M., *Law and Economics, Working Paper*, No. 165 (2nd Series) http://ssrn.com/abstract_id=344620

Taffler, R.J., & Tuckett, D. A. (2010), "Emotional Finance: The Role of the Unconscious in Financial Decisions", in Baker, H. K., & Nofsinger, J. R. (Eds.), *Behavioral Finance: Investors, Corporations, and Markets,* John Wiley & Sons, Inc., Hoboken, New York, USA.

Thaler, R.H. (1999), "The End of Behavioral Finance", in *Financial Analysts Journal*, 55, 12-17.

Tversky, A., & Kahneman, D. (1985), "The Framing of Decisions and the Psychology of Choice", pp. 107-129, in *Environmental Impact Assessment, Technology Assessment, and Risk Analysis*, Springer, Berlin.

Tversky, A., & Kahneman, D. (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty", in *Journal of Risk and Uncertainty,* 5, 297-323.

Viganò, E., Antonelli, G., Bischi, G. I., & Tramontana, F. (2015), "Consumption and Consumer of Food Products in the Postmodern Society", in *Economia agro-alimentare/Food Economy*, vol. 17 n. 1.

Von Neumann, J., & Morgenstern, O. (1947), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.

Wason, P.C. (1966), "Reasoning", in Foss, B. M., *New Horizons in Psychology*, Penguin, Harmondsworth.

Bounded Rationality and Heuristics in Humans and in Artificial Cognitive Systems

Antonio Lieto
University of Turin
ICAR-CNR, Palermo
lieto@di.unito.it

1. Introduction

The notion of bounded rationality was introduced by Herbert Simon[1] in his book *Administrative Behavior* as a constitutive element of a plausible model of human problem solving and decision making. This notion proposed, for the first time, a radically different and alternative view with respect to the idealised model of "*homo oeconomicus*" which was dominant, in that times, within the classical economic theory of decision making. Roughly speaking, the notion of bounded rationality refers to the cognitive limits (e.g. perceptual, of memory, reasoning, processing etc.) that human decision makers have, and have to face, when they have to make decisions in real life situations. Such limits lead humans to adopt "bounded", i.e. non optimal (but "*satisficing*"), solution strategies that can be cognitively dealt with. As mentioned, this approach diverged from the classical theory of decision

[1] Herbert Simon was one of the founding fathers of the AI and, along with people like Marvin Minsky, John McCarthy, Allen Newell, Nathaniel Rochester and many others, was one of the participants to the Darmouth Workshop providing the foundation of the discipline. Simon was trained as a psychologist and awarded by the Nobel Prize in Economy for his studies on the "bounded rationality" in human decision making. He was one of the main scholars pointing out, in both cognitive psychology and AI, the role played by the heuristics as decisional shortcuts to solve complex problems (for this line of research he was also awarded, with Allen Newell, by the Turing Award in 1975; see Newell & Simon,1976). On the particular meanings attributed to the term "heuristics" in the AI research, we remind the reader to the following sections.

Antonio Lieto, "Bounded Rationality and Heuristics in Humans and in Artificial Cognitive Systems", in Giorgio Grimaldi and Marialuisa Parise, Rational *Animals. Seven essays on Reason's fields of application*, pp. 53-73

making according to which humans were seen as perfectly rational decision makers, able to make optimal decisions via a maximisation of their expected utility in any give situation[2]. As I will argue in this paper, the introduction of this notion influenced not only economic model of decision making but also the field of the early Artificial Intelligence, an area in which Simon was directly involved since its foundations.

In the following, I will present two case studies in the field of automated reasoning with the intent of showing, at least in part, the impact that such notion had in this field of AI research. Before presenting these particular case studies, however, I will propose a brief historical digression concerning the influence that such idea had in the development of the early AI systems and formalisms.

2. The Early days of AI

In the early days of Artificial Intelligence, whose date of birth is historically associated to the summer of 1956, during the so called "Dartmouth Workshop"[3], the research on intelligent machines was strongly and explicitly inspired by that one coming from the experimental research in Psychology[4]. A manifestation of this connection is represented by the early AI systems/frameworks in AI developed until the '80s of the last century: most of the them were developed according to a "cognitively-oriented" inspiration. In the following we just mention few examples of such systems. and formalisms.

[2] The theory of the expected utility was introduced in game theory by Von Neumann and Morgenstern (see Morgenstern and Von Neumann, 1953). In general, in this setting, preference relations are usually modelled by means of utility functions defined on a set of alternatives with values in a suitable set of numbers (usually, real numbers). The decision model of this theory assumes an optimal (i.e. unbounded) decision maker able to calculate and chose, in each phase of problem solving, the move that maximise the utility function.

[3] This event was formally organised by John McCarthy, Marvin Minsky, Nathaniel Rochester and Claude Shannon. The workshop run for different weeks and saw the participation of many researchers interested in this emerging discipline. An archive of the Darthmouth workshop, based on the notes of Ray Solomonoff (one the participant), is available on line at the following url: http://raysolomonoff.com/dartmouth/.

[4] It must be noted that, at those times, there was not a "Cognitive Science" or a "Cognitive Psychology" field. However, all the disciplines (philosophy, psychology, computer science, anthropology, linguistics and neurophysiology) and the cultural elements that would have been later formed the interdisciplinary field of "Cognitive Science" were already present.

3. From the General Problem Solver to the Cognitive Architectures

One of the first AI systems, developed from the end of the '50s of the last century, is represented by the pioneering work of Herbert Simon, John Clifford Shaw and Allen Newell on the *General Problem Solver (GPS):* a system able to demonstrate simple logic theorems whose decision strategies were explicitly inspired by human verbal protocols[5]. The underlying idea of this approach was that, the computer system had to approximate the decision operations described by the humans in their verbal descriptions as closely as possible. In this way, when the program ran on the computer, it would have been possible to identify its problems, compare them with the description of the human verbalisation and modify them to improve its performance. In particular, the GPS system was able to implement a key mechanism in human problem solving: the so called means-ends analysis (or **M-E heuristic**). In M-E analysis the problem solver compares the current situation with a goal situation; computes the difference between the two states; finds in memory an operator that experience has taught reduces differences of this kind; and applies the operator to change the current situation. Repeating this process, the goal may gradually be attained via a reduction of the search space(although, computationally, there are generally no guarantees that the process will succeed). This kind of heuristic was used to solve, also in the decades to come, problems in a number of AI domains (from planning to diagnostics). In order to be executed, in fact, it "only" required a domain representation of the problem to solve (a problem space), operators to move through the space, and information about which operators were relevant for reducing which differences[6].

In the decade following the development of the GPS, i.e. in the '60s, another influential contribution in AI of cognitive inspiration was due to Ross Quillian (a Ph.D. student of Herbert Simon at the Carnegie Institute of Technology, now Carnegie Mellon University) that invented the *Semantic Networks,* a psychologically plausible model of human semantic memory implemented in a computer system. The idea of Quillian was that human memory was associative in nature and that concepts were represented as sort

[5] Newell *et al*. (1959). This technique in also known as "thinking aloud" protocol in the psychological literature and consists in recording the verbal explanations provided by the people while executing a given laboratory task.

[6] The ingredients required for the execution of this kind of heuristic strategy - essentially based on a "search space" approach to problem solving - explicitly supported the so called "symbolic approach" to the study, the analysis, the execution and the replication of intelligent behaviour.

of nodes in graphs and activated through a mechanism of "spreading activation", allowing to propagate information through the network to determine relationships between objects. In this setting, the highest was the activation of a node in the network, the more contextually relevant that node/concept was assumed to be for the task in focus (e.g. retrieval, reasoning, recall, planning etc.). Also in this case, the cognitive limit concerning the storage capacity of our memory triggered the inquiry about the underlying mechanisms governing our memory system. As a consequence, this led to the development of a cognitively inspired model of knowledge representation and retrieval. Evolution of the semantic networks are still nowadays widely used in the subfield of AI known as Knowledge Representation (and in particular in systems known as computational ontologies, that will be better introduced later in the paper).

In the '70s, another well know example of cognitively-inspired AI framework was represented by the introduction of the *Frames* operated by Marvin Minsky[7]. This framework was used for representing, in artificial systems, commonsense knowledge (including *default* knowledge) about the external world. This type of knowledge organization proposed in the Frames enabled the first AI systems to extend their automated reasoning abilities from the classical deduction to more complicated forms of commonsense and defeasible reasoning[8]. In this case, the idea of the Frames was directly inspired by the work of the psychologist Eleanor Rosh about the organization of conceptual information in humans known as prototype theory[9].

In the '80s, finally, other notable examples of the so called "cognitivist" tradition in Artificial Intelligence were represented by: i) the invention of the notion of *Script (Shank and Abelson)* as a data structure for representing knowledge of common sequences of events (e.g. the sequence of events used to go out for a dinner) and used in natural language processing systems as way to enable intelligent to answer questions about simple stories[10] or by

[7] Minsky (1975).

[8] All the forms of commonsense reasoning can be seen as abounded rationality phenomenon since they represent a plethora of shortcuts allowing us to make decisions in an environment with incomplete and uncertain information.

[9] According to the Rosh's theory, concepts, the building block of our thoughts, are organised in our mind in terms of prototypes (i.e. in terms of representative elements of the category) and such organization explain many types of so called typicality effects (i.e. of commonsense inferences) that we naturally perform in our everyday reasoning.

[10] A classical example to explain the notion of "Script" is the so called "restaurant situation". Let us consider to model a situation of an agent going out to a restaurant for dinner. A script representing the restaurant situation is a data structure that would record the typical events associated to this scenario: e.g. entering in the restaurant, asking for a table, sitting

ii) the work operated by Allen Newell and colleagues at Carnegie Mellon University on the first cognitive architecture for general intelligence: SOAR[11].

This list of examples is, of course, not exhaustive but it can be useful to indicate the main assumptions made by the early AI researchers and synthesized by Pat Langley[12] in this way: "(early) AI aimed at understanding andreproducing in computational systems the full range of intelligent behavior observed by humans"[13]. The underlying assumption of that period was that, in order to reach this broad goal, the only way to proceed in the analysis of intelligent systems (natural and artificial) was that one of discovering the heuristic strategies adopted by humans (or other animals) that could have been later operationalised and replicated in artificial systems.

It is worth-noticing, however, that this kind of approach to the "study of the artificial" did not came out *ex-abrupto.* It borrowed its original inspiration –from a historical perspective–from the methodological apparatus developed by the scholars in Cybernetics[14].

The origins of cybernetics are usually traced back to the middle of the 1940s,with the release of the 1948 book of Norbert Wiener entitled *Cybernetics: Or Control and Communication in the Animal and the Machine*. One of underlying idea of cybernetics was that one of building mechanical models to simulate the adaptive behavior of natural systems. As indicated in Cordeschi[15]: "the fundamental insight of cybernetics was in the proposal of a unified study of organisms and machines". In this perspective, the computational simulation of biological processes was assumed to play a central epistemological role in the development and refinement of theories about the elements characterizing the nature of intelligent behaviour in natural and artificial systems. Such simulative approach, as mentioned, was inherited by the early AI research and the adoption of such perspective was crucial for

down, consulting a menu, eating the food, pay the check etc. This kind of representational structure enabled early AI system to answer questions to simple stories. For a story like this, for example: "Mary went to a restaurant, ordered a salmon. When she was paying she noticed that she was late for the next appointment", computerised systems were able to answer to a question like: "Did Mary eat dinner last night?" in a positive way (as we do). It is worth-noticing that this information is not explicitly provided in the story. Answering to these types of questions was possible through the use of a script of the restaurant situation.

[11] Laird *et al*. (1987). On the role of the Cognitive architectures for general intelligent systems we remind to (Lieto *et al*., 2018).

[12] Langley (2012).

[13] Ibid.

[14] Cordeschi (1991).

[15] Cordeschi (2002).

the development of both intelligent solutions inspired by human processes and heuristics[16] and for the realisation of computational models of cognition realized with the aim of providing a deeper understanding of human thinking, as originally suggested in the manifesto of the Information Processing Psychology (IPP)[17]. These two sides of the cognitivist tradition are nowadays still alive. They correspond, roughly, to the research areas of Cognitive Systems or cognitively-inspired AI (Cognitive-AI) and to that one of cognitive modelling (or Computational Cognitive Science) respectively.

4. Heuristics and AI Eras

The notion of heuristics deserves, in this historical digression, and in the light of the purposes of this paper, a special attention. It was used, in fact, with two different meanings since the times of the first AI researches. In its first sense, the term referred to the most detailed simulation possible of human cognitive processes, and it characterized the above mentioned Information Processing Psychology (IPP) introduced by Newelland Simon. In this view, a computer program was considered as a model providing a test of the hypothesis that the mind is an information-processing system. More precisely, "the program was considered to be a highly specific behavioral theory, concerning the behavior of an individual human problem-solver: a microtheory"[18].

In another sense, the term referred to the possibility of obtaining the most efficient (and efficacious) performance possible from computer programs, by allowing also for typically non-human procedures, such as those where the computer can excel. Before the introduction of the term "heuristics" in AI, operated by Newell, Shaw and Simon, there were already available algorithmic procedures which might have been defined as heuristic in the second of these senses had already been tried out experimentally. The first among them were the procedures that allowed the program developed by Arthur Samuelto play chess despite the combinatory explosion of moves[19].

[16] Newell & Simon (1976), Gigerenzer & Todd (1999).
[17] Newell & Simon (1972).
[18] Cordeschi (2002: 182). In this view, the general theory of human information-processing was assumed to be derivable from a body of qualitative generalizations coming from the study of individual simulative programs, or microtheories.
[19] Samuel (1959).

The fact that these two tendencies, reflected in this double meaning of the term "heuristics", coexisted in AI was immediately clear. As reported in Cordeschi (2002: 190), in 1961, while discussing a presentation of GPS given by Simon during a seminar at MIT, Minsky drew a distinction in AI research between those who were willing to use "non-human techniques" in constructing intelligent programs and those, like the Carnegie-Mellon group, who were interested in simulating human cognitive processes[20]. This distinction is crucial, since outlines the raise of different research agendas already in those years. This divergence started to be very significant in the mid'80s of the last century. After decades of pioneering collaborations, in fact, several sub-fields, each with its own goals, methods and evaluation criteria were produced. On the one hand this fragmentation led AI to reach remarkable results in a variety of narrow fields by focussing on quantitative results and metrics of performance, and on a machine-oriented approach to the intelligent behaviour (i.e., without taking into account human-inspired heuristics). On the other hand, however, it significantly inhibited the cross-field collaborations and the research efforts targeted at investigating a more general picture of what natural and artificial intelligence is, and how intelligent artifacts can be designed by taking into account the insights coming from human cognition.

In the last few years, however, the cognitive approach to AI gained a renewed consideration, both from academia and industry, in wide research areas such as Knowledge Representation and Reasoning, Robotics, Machine Learning, Bio-Inspired Cognitive Computing, Computational Creativity and further research fields that aspire to Human Level Intelligence. Nowadays, in fact, artificial systems endowed with human-like and human-level intelligence[21] are still far from being achieved and, as I will try to show in the following section, the adoption of cognitively-inspired heuristics (all directly based on bounded rationality assumption) seems to be a suitable way to handle problems that non-human oriented algorithmic processes are not able to deal with.

In the next section, I will decline this argument by presenting two case studies in the field of knowledge representation and automated reasoning.

---

[20] As reported in Cordeschi (2002), Minsky emphasized that these two tendencies were distinguished "in methods and goals" from a third tendency, that which "has a physiological orientation and alleges to be based on an imitation of the brain," i.e. neural net and self-organizing system approaches. This tendency would have lead to the development of the neural and brain inspired methods which are the basis of the so-called "connectionist agenda" (nowadays wide-spreading due to the recent success of deep learning).

[21] McCarthy (2007).

5. Cognitive heuristics in modern AI: two case studies in Automatic Reasoning

The representation of conceptual information, and the corresponding reasoning mechanisms that can be built on such representations, is a crucial problem to deal with in the subfield of AI known as Knowledge Representation (KR). In this field, one of the most widespread system currently used are the so called computational ontologies, defined as"an engineering artifact, constituted by a specific vocabulary used to describe a certain reality, plus a set of explicit assumptions regarding the intended meaning of the vocabulary words"[22]

The main building blocks of ontological models are, therefore, concepts (or classes), roles (or properties), and individuals describing a given domain. In other words: ontologies provide an explicit and axiomized reference domain model. Such model is used to interpret and organise the information and to perform simple forms of automatic reasoning.

Ontological models are built on by using a class of logical formalisms that are known as Description Logics (DLs)[23]. Technically, DLs are decidable subsets of first-order logic in which a standard Tarskian semantics is directly associated to the syntax of the language. The decidability is usually obtained by restricting the set of variables and operators that are allowed inside formulae. In DL systems, inferential knowledge is formally expressed by means of "terminological axioms", which are simply notational variants of classical meaning postulates - i.e. universally quantified (bi)conditional statements that constrain the extensions of the constant that appear in the antecedent (e.g. see the figure below).

[22] Guarino (1998).
[23] Badeer *et al*. (2010).

Figure 1. Examples of terminological axioms in Description Logics

The meaning postulates approach assumed in the tarskian semantics of Description Logics, however, is not immune from problems. For instance, it is known that meaning postulates can formally express only rigid relations between concepts, i.e. relations that do not allow for exceptions (e.g. if x is a bachelor, than x is not married). As a consequence, standard DL system and ontologies are not capable to model prototypical knowledge and defeasible inference[24], which - on the other hand - represents the way in which the humans encode and reason on conceptual information.

In the field of logic-oriented KR, various non-monotonic extensions of DL formalisms have been designed to deal with some aspects of typicality and automated commonsense reasoning[25]. A proposed solution has been that one of integrating a set of meaning postulates with a non-monotonic logic like Default Logics[26]. This is the solution adopted by Baader & Hollunder (1995); the same authors, however, point out both the computational difficulties of this integration. Current literature also offers several attempts to formulate DLs extensions based on fuzzy logic[27] and some "probabilistic" versions of DL languages have been also proposed, which try to combine the DL language with probabilistic theories based on Bayesian formalisms[28]. Despite these computational advancements, however, the use of non-

[24] Lieto (2012).
[25] E.g. Bonatti *et al*. (2006), Lukasiewicz & Straccia (2008), Giordano *et al*. (2013).
[26] Reiter (1980).
[27] E.g., Caligari *et al*. (2007).
[28] Ding *et al.* (2006).

monotonic and, in general, non-classical logics in the field is not universally accepted due to the high computational complexity, and therefore their practical intractability, intrinsic to such formalisms. In other words: the use of non-human inspired procedures does not seem to provide any practical advantage for modelling the problem of commonsense reasoning related to typicality.

An alternative move was proposed by adopting a cognitively-inspired system explicitly implementing a plethora of cognitive assumptions and heuristics for commonsense reasoning. This has recently led to the development of a system called DUAL-PECCS (Dual Prototype and Exemplars Based Conceptual Categorization System)[29], able to combine on a large scale, both commonsense representation and reasoning with standard ontological DL-based semantics. The main merit of such proposal lies in the adoption of the representational component of Conceptual Spaces (a geometric and cognitively inspired framework for concept representation and formation, see Gärdenfors, 2000) that has been successfully integrated with ontological formalisms[30].

DUAL PECCS, explicitly implement the following cognitive assumptions: i) the fact that conceptual representations are heterogeneous co-referring representational structures demanded to different computational frameworks and ii) the fact that the different reasoning strategies executed on such representations are harmonised according to the dual process theory or reasoning and rationality[31].

The heterogeneous conceptual architecture of DUAL PECCS currently includes "typical" commonsense representations (i.e. prototypes and exemplars) and classical "rigid" representations allowed by the Description Logics Semantics. All these different bodies of knowledge point to the same conceptual entity[32]. An example of the heterogeneous conceptual architecture of DUAL PECCS is provided in the figure 2 below. Such figure shows

[29] Lieto *et al*. (2015), Lieto *et al*. (2017).

[30] The solution adopted in DUAL-PECCS is one of the three proposals currently suggested to overcome the current problems affecting the knowledge level of cognitive artificial systems. Alternative proposals are represented by the Semantic Pointers (Eliasmith *et al*., 2012) and by the neuro-symbolic approaches à la ACT-R (Anderson *et al*., 2004). A detailed comparative analysis among the different approaches is in Lieto *et al*. (2018b).

[31] For the details see Lieto (2014) and Lieto (2019).

[32] The procedures for the automatic anchoring for these different types of representations are described in Lieto *et al*. (2016).

how it is represented the concept DOG. In this case, the prototypical representation grasps information such as that dogs are usually conceptualized as domestic animals, with typically four legs, a tail etc.; the exemplar-based representations grasp information on individuals. For example, it is represented the individual of Lessie, which is a particular exemplar of dog with white and brown fur and with a less domestic attitude w.r.t. the prototypical dog (e.g. its typical location is lawn).

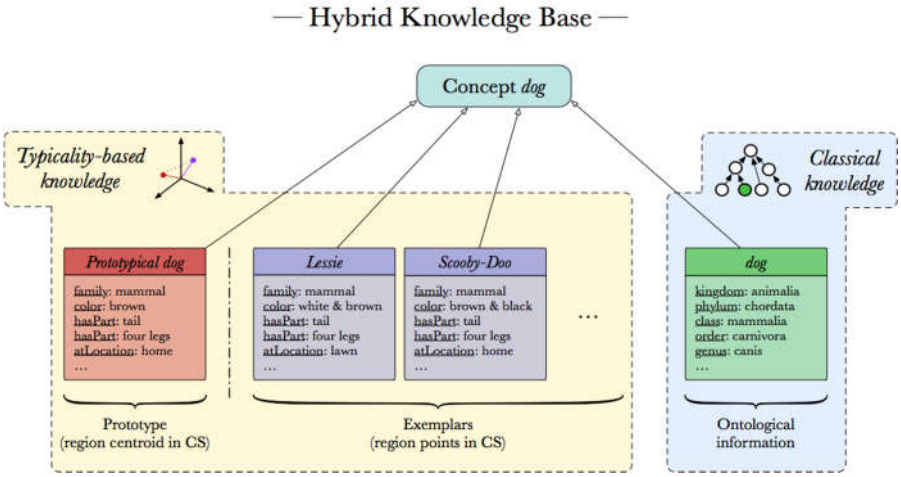


Figure 2. An example of the hybrid conceptual architecture in DUAL PECCS.

Within the system, both these two types of typicality based representations are represented by using the framework of the conceptual spaces proposed by Gärdenfors[33]. Such framework allows to adopt standard similarity metrics to determine the distance between instances and concepts within the space. The representation of classical information (e.g. the fact that Dog ⊑Animal , that is to say that "Dogs are also Animals") is, on the other hand, demanded to standard ontological formalisms.

From a reasoning perspective, in DUAL-PECCS, different kinds of reasoning strategies are associated to these different bodies of knowledge. In particular, the system combines non-monotonic commonsense reasoning (executed on conceptual spaces representations and based on measures of "semantic similarity") and standard monotonic categorization procedures (executed on the classical, ontological, body of knowledge). These different

[33] Gärdenfors (2000).

types of reasoning are harmonized according to the theoretical tenets coming from the dual process theories of reasoning and rationality according to which fast - non monotonic - reasoning is executed first and more logical - deliberative - reasoning is executed later.

Given this explicit cognitive assumptions, DUAL PECCS has been tested in conceptual categorisation of commonsense linguistic descriptions similar to riddles (i.e. of the form "which is the big feline with yellow fur and black stripes?"etc.).

This kind of task is a very difficult one from an AI perspective since, in commonsense reasoning, also state-of-the-art AI systems like IBM Watson obtain very poor results[34]. On the other hand, the results obtained by DUAL PECCS are promising both when compared with human performances (with an overlapping of the 89% of the responses, see Lieto et al., 2017 for the results in details) and when compared to other artificial systems like Google, Bing or Wolphram Alpha[35]. An additional advantage of the proposed approach lies in the fact that it limits the computational complexity of the commonsense reasoning strategy to linear time (i.e. O(n)), since the mechanism of typicality based categorisation technically corresponds to measuring the semantic relatedness of a query vector - in which the linguistic input is transformed - with respect to a vectorial knowledge-base of conceptual spaces (and this process can be solved in linear time with respect to the size of the knowledge base).

As a consequence, this case study seems to suggest that, also in the modern AI, the adoption of bounded rationality heuristics in the design and implementation of intelligent systems can represent an important aspect to consider in order to progress forward more human-like and human-level AI systems.

As a second case study, which also related to the above mentioned problem of commonsense representation and reasoning in AI systems, let us consider the problem of obtaining, via compositionality heuristics, a commonsense compound concept obtained by combining the typical knowledge of the composing concepts.

Dealing with such ability requires, from an AI perspective, the harmonization of two conflicting requirements that are hardly accommodated in

[34] See Davis & Marcus (2015).
[35] See Lieto *et al*. (2017b) for a detailed analysis of the obtained results.

symbolic systems: the need of a syntactic compositionality (typical of logical systems) and that one concerning the exhibition of typicality effects[36]. According to a well-known argument[37], in fact, prototypical concepts are not compositional. The argument runs as follows: consider a concept like PET FISH. It results from the composition of the concept pet and of the concept fish. However, the prototype of pet fish cannot result from the composition of the prototypes of a pet and a fish: e.g. a typical pet is furry and warm, a typical fish is grayish, but a typical pet fish is neither furry and warm nor grayish (typically, it is red).

Recently, a logical framework able to account for this type of human-like concept combination has been proposed in the field of nonmonotonic Description Logics of typicality: T^{CL} (Typicality-based Compositional Logic, for the details see Lieto & Pozzato, 2018).

This logic combines three main ingredients. The first one relies on the DL of typicality ALC + TR introduced in Giordano *et al*.(2015). In this logic, "typical" properties can be directly specified by means of a "typicality" operator T enriching the underlying DL, and a knowledge base (KB) can contain inclusions able to represent that "typical Cs are also Ds". In ALC + TR one can consistently express exceptions and reason about defeasible inheritance as well.

A second ingredient is represented by a distributed semantics similar to the one of probabilistic DLs known as DISPONTE[38], allowing to label ontological axioms with degrees representing probabilities, but restricted to typicality inclusions. The basic idea is to label inclusions of the type "typical Cs are also Ds" with a real number between 0.5 and 1, representing its probability, assuming that each axiom is independent from each others (the actual probabilistic values are assumed to come from an application domain). The resulting knowledge base defines a probability distribution over scenarios.

The third element of the proposed formalization is represented by the adoption of an heuristics method inspired by cognitive semantics[39] called HEAD-MODIFIER heuristics. Such heuristics allows to characterise a dominance effect between the concepts to be combined. In particular, for every combination, it allows to distinguish a HEAD, representing the

[36] See Lieto (2012) and the references in Lieto *et al*. (2018) for a contextualized overview of such a problem.
[37] Osherson & Smith (1981).
[38] Riguzzi *et al*. (2015).
[39] See Hampton (1987) for a review.

stronger element of the combination, and one (or more) MODIFIERs. The basic idea is: given a KB and two concepts CH (HEAD) and CM (MODIFIER) occurring in it, only some scenarios are considered in order to define a revised knowledge base, enriched by typical properties of the combined concept.

The selection criteria of such scenario is obtained as follows: given a KB K and given two concepts CH and CM occurring in K, $\mathbf{T}^{\mathbf{CL}}$ allows defining the compound concept C as the combination of the HEAD (CH) and the MODIFIER (CM), where $C \sqsubseteq CH \sqcap CM$.

The typical properties of the form $T(C) \sqsubseteq D$ to ascribe to the concept C are obtained in the set of scenarios, obtained by applying the DISPONTE semantics, that: i) are consistent with respect to K; ii) are not trivial (i.e. those with the highest probability, in the sense that the scenarios considering all properties that can be consistently ascribed to C, or all the properties of the HEAD that can be consistently ascribed to C are discarded); iii) are those giving preference to the typical properties of the HEAD CH (with respect to those of the MODIFIER CM) with the highest probability.

An additional element in $\mathbf{T}^{\mathbf{CL}}$, inherited from the representational assumptions considered in DUAL PECCS, is that a KB in such a formalism combines both "rigid" and "typical" knowledge.

In this way, as shown in detail in Lieto & Pozzato(2018), it is possible to model phenomena like the above described PET FISH composition. In particular, this problem is modelled as follows: let K be a Knowledge base containing the rigid inclusion (*) $Fish \sqsubseteq \forall livesIn.Water$ and the following typical inclusions equipped with probabilities:

1. 0.9 :: $\mathbf{T}(Pet) \sqsubseteq \forall livesIn.(\neg Water)$
2. 0.8 :: $\mathbf{T}(Pet) \sqsubseteq Affectionate$
3. 0.7 :: $\mathbf{T}(Fish) \sqsubseteq \neg Affectionate$
4. 0.8 :: $\mathbf{T}(Pet) \sqsubseteq Warm$
5. 0.6 :: $\mathbf{T}(Fish) \sqsubseteq Greyish$
6. 0.9 :: $\mathbf{T}(Fish) \sqsubseteq Scaly$
7. 0.8 :: $\mathbf{T}(Fish) \sqsubseteq \neg Warm$

In this case, by applying the DISPONTE semantics, we have $2^7 = 128$ different scenarios. In $\mathbf{T}^{\mathbf{CL}}$, inconsistent scenarios are discarded along with those that are trivial and privilege the MODIFIER with respect to the HEAD. It turns out that the logic $\mathbf{T}^{\mathbf{CL}}$ is able to select the scenario with the following typical properties (which are those required for handling the PET FISH prototypical composition):

3. 0.7 :: $\mathbf{T}(Fish) \sqsubseteq \neg Affectionate$
6. 0.9 :: $\mathbf{T}(Fish) \sqsubseteq Scaly$
7. 0.8 :: $\mathbf{T}(Fish) \sqsubseteq \neg Warm$

On the other hand, the composed concept PET FISH also inherits the rigid inclusion $\sqsubseteq \forall$livesIn.Water from (*).

In addition, in such a logical framework, adding a new inclusion $\mathbf{T}$(PET $\sqcap$ FISH) $\sqsubseteq$ Red, would not be problematic (i.e. this logic tackles the phenomenon of prototypical attributes emergence). The proposed logic has been recently applied to a number of cognitive phenomena including: conjunction fallacy, metaphors generation, and iterative conceptual combination[40].

An important element emerging from this particular application, and relevant with respect to the overall overview proposed in this paper, lies in the fact that it would have been not possible to model all such phenomena without considering all the 3 ingredients of such logic (including the HEAD-MODIFIER heuristics). In other words, all such elements are individually necessary but only jointly sufficient to tackle a complex problem like the one described above. This is a symptom of the fact that, the applica-

[40] Lieto & Pozzato (2019).

tion (and integration) of cognitive heuristics in artificial systems and formalisms, can still play an important role in AI.

6. Concluding Remarks

In this paper I have argued that the notion of bounded rationality heuristics has influenced, directly or indirectly, the theoretical and applied research in the field of Artificial Intelligence. I have provided a brief historical contextualization to the rise, the fall and the recent renewed interest in the design approach based on cognitive heuristics and historically ascribable to the cognitivist tradition in AI. Finally I have provided two simple case studies in the field of ADM by showing that the role that cognitive heuristics can play in the so-called "science of the artificial"[41] is still a relevant one in the years to come. In particular, a cognitive-based approach to development of AI systems seems to be still relevant in all those tasks that are easily solvable for humans but very hard to solve for machines (and the problem of commonsense reasoning is paradigmatic in this perspective). Other relevant research areas that may benefit from this kind of approach concern, for example, the problems of: analogical reasoning; learning from few examples (differently from what current deep learning techniques do); transfer learning; multimodal perception and attentive mechanisms; computational creativity and knowledge invention; narrative and story understanding; integration of mechanisms involving planning, acting, monitoring and goal reasoning; emotion modelling and the so called area of explainable AI or XAI (a novel name for a very old problem consisting in truing to provide human-understandable explanations of algorithmic decisions). As this illustrative lists shows, such problems involve both low-level (e.g. perceptual) and high level (e.g. reasoning) cognitive capacities (usually modelled in AI by adopting connectionist and symbolic approaches respectively). That is to say that the cognitive approach is agnostic with respect to the classes of formalisms applied to model a given phenomenon and can be applied to both the symbolic and the connectionist research agendas, as the recent history of Cognitive Science has successfully showed in the last 30 years.

[41] Simon (1981).

Acknowledgements

The content of this paper has benefited from many discussions with Roberto Cordeschi and Marcello Frixione. Of course eventual errors can be ascribed only to myself. I am also indebted to the feedback and comments received by Amedeo Cesta, Antonio Chella, Fabio Paglieri, Oliviero Stock and Giuseppe Trautteur during the panel
"Can AI and Cognitive Science still live together happily ever after?"
(link: http://aiia2017.di.uniba.it/index.php/joint-panel-aiia-and-aisc/)
organised at the international conference AI*IA 2017 in Bari. The work on DUAL PECCS and on the $\mathbf{T}^{\mathbf{CL}}$ logic has been carried out with Daniele Radicioni, Valentina Rho and Gian Luca Pozzato.

References

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004), "An Integrated Theory of the Mind", in *Psychological review*, 111(4), 1036.

Baader, F., Calvanese, D., McGuinness, D., Nardi, D., & Patel-Schneider,P. (2010), *The Description Logic Handbook: Theory, Implementations and Applications*, 2nd edition, Cambridge, University Press.

Baader, F., & Hollunder, B. (1995), "Embedding Defaults into Terminological Knowledge Representation Formalisms", in *J. Autom. Reasoning* 14, 1: 149–180.

Bonatti, P.A., Lutz C., & Wolter, F. (2006), "Description Logics with Circumscription", in *Proceedings of KR*, pp. 400–410.

Calegari, S., & Ciucci, D. (2007), "Fuzzy Ontology, Fuzzy Description Logics and Fuzzy-OWL", in *Proc. WILF 2007* , LNCS, volume 4578.

Cordeschi, R. (1991), "The Discovery of the Artificial. Some Protocybernetic Developments", in 1930–1940. *AI & society*, 5 , 218–238.

Cordeschi, R. (2002), *The Discovery of the Artificial: Behavior, Mind and Machines Before and Beyond Cybernetics*, Springer Science & Business Media, Volume 28.

Davis E., & Marcus, G. (2015), "Commonsense Reasoning and Commonsense Knowledge in Artificial Intelligence", in *Communications of the ACM*, 58(9):92–103, 2015.

Ding, Z., Peng, Y., & Pan, R. (2006). "Bayes OWL: Uncertainty Modeling in Semantic Web Ontologies", in *Soft Computing in Ontologies and Semantic Web , Z. Ma* (ed.), Studies in Fuzziness and Soft Computing, Volume 204, Springer.

Eliasmith, C., Stewart, T. C., Choo, X., Bekolay, T., DeWolf, T., Tang, Y., & Rasmussen, D. (2012), "A Large-scale Model of the Functioning Brain", in *Science*, 338(6111), 1202-1205.

Evans, J. S.B.T., & Frankish, K. (eds.) (2008), *In Two Minds: Dual Processes and Beyond* , Oxford UP, New York, NY.

Gärdenfors, P. (2000), *Conceptual Spaces: The Geometry of Thought,* Cambridge, MA: MIT Press.

Gigerenzer, G., & Todd, P. M. (1999), *Simple Heuristics that Make Us Smart*, Oxford University Press, USA.

Giordano, L., Gliozzi, V., Olivetti, N., & Pozzato, G. L. (2013), "A Nonmonotonic Description Logic for Reasoning about Typicality.", in *Artificial Intelligence*, 195, 165–202.

Giordano, L., Gliozzi, V., Olivetti, N., & Pozzato, G.L. (2015), "Semantic Characterization of Rational Closure: from Propositional Logic to Description Logics", in *Artificial Intelligence* 226, 1–33.

Guarino, N. (1998), "Formal Onthology in Information Systems", in *Proceedings of the First International Conference* (FOIS'98), June 6-8, Trento, Italy, Volume 46. IOS Press.

Hampton, J.A. (1987),"Inheritance of Attributes in Natural Concept Conjunctions", in *Memory & Cognition* 15(1), 55–71.

Langley, P. (2012),"The Cognitive Systems Paradigm", in *Advances in Cognitive Systems*, 1, 3–13.

Laird, J. E., Newell, A., & Rosenbloom, P. S. (1987), "Soar: An Architecture for General Intelligence", in *Artificial Intelligence*, 33, 1–64.

Lieto, A. (2012), *Non Classical Concept Representation and Reasoning in Formal Ontologies*, PhD thesis, Università degli studi di Salerno.

Lieto, A. (2014), "A Computational Framework for Concept Representation in Cognitive Systems and Architectures: Concepts as Heterogeneous Proxytypes", in *Procedia Computer Science*,41, 6–14.

Lieto, A. (2019), "Heterogeneous Proxytypes Extended: Integrating Theory-like Representations and Mechanisms with Prototypes and Exemplars", in *Biologically Inspired Cognitive Architectures 2018. BICA 2018.Advances in Intelligent Systems and Computing*, Volume 848. Springer, Cham.

Lieto, A., Bhatt, M., Oltramari, A., & Vernon, D. (2018), "The Role of Cognitive Architectures in General Artificial Intelligence", in *Cognitive Systems Research*, 48, 1-3.

Lieto, A., Lebiere, C.,&Oltramari, A. (2018b), "The Knowledge Level in Cognitive Architectures: Current Limitations and Possible Developments", in *Cognitive Systems Research*, 48, 39-55.

Lieto, A., Mensa, E., & Radicioni D.P. (2016), "A Resource-Driven Approach for Anchoring Linguistic Resources to Conceptual Spaces", in *XVth International Conference of the Italian Association for Artificial Intelligence*, Genova, Italy, November 29-December 1, 2016, *volume 10037 of Lecture Notes in Artificial Intelligence*, pp. 435–449. Springer, 2016.

Lieto, A., & Pozzato, G. L. (2018), "A Description Logic of Typicality for Conceptual Combination", in *Proceedings of the 24th International Symposium on Methodologies for Intelligent Systems, ISMIS 2018*, Lecture Notes in Artificial Intelligence, Springer.

Lieto, A., & Pozzato, G. L. (2019), "A Description Logic Framework for Commonsense Conceptual Combination Integrating Typicality, Proba-

bilities and Cognitive Heuristics", in *arXiv*, preprint arXiv:1811.02366, to appear in *Journal of Experimental & Theoretical Artificial Intelligence*.

Lieto, A., Radicioni, D.P., & Rho, V. (2015), "A Common-Sense Conceptual Categorization System Integrating Heterogeneous Proxytypes and the Dual Process of Reasoning", in *Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI)*, Buenos Aires, 875-881, AAAI Press.

Lieto, A., Radicioni, D.P., & Rho, V. (2017), "Dual PECCS: a Cognitive System for Conceptual Representation and Categorization", in *Journal of Experimental & Theoretical Artificial Intelligence*, 29(2):433-452.

Lieto, A., Radicioni, D., Rho, V., & Mensa, E. (2017b), "Towards a Unifying Framework for Conceptual Representation and Reasoning in Cognitive Systems", in *Intelligenza Artificiale*, 11(2), 139-153.

Lukasiewicz, T., & Straccia, U. (2008), "Managing Uncertainty and Vagueness in Description Logics for the Semantic Web", in *Journal of Web Semantics* 6: 291–308.

McCarthy, J. (2007), "From Here to Human-level AI", in *Artificial Intelligence*, 171, 1174–1182.

Minsky, M. (1975), "A Framework for Representing Knowledge", in Patrick Winston, H. (ed.), *The Psychology of Computer Vision* , McGraw-Hill, New York.

Morgenstern, O., & Von Neumann, J. (1953), *Theory of Games and Economic Behavior*, Princeton University Press.

Newell, A. (1990), *Unified Theories of Cognition*, Harvard University Press, Cambridge, Mass.

Newell, A., Shaw, J.C., & Simon, H. A. (1959), "Report on a General Problem-solving Program*",* in *Proceedings of the International Conference on Information Processing*, pp. 256–264.

Newell, A., & Simon, H. A. (1972), *Human Problem Solving*, Englewood Cliffs, NJ: Prentice-Hall.

Newell A., &Simon, H. A. (1976), "Computer Science as Empirical Inquiry: Symbols and Search", in *Communications of the ACM*, 19(3):113–126.

Osherson, D.N., & Smith, E.E. (1981), "On the Adequacy of Prototype Theory as a Theory of Concepts", in *Cognition* 9(1) 35–58.

Quillian, R. (1968), *Semantic Memory*, Cambridge, MA: MIT Press.

Riguzzi, F., Bellodi, E., Lamma, E., & Zese, R. (2015), "Reasoning with Probabilistic Ontologies", in Yang, Q.,& Wooldridge, M. (eds.), *Proceedings of IJCAI 2015*, AAAI Press, 4310–4316.

Samuel, A.L. (1959), "Machine Learning", in *The Technology Review*, 62 (1): 42-45.

Schank, R.C. & Abelson, R.P. (1977), *Scripts, Plans, Goals and Understanding*, Erlbaum, Hillsdale, N.J.

Simon, H. A. (1947), *Administrative Behavior: A Study of Decision-making Processes in Administrative Organization*, New York: Macmillan.

Simon, H. A. (1981), *The Sciences of the Artificial*, MIT Press, Cambridge, Mass., 2nd edition: MIT Press, Cambridge, Mass.

Reiter, R. (1980), "A Logic for Default Reasoning", in *Artificial intelligence*, 13(1-2), 81-132.

Rosch, E.(1975) "Cognitive Representation of Semantic Categories", in *Journal of Experimental Psychology* 104: 573–605.

Wiener, N. (1948/1961), *Cybernetics, or Control and Communication in the Animal and the Machine*, MIT Press, Cambridge, Mass. (2nd edition: MIT Press, Cambridge, Mass., 1961).

In What Sense (If Any) Can Rationality Be Said to Be Universal?

Paolo Labinaz
University of Trieste
plabinaz@units.it

1. Introduction

This paper examines whether, and if so, how rationality can be said to be universal in light of the most recent developments in the studies on human reasoning. Over the last fifty years, a large number of psychological experiments have provided evidence that people are apparently prone to systematic biases and errors in reasoning and decision-making[1]. Inasmuch as rationality is equated with one's capacity to reason in a normatively appropriate way[2], this evidence seems to call into question the traditional characterization (dating back to Aristotle and shared by most Western philosophers) of man as the rational animal[3]. Philosophers and psychologists have worked hard to try to explain the implications of the evidence emerging from psychological experiments for our conception of human rationality[4]. While some scholars initially held that on the basis of this evidence, humans may be regarded as basically irrational, denying that

[1] For an overview, see Evans *et al.* (1993), Kahneman *et al.* [eds.] (1982), Gilcovich *et al.* [eds.] (2002).
[2] It is commonly agreed that rationality requires normativity in the sense that there must be a difference between doing a rational thing and doing an irrational one. Accordingly, any normative conception of rationality presupposes the existence of a space of alternatives in which the right thing to do always contrasts with some other actions that could have been performed: it cannot be the case that anything people might do would qualify as rational.
[3] Aristotle's standpoint on this topic is to be found in his *De Anima* (*On the Soul*). According to him, what distinguishes humans from other animal species is the fact that only humans have the capacity to think and reason, i.e. they possess a rational soul.
[4] See, e.g., Stein (1996), Samuels & Stich (2004), Rysiew (2008).

Paolo Labinaz, "In What Sense (If Any) Can Rationality Be Said to Be Universal?",
in Giorgio Grimaldi and Marialuisa Parise, Rational *Animals. Seven essays on Reason's fields of application*, pp. 75-103

rationality is an essential characteristic of human behaviour, the two leading theories in the current debate, namely the evolutionary (or ecological) and dual-process theories, maintain that once we take into consideration the origin and function(s) of reasoning, many of its instances that seem to be irrational can be judged as perfectly rational. In so doing, these theories have provided a new avenue for understanding rationality, not conceiving it as a mere conformity to norms (as is presupposed in the most influential psychological experiments on reasoning, such as the Wason selection task and the Linda problem), but as an effectiveness in achieving goals, be they personal, evolutionary, or some combination thereof. However, as I will argue, their ways of conceiving rationality run the risk of making it meaningless: while evolutionary (or ecological) theories seem to collapse rationality into a more general notion of adaptiveness, dual-process theories narrow its scope of application too much. If so, neither of them can be a good candidate to account for the universal character of rationality.

Here, my aim is to show that, insofar as we may wish to preserve the role of rationality as a defining feature of humanity, we should turn our attention to the argumentative (and hence social) function of reasoning. If reasoning can indeed be taken to be argumentative in nature, and this is what recent work in various fields of research (such as cognitive psychology, social epistemology, philosophy of logic and moral philosophy) seems to suggest[5], rationality itself should be conceived as having an essential connection with argumentation. In order to account for this connection, I will rely on some ideas about human rationality presented (in a non-systematic way) by Paul Grice[6], which Marina Sbisàhas recently presented under the label *argumentative rationality*[7]. I shall then show how this way of conceiving rationality can help us to understand the universal character of rationality, without encountering the difficulties faced by evolutionary (or ecological) and dual-process theories. In conclusion, I shall explain why relying on this "argumentative" picture of rationality may give us a deeper insight into our nature as human beings.

[5] See, e.g., Dutilh Novaes (2013), Hahn & Oaksford (2006), Hahn & Oaksford (2007), Haidt (2001), Mercier & Sperber (2011a), Mercier & Sperber (2017).
[6] Grice (1991), Grice (2001).
[7] Sbisà (2006), Sbisà (2007).

2. The standard picture of rationality

Traditionally, it has been assumed that rationality demands stable and firm standards so that we can distinguish it from subjective and changeable opinions about what is the best thing to believe or do. If no such standards are available, we have no way of adjudicating between rational and irrational instances of reasoning. With this (probably) in mind, most psychologists interested in making an empirical study of reasoning and decision-making have relied upon what Edward Stein (1996: 4) has called *the standard picture of rationality*, according to which

> to be rational is to reason in accordance with principles of reasoning that are based on rules of logic, probability theory and so forth. If the *standard picture* of reasoning is right, principles of reasoning that are based on such rules are *normative principles of reasoning*, namely they are the principles we *ought* to reason in accordance with.

Note that the *standard picture* has been widely adopted not only among psychologists, but also among philosophers and economists when studying reasoning and decision-making. But how is it that so many scholars have been attracted to this picture of rationality? There are probably two main reasons. Firstly, the *standard picture* fits well with the idea (widely accepted among educated people) that deductive logic, standard probability theory and expected utility theory prescribe how people ought to reason if they are to reason well[8]. Secondly, it captures a received idea about rationality, which is that its norms are universal and apply in any context. Deductive logic is perhaps the best-known example of universality. Depending only on its formal structure, the validity of a deductive argument is independent of the matter or meaning of the premises and a deductively valid argument is valid in every context. This has two important advantages for scholars studying reasoning empirically: (i) the universality of the norms implies that, when faced with the same reasoning problem, any rational subject must draw the same conclusion: in other words, if they are given the same premises, good reasoning can only lead to one correct conclusion; (ii) if the same norms are applicable in every context, there is no need to debate over which norms should be relied on when examining reasoning performances. Indeed, when different subjects dealing with the same reasoning problem draw

[8] By way of example, consider the following quotation from Kahneman and Tversky (1982: 493): "[t]he presence of an error of judgment is demonstrated by comparing people's responses either with an established fact (e.g., that the two lines are equal in length) or with an accepted rule of arithmetic, logic, or statistics".

different conclusions, either they are relying on different information or assumptions, or at least one of them is not reasoning in accordance with the appropriate norm. There is no other option, because a rational conclusion will always necessarily follow from a given set of premises.

If we accept the *standard picture*, and (as said above) most psychologists, as well as philosophers and economists, have more or less explicitly adopted it in studying reasoning, the right standards against which to assess people's reasoning are the norms stemming from deductive logic, standard probability theory and expected utility theory. In this sense, to the extent that people reason in accordance with these norms, they can be conceived as rational, but if they violate such norms, they should be regarded as irrational. Over the years, cognitive psychologists have collected a vast amount of evidence purporting to show that people systematically fail to reason in accordance with these norms. Think of the results of the research program based on Wason's selection task, which tests people's reasoning performances in conditional reasoning[9], or those coming from the "heuristics and biases" research program, which is concerned with probabilistic reasoning and decision-making[10].If it is true, however, that people do not systematically reason in accordance with the norms of the *standard picture*, and we assume, in line with this picture, that to be rational means to reason according to its norms, we must conclude that humans are basically irrational. In fact, some psychologists and philosophers have actually pointed out that the evidence coming from psychological studies on reasoning has "bleak implications" for human rationality[11], claiming, for example, that "[i]rrationality, not rationality, is the norm"[12], "[…] individuals are generally affected by systematic deviations from rationality"[13], "people's judgments of important probabilistic phenomena are not merely biased but are in violation of fundamental normative rules"[14], and so on.

In any event, these pessimistic claims about human rationality have been widely criticized[15]. For example, even while accepting the standard picture, some scholars have claimed that any deviation from its norms can easily be explained away by appealing to various (cognitive and non-cognitive) factors (e.g., conceiving such deviation as an error performance

[9] Wason (1966); for an overview, see Evans (2002).
[10] For an overview, see Kahneman *et al.* [eds.] (1982), Gilcovich *et al.* [eds.] (2002).
[11] Nisbett & Borgida (1975: 395).
[12] Wason (1983: 59).
[13] Bazerman & Neale (1986: 317).
[14] Slovic *et al*. (1976: 173-174).
[15] See, e.g., Adler (1991), Cohen (1981), Hilton (1995), Oaksford & Chater (1994).

or as due to a misunderstanding between experimenters and subjects). Be that as it may, there is a more fundamental question that needs to be addressed by those supporting the standard picture, regardless of their position about human rationality, and namely: why is reasoning according to its norms a goal that we as humans should pursue? To put it another way, why should reasoning in this normatively appropriate way matter to us? Indeed, according to critics of the standard picture, it is not compliance with an arbitrary set of norms that makes people's way of reasoning normatively appropriate, but instead, its effectiveness in attaining specific goals such as personal utility, reproductive success, happiness and so on. This means that conforming to some norms may be useful when it helps to attain these goals, but it is not in itself constitutive of good reasoning[16]. Therefore, the fact that people do not systematically reason in accordance with the norms derived from deductive logic, standard probability theory and expected utility theory, should not be considered as evidence of their irrationality. Moreover, norms of rationality such as having logically consistent beliefs and consistent probability assignments, behaving in such a way as to maximize one's expected utility, taking all relevant information into account when arriving at conclusions, intuitively seems to be demanding too much of human reasoners. Humans do indeed have very limited time, computational capacities and working memory, and precisely because of these limitations, cannot perform all the inferences required by the norms deriving from deductive logic, standard probability theory and expected utility theory[17]. Hence it must be open to debate whether these norms can really be considered "normative" for human beings. These are just some of the reasons – but they may well suffice for our purposes – that have led scholars interested in studying reasoning, especially psychologists, to abandon the standard picture and its norms as a benchmark for rationality. As an alternative to this picture, they have relied upon empirically informed pictures of rationality, which seem to be more sensitive to the subtlety and complexity of human reasoning. In the next two sections, I will focus in particular on the pictures of rationality provided by the two most important theories in contemporary

[16] See, e.g., Baron (2008), Samuels *et al.* (2004). As pointed out by Jonathan Baron (2008: 61), "[i]f it should turn out that following the rules of formal logic leads to eternal happiness, then it is 'rational thinking' to follow the laws of logic (assuming that we all want eternal happiness). If it should turn out, on the other hand, that carefully violating the laws of logic at every turn leads to eternal happiness, then it is these violations that we shall call 'rational'".

[17] See Cherniak (1986).

psychological studies on reasoning, and namely, evolutionary (or ecological) and dual-process theories.

3. Evolution, reasoning and rationality

If, as suggested by evolutionary psychologists[18], the human brain evolved by natural selection just like any other biological organ, then our capacity to reason, too, must be a product of evolution by natural selection. If this is true, however, what was it selected for? From an evolutionary standpoint, it seems unlikely to have happened because of some intrinsic value of that capacity. In fact, evolutionary psychologists would argue that the human capacity to reason is about as valuable as the long neck of a giraffe: both characteristics were selected because they were highly effective tools for increasing the reproductive success of humans and giraffe respectively. So how is rationality to be conceived within this evolutionary framework?

3.1. Evolution and reasoning

Before approaching the rationality question, we must take a step backwards and look more closely at how human reasoning is conceived from an evolutionary standpoint.

While agreeing that the human mind is composed of psychological adaptations, evolutionary psychologists usually differ in the ways they characterize these adaptations. Here, for the sake of simplicity, I present the characterization provided by Leda Cosmides and John Tooby[19], although I am aware that it involves certain specific assumptions (such as the massive modularity hypothesis) not widely accepted in this field of research[20]. Note,

[18] As pointed out by Leda Cosmides and John Tooby (Cosmides & Tooby, 1987; Cosmides & Tooby, 1994), probably the world's most prominent evolutionary psychologists, the main goal of evolutionary psychology is to discover the psychological adaptations of the human mind, which are the "proximate mechanisms" that cause human behaviour.

[19] Cosmides & Tooby (1987), Cosmides & Tooby (1994).

[20] In this field of research, the main competitor of Cosmides and Tooby's evolutionary approach to the human mind is probably Gerd Gigerenzer's simple heuristics research program. According to Gigerenzer (Gigerenzer, 2000; Gigerenzer, 2007), evolution has endowed humans with an "adaptive toolbox" of specialized heuristics which, while being fast and frugal, are quite effective if applied to the right range of environments. Although similar in some respects, there are significant differences between these two approaches, especially as to how they conceive the cognitive architecture of the human mind (see Carruthers, 2014).

however, that I am not interested in arguing for or against one or another evolutionary model of cognition, but in understanding how reasoning and rationality are generally conceived within the evolutionary framework. Let us return to Cosmides and Tooby's characterization, then, which conceives psychological adaptations as domain-specific cognitive modules[21]. In their view, these cognitive modules were selected because they were highly effective at solving adaptive problems, that is, recurring evolutionary problems "[...] whose solution promoted reproduction, however long or indirect the chain by which [they] did so"[22]. In particular, any given cognitive module is functionally dedicated to solving a restricted array of adaptive problems in a very restricted domain. Accordingly, if a cognitive module works well for solving a given adaptive problem belonging to one domain, it will not work equally well when applied to another adaptive problem which is similar but belongs to a different domain. Within this evolutionary framework, what we ordinarily call "reasoning" does not amount to a domain-general system, involving a unitary reasoning competence, but emerges from a set of highly specialized and effective cognitive modules, each of which evolved in response to specific recurring adaptive problems posed by the environment, be it social or physical, which our ancestors were faced with in a fairly distant past (that is, in the Pleistocene era). In this sense, what we consider to be an instance of reasoning is nothing but the result of the processing performed by these cognitive modules individually, according to the problem we are faced with. Obviously, the cognitive modules underlying reasoning work well not only when used to solve problems that were relevant in the evolutionary history of the human species, but also when applied to problem situations that go beyond their original domain of application, so long as the structure of the problem resembles that of said original domain of application.

In the last few decades, evolutionary psychologists, especially Leda Cosmides and John Tooby, have provided both theoretical considerations and empirical data in order to show that, contrary to the pessimistic claims about human rationality, people may draw normatively appropriate conclusions when the information of the reasoning problem presented to them is given in the appropriate format and when the reasoning problem itself resembles a problem from the so-called environment of evolutionary adapta-

[21] In what follows, I am not committing to any specific conception of modularity. In speaking of "domain-specific cognitive modules", I am simply referring to computational mechanisms that guide human behavior in different ways appropriate to specific domains.
[22] Cosmides & Tooby (1994: 87).

tion[23]. More generally, from an evolutionary standpoint, whenever classic norms of rationality and subjects' responses in a reasoning task do not agree, we should prefer the latter to the former. For example, if people do not reason in accordance with norms derived from deductive logic, this does not mean that they are bad at logical reasoning, but that they are not equipped with a cognitive module devoted to deductive reasoning. This is probably due to the fact that having a cognitive module of this kind could not have increased our ancestors' chances of survival and reproductive success. In this sense, there are no other standards than those which come from cognitive capacities because a cognitive mechanism that operates normally is operating as it ought to operate. More specifically, the "ought to" should always be understood in evolutionary terms, in the sense that a properly operating cognitive module should be regarded as solving the problems for which it evolved: such is the fundamental form of normativity.

3.2. Rationality as adaptiveness

Within this evolutionary framework, rationality is to be conceived as the effect of the "normal" functioning of our cognitive modules. In particular, it is their adaptive design that makes us rational. This means that humans are rational simply because they possess a large amount of cognitive modules that are well adapted to the structure of the environment in which they normally operate. From an evolutionary standpoint, these modules gave a reproductive advantage to those of our ancestors who were endowed with them and so (in the long run) went to fixation in early human populations. If we go back now to the initial question, namely in what sense (if any) rationality can be said to be universal, ecological (or evolutionary) theories have a clear and definite answer. Indeed, inasmuch as evolution has endowed all humans with these cognitive modules, they are universal, species-typical cognitive traits. They can therefore be conceived as defining features of our humanity. For example, Tooby and Cosmides (1990: 19) solemnly state that cognitive modules are "[...] the psychological universals that constitute human nature"[24].

[23] See especially Cosmides (1989), Cosmides & Tooby (1996), Fiddick *et al.*(2000).

[24] This does not mean that there cannot be behavioral and attitudinal differences among humans, but only that these differences do not reflect differences in the underlying psychological adaptations, which are instead cognitive invariants of the human mind (see Tooby & Cosmides, 1992: 45).

Though attractive, this evolutionary explanation runs the risk of making rationality meaningless. Indeed, appealing to the adaptive value of cognitive modules in order to account for human rationality is rather problematic. In particular, the shift from the adaptive design of cognitive modules, which is claimed to make them "ecologically" rational, to the rationality of their possessors (that is, human beings) leads to some undesirable consequences. Let us consider the first: remember that cognitive modules are said to be adaptive because they lead to successful outcomes when applied to the problems for which they evolved. From an evolutionary standpoint, the success is measured in terms of reproductive success or inclusive fitness[25]: in other words, the design of a cognitive module is adaptive insofar as its operations lead to successful outcomes that have been conducive, directly or indirectly, to reproductive success (or inclusive fitness) of its possessor. However, if rationality is defined in terms of success, and success is understood as the tendency to increase reproductive success or inclusive fitness, directly or indirectly, we have no way of distinguishing between processes that are supposed to make us rational and those that are not in the business of rationality. Think of our stomachs[26]. Our stomachs, like our cognitive modules, are well adapted to solving certain recurring adaptive problems, that is, metabolic and cognitive problems, respectively, and the fact that we are endowed with them means that both have proved to be highly efficient in this respect. In other words, we are equipped with them because of the benefits their adaptive design provided to our ancestors in terms of reproduction and still provide to us today. But if rationality is characterized as described above, there would be nothing wrong if someone felt entitled to say that we are rational because of the adaptive design of our stomachs. The same may well be said by appealing to the adaptive design of any other organ of our body. This way of conceiving rationality thus deprives it of all its explanatory power. Moreover, there is a second, highly undesirable consequence in characterizing rationality in adaptive terms. If we assume that it is thanks to the adaptive design of our cognitive modules that we are rational, then it is unclear why no other animal species can be considered to be endowed with rationality. After all, just like us, they are the product of evolution by natural selection, and in this sense are endowed with psychological adaptations that are no doubt highly effective in solving the adaptive

[25] Inclusive fitness can be (roughly) characterized as the number of genes one transmits to subsequent generations as a result of direct and indirect reproduction (for a technical explanation, see Hamilton, 1963).

[26] With some modifications, this example is borrowed from Chater and Oaksford (2002: 141).

problems faced by their ancestors. The adaptive design of their cognitive modules therefore makes them rational like us.

To sum up, if we follow the evolutionary strategy, rationality becomes one among many other features that we possess, and that we may share with other animal species.

4. Two minds, one rationality?

As seen in the previous section, to describe the mind, as evolutionary psychologists do, as operating exclusively on the basis of evolutionarily adaptive cognitive mechanisms, does not really help us to shed light on the nature of human rationality. As an alternative to this model of the mind, some cognitive psychologists and philosophers have developed a family of dual-process theories, which divide mental processes into two general categories, depending on how they operate. The question is whether we can improve our understanding of rationality, specifically as a defining feature of our humanity, by conceiving the human mind in this way.

4.1. Dual-process theories of reasoning

Dual-process theories, which began in the nineteen-seventies with the works of Wason and Evans (1975), have been developed more recently in different ways by various scholars[27]. While differing as to details, their theories agree on a set of distinctive features deemed to characterize two categories of processes operating in the human mind. The first category comprises those processes labelled as "type 1 processes" which are typically fast, automatic, often unconscious (that is, occurring at a sub-personal level), relatively inflexible and require relatively little cognitive effort. Note that, due to their automatic, effortless nature, they are supposed to guide most of our actions and judgements. The most commonly cited example of this kind of processes is domain-specific modules, as well as Gigerenzer's fast and frugal heuristics[28]. Dual-process theorists regard type 1 processes as universal among humans, but also as shared to a certain extent with other animal species. In contrast, processes belonging to the second category, known as "type 2 processes", are more "reflective", slow at processing information

[27] See, e.g., Evans (2007), Evans (2010), Evans & Over (1996), Kahneman & Frederick (2002), Stanovich (1999), Stanovich (2004), Stanovich (2011).
[28] See footnote 20.

and often rule-based; moreover, they require the allocation of significant cognitive resources, operate consciously and, because of their malleability, can differ considerably between individuals, as well as among cultures and societies[29]. More generally, type 2 processes are deemed to have evolved recently, and unlike type 1 processes, are an exclusively human property. In particular, according to Keith Stanovich (2004: 47), they allow us "[…] to sustain the powerful context-free mechanisms of logical thought, inference, abstraction, planning, decision-making, and cognitive control". Similarly, Jonathan Evans (2007) holds that the function of these processes is to make possible hypothetical thinking, which is fundamental for high-order cognitive activities such as reasoning and decision-making.

More recently, in a development of his dual model of cognition, Stanovich (2011) has made a distinction within the second category between low-level, algorithmic processes, which are part of what he calls "the algorithmic mind", and higher-level reflective processes characterizing what he calls "the reflective mind". While the reflective mind is concerned with "[...] higher-level goal states and epistemic thinking dispositions"[30], which exercise reflective control over our reasoning, the algorithmic mind corresponds to the machinery which lends its support to the reflective-level activities[31].

4.2. Rationality, processes and goals

Although dual process theories are psychological theories, some of their proponents have been closely engaged in exploring what their theories might say about human rationality[32]. Here I shall focus on Keith Stanovich's analysis, which is probably the most prominent application of dual-process theories to issues concerning rationality. According to him, we can distinguish (in principle) between two senses of rationality, both of which are characterized in terms of optimization. Where they differ is in the goals against which their degree of optimization can be measured. While the first

[29] For empirical data showing individual differences in people's abilities to solve reasoning tasks, see Stanovich & West (1998), Stanovich *et al*. (2011). With regard to socio-cultural differences in reasoning strategies, Richard Nisbett and his collaborators claim to have provided evidence that Westerners and East Asians think and reason differently when faced with certain cognitive tasks (see, e.g., Nisbett *et al.,* 2001; Nisbett, 2003). According to Nisbett (2003: xx), this is not due to genetic differences, but depends on "[...] different social relations, views about the nature of the world, and characteristic thought processes".

[30] Stanovich (2011: 32).

[31] See Stanovich (2011: 47-60).

[32] See, e.g., Evans & Over (1996), Stanovich (1999), Stanovich (2004), Stanovich (2011).

sense is connected with the achievement of adaptive goals, and is therefore measured in terms of fitness maximization, the second one is concerned with the achievement of "personal" goals, characterizing rationality as utility maximization[33]. According to Stanovich, however, a type of behaviour is fully rational only when it optimizes one's goals at a personal level, that is, according to the second sense described above: this happens when type 2 processes override type 1 processes. Evolutionary adaptive behavior, on the other hand, can be judged as being more or less efficient, but in itself does not display rational or irrational properties[34]. This is because "[r]ationality concerns the actions of an entity in its environment that serve its goals"[35], and not evolutionary adaptive behavior caused by type 1 processes[36]. As Stanovich rightly observes, we cannot reduce our being rational to the "normal" functioning of type 1 processes because it is only at the personal level that issues of rationality come into play. In fact, according to him, rationality needs to be defined "[…] over a person, not a subpersonal entity (the genes)"[37]. In particular, humans manifest their rationality only when they behave in ways that maximize their chances of fulfilling their personal goals (given the resources available, whether they be cognitive or physical).

Although this way of conceiving rationality appears to be promising, there is a fundamental limitation to Stanovich's distinction between personal and adaptive goals. In particular, if we follow Stanovich's strategy (but this can be extended to include other analyses proposed by dual-process theorists), the field of application of rationality runs the risk of being too restricted. Indeed, insofar as one is considered to be rational only if one acts on the basis of type 2 processes, rationality becomes a property of a certain kind of process, and not of the individual as a whole. But this is open to question. There can indeed be situations (as we shall see) where we can act rationally even if type 2 processes do not override type 1 processes. Note that I am not claiming that following type 1 processes is rational *per se*, but that in order to achieve full rationality we need to understand how type 1

[33] Stanovich (1999: 148-152), Stanovich (2004).
[34] As pointed out by Stanovich (2011: 18), the main error of evolutionary psychologists is to equate "rational" with "evolutionary adaptive".
[35] Stanovich (2011: 59).
[36] On this point, Stanovich is in agreement with other two prominent dual-process theorists, namely Jonathan Evans and David Over. According to them, "[…] the starting point for any understanding of human rationality must be behavioural: we must ask how decisions taken and actions performed serve the goals of the individual" (Evans & Over, 1996: 1).
[37] Stanovich (2004: 86).

and type 2 processes can work well together. Even Stanovich is conscious of that: indeed, he claims that people should feel the need for rational integration by eliminating the conflict between type 1 and type 2 processes[38]. But in his opinion, this integration should occur across our preference hierarchies. If so, in presence of a conflict between type 1 and type 2 processes, what is rationally required of an individual is to favor the latter over the former. We might easily conceive of cases, however, in which type 1 processes can maximize one's personal utility. Think of an anorexia patient. If we consider her individual preferences, what she wants is to become leaner and thinner, although she is probably lean and thin already. In order to achieve this goal, the best thing to do (among other things) is to eat as little as possible. Fulfilling this particular goal, however, involves setting aside other fundamental ones, such as being healthy, regardless of whether she is conscious of that. Fortunately, like us, she has an unconscious tendency to feel hungry, and this tendency may stimulate her to look for food. Obviously, looking for food goes against her individual preferences, but is conducive to maintaining her health. In this case, then, following type 1 processes seems to be the rational choice to make, that is, the choice which maximizes the anorexic's personal utility, even if it goes against her individual preferences, and it would be strange if someone were to consider it a case of irrationality due to the fact that type 1 processes have not been overridden by type 2 processes. It follows that it is not quite so obvious that type 1 processes cannot maximize one's personal utility, or that all type 2 processes really do serve the interests of the individual as a whole. Indeed, one can obtain benefits, whatever they may be, from both kinds of processes, as well as from their interplay: if so, rationality requires more than simply overriding our type 1 processes.

To sum up, while it is fundamental to assume, as dual process theorists do, that rationality pertains to the personal level, their attempt (exemplified here by Stanovich) to remove type 1 processes from the domain of rationality is unsuccessful: if we were to follow their strategy, rationality would have to be treated as a property of a certain kind of process (that is, type 2 processes) and not of the individual as a whole. At that point, however, it is unclear in what sense (if any) it can be taken to be a global feature of our humanity.

[38] Stanovich (2011: 81-94).

5. Reasoning and argumentation

So far, I have examined the pictures of rationality provided by the two leading theories in contemporary psychological studies on reasoning, and namely evolutionary (or ecological) and dual-process theories, respectively. More specifically, I have evaluated whether they might be good candidates in order to account for the universal nature of rationality, and I have reached the conclusion that they are not. In what follows, I would like to take a different route, considering the argumentative (and hence social) function of reasoning. Recent work in various fields of research (such as cognitive psychology, social epistemology, philosophy of logic and moral philosophy) seems to suggest, from both the theoretical and experimental point of view, that reasoning is argumentative in nature. There are four theories in particular that have contributed to the development of this new perspective on reasoning and its relationship with argumentation. Let us consider each of them briefly. In the interdisciplinary field of cognitive pragmatics, Hugo Mercier and Dan Sperber[39] have proposed the so-called "argumentative theory of reasoning", according to which reasoning has evolved so as to devise and evaluate arguments in communicative contexts. Ulrike Hahn and Mike Oaksford[40] have developed a Bayesian approach to argument strength that aims to explain from an argumentative point of view why people are prone to biases and fallacies. In the context of the history of logic, Catarina Dutilh Novaes (2013) has argued for the dialogical nature of deductive reasoning. According to her, our ability to reason develops thanks to the dialogical and argumentative practices that we learn, particularly in the context of schooling. Finally, in supporting an intuitionist model of moral judgment, Jonathan Haidt (2001) has held that people reason to find arguments to justify (or rather, to provide post-hoc justifications for) their moral intuitions.

For reasons of space, I shall focus here only on the most representative of these theories, and namely Mercier and Sperber's argumentative theory of reasoning. Interestingly enough, their theory is developed within an evolutionary framework, but in contrast with other evolutionary theories it acknowledges that reasoning has evolved for a specific function, which is that of devising and evaluating arguments in communicative contexts (as said above). In particular, according to Mercier and Sperber, reasoning is to

[39] Mercier & Sperber (2011a), Mercier & Sperber (2017).
[40] Hahn & Oaksford (2006), Hahn & Oaksford (2007).

be conceived as "the ability to produce and evaluate reasons"[41]. Within this evolutionary framework, reasoning is a function (and probably the primary one) of an evolved, cognitive module that deals with reasons and their relationship with the claims they purportedly support. While the output of this module is an intuitive representation about whether a certain claim is well supported by the reasons provided to accept it, acceptance or rejection of that claim is done reflectively in light of the strength of this intuitive representation, and herein lies the defining feature of reasoning: unlike all the other cognitive processes that occur without any attention being paid to reasons, reasoning deals essentially with reasons, and probably this makes it an exclusive property of humans[42]. As we shall see, this observation has very interesting implications for our understanding of human rationality and its universal character. Indeed, insofar as we possess the cognitive module mentioned above, and (if we follow Mercier and Sperber's line of thought) if this is something that only humans possess, as humans, we have a special attention for, and interest in reasons. This is a very intriguing idea which deserves to be developed more fully, and I will elaborate further on it in the next section. Before moving on, though, I must give substance to Mercier and Sperber's claims by providing a more detailed presentation of their theory, as well as of the evidence they offer in support of it. In their view, "the emergence of reasoning is best understood within the framework of the evolution of human communication"[43]. Indeed, whether in the Pleistocene epoch or today, there is a fundamental problem in communicative contexts due to the fact that the interests of communicators and receivers usually diverge. If communication has evolved, however, then it should have been advantageous to both communicators and receivers. As pointed out by Mercier and Sperber (2011b: 96),

> What makes communication advantageous to receivers is that it provides them with rich information that they could not, or not easily, have obtained on their own. For this, the information they receive has to be genuine information; that is, close enough to truth. What makes communication advantageous to communicators is that it allows them to achieve some desirable effect in the receivers. For this, the information they emit has to be conducive to this effect, whether it is true or false.

It is thanks to reasoning if communication is more reliable and hence more advantageous for communicators and receivers: while it enables

[41] Sperber & Mercier (2012: 381).
[42] See Mercier & Sperber (2009).
[43] Mercier & Sperber (2011a: 60).

communicators to produce arguments to convince receivers (especially when receivers are unlikely to take the communicated information on trust), receivers can use reasoning to evaluate these arguments, thereby minimizing the risk of being misled. If this is true then, reasoning evolved not to help us get better at thinking on our own, but to justify our thoughts and actions to our interlocutors in order to convince them to think and act as we suggest, as well as to assess the reasons they provide so as to defend ourselves from their attempts to mislead us. Accordingly, reasoning should work better when used in dialogical situations, particularly in group discussions, where participants offer each other arguments in support of their viewpoints with the aim of convincing one another, compared to when one is thinking on one's own (and does not take into account other aspects or perspectives) or taking part in non-deliberating groups, in which people are not interested in comparing their opinions with those of their interlocutors. In fact, various experimental results presented by Mercier and Sperber (2011a: 61-71) collected from different fields of research (such as social and developmental psychology, as well as studies on informal reasoning) seem to confirm that people reason better in situations where they are part of groups which reason, discuss or argue collectively. In particular, reasoning tasks in which subjects taken in isolation tend to give wrong answers are more frequently solved correctly when they are approached in groups and discussed collectively. For example, as reported by Mercier and Sperber (2011a: 61), when taken in isolation, subjects are not usually able to apply or recognize the well-known *modus tollens* (if *p* then *q*, not-*q*, so not-*p*), but when engaged in argumentative contexts, they easily recognize and apply this argumentative form in order to criticize the claims of their opponents[44]. When they discuss how to solve it with others, even subjects performing the standard version of the Wason selection task for the first time give the correct response in about 70% of cases[45]. This is probably due to the fact that in the case of solitary reasoning, people tend to accumulate evidence in support of their pre-existing opinions or hypotheses and fail to look for reasons that may question them, while in dialogical situations, when confronting each other's arguments, people are able to improve their argumentative abilities by collecting new evidence, anticipating counter-arguments, taking into consideration the other person's viewpoint, and so on.

[44] See, e.g., Thompson *et al*. (2005).
[45] Moshman & Geil (1998).

6. Argumentative rationality and universality

If Mercier and Sperber are right and the primary function of reasoning is argumentative, and there are good empirical (as well as theoretical) reasons to suppose that this is so (as seen above), rationality itself should be conceived as having an essential connection with argumentation. However, when we look at Mercier and Sperber's works, there is no suggestion as to how rationality should be characterized in light of this connection. As pointed out above, all we can surmise is that, according to them, as humans we have a special attention for, and interest in, reasons due to the fact that we are endowed (apparently uniquely among the animal species) with a specific module devoted to reasons and to their relationship with the claims they purportedly support[46]. But we might then wonder why it is that people have this special interest in reasons. On the one hand, our interest in reasons might depend on their effectiveness in changing people's minds, regardless of whether they are good or not. In conceiving our interest in reasons in this way, one presupposes an instrumental conception of rationality, according to which giving reasons in support to one's thoughts and actions is not valuable *per se* but only insofar as it succeeds in convincing other people to do or think something. Obviously, the ultimate goal is not to convince other people to do or think something, but to achieve one's personal utilities by means of persuading them. This places the emphasis on the persuasive function of reasoning. On the other hand, our interest in reasons can be conceived differently if we focus on the reason-giving function of reasoning and here, the central notion may be accountability, rather than personal utility. Philip E. Tetlock (together with Jennifer Lerner), who first defined this term, characterizes it as "the implied or clearly expressed expectation that one will be called upon to justify one's beliefs, feelings or actions to others"[47]. It is thanks to reasoning that we can give reasons to explain and justify ourselves. More specifically, insofar as we are held to be accountable for our beliefs, feelings or actions, reasoning enables us to motivate and justify these beliefs, feelings and actions to others, thereby letting people know what to expect of us.

In my view, reducing the argumentative function of reasoning to the production of persuasive arguments does not really explain much about our

[46] Mercier has provided a great deal of evidence (coming from experimental studies in which he has been engaged, as well as from those of other groups of research) to show that the ability to produce arguments is universal among humans and that it develops early on in life (see, respectively, Mercier, 2011a and Mercier, 2011b).

[47] Lerner & Tetlock (2003: 434).

interest in reasons[48]. We can provide reasons for a variety of motivations, such as giving sense to our speech or actions, making explicit the premises of what we say or do, collaborating with others, or even improving our self-image. Of course, we can also give reasons to persuade others. What is not very clear, however, is why the argumentative function of reasoning should be considered as primarily persuasive. Actually, our ability to reason not only seems to be independent of our ability to produce persuasive arguments, but it is the latter that depends in some way on the former. In fact, devising arguments intended to persuade is only one of various things that we can do thanks to our ability to reason. Paradoxically, one may be interested in giving reasons in support of a certain claim or decision whenever one thinks it is relevant to do so, without any particular ulterior motive (such as that of persuading others of something). Accordingly, the instrumental conception of rationality does not appear suitable to account for the argumentative nature of reasoning.

We can now turn to the reason-giving function of reasoning. Insofar as we want to consider this function of reasoning as being more fundamental than its alleged persuasive function, we need to explain how and why giving reasons in support to our thoughts and actions should be valuable *per se*. This calls for a thoroughly argumentative conception of rationality. One solution may come from a conception of rationality inspired by Paul Grice's later works[49], which Marina Sbisà (2006: 241) has recently presented under the label *argumentative rationality*[50].

According to this conception, as characterized by Sbisà (2006: 241-242), rationality amounts to "[…] a concern that one's moves are justified and a capacity (to some degree) to give effect to that concern"[51]. What is interesting here is the rather unusual distinction (at least among psychological theories of reasoning) between people's concern for justifying, say, their action and thoughts, and their actual capacity to give effect to that concern, that is, their capacity to reason. When dealing with issues of rationality, scholars such as those supporting evolutionary (or ecological) and dual-

[48] In a previous work, I have argued that the argumentative theory of reasoning sticks to the persuasive function of reasoning and so appears to presuppose an instrumental conception of rationality (see Labinaz, 2014). In their recent book *The enigma of reason* (2017), Mercier and Sperber appear to give more weight to the reasoning-giving function of reasoning. However, they do not seem to assume any specific notion of rationality in support of this view.

[49] See Grice (1991), Grice (2001).

[50] See also Sbisà (2007).

[51] See Grice (1991: 82-83).

process theories focus almost exclusively on the second aspect. As highlighted above, while some reduce our being rational to the proper functioning of some evolved cognitive modules, others describe it as the effectiveness of our type 2 processes. What they both leave aside is the role of motivations in reasoning, which corresponds to the first aspect of the characterization of rationality presented above. Indeed, being able to provide reasons supporting one's thoughts or actions is not enough: one also needs to care about actually having such reasons; it needs to matter to us to provide reasons in support of our thoughts or actions. In other words, without motivations, our ability to reason is neither useful nor relevant to us[52]. Rationality therefore involves not only the capacity to account for what one says or does, but also one's motivation to do so. Let me explain this point in more detail. If you ask yourself why you engage in reasoning, you probably will not have a clear and definite answer. But it would be very strange if you were to say that you engage in reasoning because of its effectiveness in persuading others. You might say instead that you care about having reasons for your thoughts and actions, probably because since we tend to be held accountable for our thoughts and actions (as pointed out above), we find it valuable to provide reasons in support of what we think and do. One might suppose, then, that it is in order to satisfy this requirement that our ancestors were endowed with a capacity to deliver reasons in support of their thoughts and actions. From this perspective, our ability to reason is a response to a very human concern to provide reasons to our interlocutors in support of what we say and do. While this ability can be exploited to persuade others to think or do something (thus becoming part of a more complex social situation which may be adversarial but also cooperative), its primary function could be deemed to be communicative. To sum up, being rational here amounts to being a creature that can be acknowledged to act and think on the basis of reasons because of its concern for justifying its own thoughts and actions: for this reason, we expect each other to be accountable for our own thoughts and actions.

If we go back now to the initial question, namely in what sense (if any) rationality can be said to be universal, it seems to me that the argumentation-based picture of rationality proposed enables us to account for its universal character without reducing it merely to adaptiveness or effectiveness in achieving one's personal goals. Although it is our concern for justifying our thought and actions that lies at the core of this picture of

[52] Of course, the opposite is also true: if we have no ability to reason, then we cannot give effect to our concern for justifying our thoughts and actions.

rationality, which can considered to be universal among humans, this concern needs to be supported by some capacity to actually provide justifications (as said above). Note that this capacity can vary among humans, without undermining the universal character of their rationality. As pointed out by Mercier and Sperber (2017: 286),

> [t]o say that reason is a universal mechanism does not imply that it works in exactly the same manner in all places or that specific societies can affect reason only in superficial, merely cosmetic ways. Reason is deployed in a variety of cultural practices and institutions that may inhibit some of its uses, enhance others, and provide socially developed cognitive tools that complement and extend naturally evolved abilities.

Clearly, even the concern for justifying one's thoughts and actions can differ among different individuals, groups and societies, depending on the attitudes, dispositions and motivations stemming from one's own personal experience, as well as from one's engagement in certain social practices or institutions. Be that as it may, the fact remains that as humans, we have in common an attention for (and interest in) reasons, regardless of the undeniable differences in our attitudes, dispositions and motivations.

Let me now explain why the proposed argumentation-based picture of rationality does not encounter the difficulties faced by evolutionary (or ecological) and dual-process theories. As seen above, in contrast with evolutionary (or ecological) theories, being rational here is not reduced to the normal functioning of some evolved cognitive modules. As argued in Section 3.2, every animal species is endowed with psychological adaptations that are undoubtedly highly effective in solving the adaptive problems faced by their ancestors, but it would be strange if we were then to conclude that this makes their members rational like us. Moreover, even though evolution endowed us with a well-functioning ability to reason, we need something more, namely, an attitude of willingness, or a certain cognitive disposition, to use this ability in order to provide reasons for what we think or do. In other words, as also pointed out by Stanovich, rationality should be linked to the personal level, because it is at this level that issues of rationality come into play. If we now turn to dual-process theories, while agreeing that rationality is linked to the personal level, the argumentative picture of rationality disagrees with them on where rationality lies, and in my view, this difference is crucial. Indeed, insofar rationality lies in one's concern to justify one's thoughts and actions, and not in the effectiveness of some cognitive processes (specifically, type 2 processes) in supporting the achievement of one's personal goals, it becomes clear that being rational

does not depend completely on one's actual behavior. In Stanovich's framework, the fact that we do not achieve our goals, or that a certain behavior performed is not effective as we might think, suggests that type 2 processes have not overridden type 1 processes, and this makes that behavior irrational (at least, to a certain extent). Instead, failures to provide good arguments do not by themselves cancel out our original concern or intention to justify our thoughts or actions, because one's concern or intention involves something more than actual behavior: they require "[…] a certain first-person perspective on what is to be achieved or to be done"[53]. So even while pertaining to the personal level, rationality here does not refer to expected utility maximizing behavior but focuses on what is expected to be achieved or to be done from humans as we are and, as said above, we expect each other to be accountable for our own thoughts and actions. In this sense, for each of us, what is to be achieved or done becomes a task that our fellow human beings expect us to carry out, regardless of whether we perform it by providing "genuine" reasons or by constructing post-hoc justifications. As Marina Sbisà has pointed out (2007: 92), "conceiving of rationality as 'argumentative' amounts to claiming that we are constituted as rational beings in the first place by the fact that we accept it as a task that we have to justify our moves". In this sense, rationality cannot be considered as a given or a feature that we as humans do or (for an irrationalist) do not possess, but as a task that we should take seriously because it reflects what our fellows expect from us (but not from the members of any other animal species) and namely, that of being accountable for our own thoughts and actions.

7. Concluding remarks

The goal of this paper was to examine whether, and if so how, rationality can be said to be universal in light of the most recent developments in the studies on human reasoning. I first examined whether the two leading theories in contemporary psychological studies on reasoning, and namely evolutionary (or ecological) and dual-process theories, might be good candidates in order to account for its universal character, and I reached the conclusion that they are not. If we were to follow the evolutionary strategy, rationality would become one among many other features that we possess, and that we may share with other animal species. Instead, in the framework of dual process theories, rationality coincides with the effectiveness of type 2 pro-

[53] Sbisà (2007: 92).

cesses in supporting the achievement of one's personal goals, making it a property of those processes, and not of the individual as a whole. I then turned to the hypothesis, supported by scholars from various fields of research, that reasoning is argumentative in nature, suggesting that there are good empirical and theoretical reasons for accepting it. In the final section, I proposed an argumentation-based picture of rationality which, by focusing on the reason-giving function of reasoning, seems to be able to account for its universal character without reducing it merely to adaptiveness or effectiveness in achieving one's personal goals. According to this perspective, rationality is not a given or a feature that we do or do not possess, but is instead conceived as a task that we as humans consider as worth pursuing. Of course, there is still much work to be done to provide a satisfactory analysis of rationality in accordance with this argumentative-based perspective: more specifically, some crucial issues still need to be tackled, particularly how to define and attribute capacity to provide justifications. For the moment, however, I hope at least to have shown that conceiving of rationality from this alternative perspective may help us, in the words of Nozick (1993: 181) to see it as something that "[…] enables us to transform ourselves and hence transcend our status as mere animals" and not as the "normal" functioning of the human mind or the effectiveness of a sub-part of it. It is our concern for having reasons for what we think and do and the interpersonal commitment to offering them in support of our thoughts and actions that enables us to transform ourselves: that is what makes rationality a distinctively human attribute, and something that sets us apart from the other animal species.

Funding: This work was supported by the Croatian Science Foundation under grant number IP-06-2016-2408.

References

Adler, J.E. (1991), "An Optimist's Pessimism: Conversation and Conjunctions", in *Probability and Rationality: Studies on L. Jonathan Cohen's Philosophy of Science*, ed. by E. Eells & T. Maruszewski, Amsterdam, Editions Rodopi, pp. 251-282.

Aristotle (1961), *De Anima*, edited, with introduction and commentary by W.D. Ross, Oxford, Clarendon Press.

Baron, J. (2008), *Thinking and Deciding*, 4th ed., Cambridge, Cambridge University Press.

Bazerman, M.H., & Neale, M.A. (1986), "Heuristics in Negotiation: Limitations to Effective Dispute Resolution", in *Judgment and Decision Making: An Interdisciplinary Reader*, ed. by R. Arkes & R.R. Hammond, Cambridge, Cambridge University Press, pp. 311-321.

Carruthers, P. (2014), "The Fragmentation of Reasoning", in *La Coevolución de Mente y Lenguaje: Ontogénesis y Filogénesis*, ed. by P. Quintanilla, Lima, Fondo Editorial de la Pontificia Universidad Católica del Perú. Available at: http://faculty.philosophy.umd.edu/pcarruthers/The%20Fragmentation%20of%20Reasoning.pdf (accessed September 5, 2017).

Chater, N., & Oaksford, M. (2002), "The Rational Analysis of Human Cognition", in *Reason and Nature: Essays in the theory of rationality*, ed. by A. Millar & J.L. Bermúdez, Oxford, Oxford University Press, pp. 135-174.

Cherniak, C. (1986), *Minimal Rationality*, Cambridge (MA), The MIT Press.

Cohen, L.J. (1981), "Can Human Rationality Be Experimentally Demonstrated?", in *Behavioral and Brain Sciences*, 4, pp. 317-331.

Cosmides, L. (1989), "The Logic of Social Exchange: Has Natural Selection Shaped How Humans Reason? Studies with Wason Selection Task", in *Cognition*, 31, pp. 187-276.

Cosmides, L., & Tooby, J. (1987), "From Evolution to Behavior: Evolutionary Psychology as the Missing Link", in *The Latest on the Best: Essays on Evolution and Optimality*, ed. by J. Dupré, Cambridge (MA), The MIT Press, pp. 277-306.

Cosmides, L., & Tooby, J. (1994), "Origins of Domain Specificity: The Evolution of Functional Organization", in *Mapping the Mind*, ed. by L. Hirschfeld & S. Gelman, Cambridge, Cambridge University Press, pp. 85-116.

Cosmides, L., & Tooby, J. (1996), "Are Humans Good Intuitive Statisticians after All? Rethinking Some Conclusions from the Literature on Judgment under Uncertainty", in *Cognition*, 58, pp. 1-73.

Dutilh Novaes, C. (2013), "A Dialogical Account of Deductive Reasoning as a Case Study for How Culture Shapes Cognition", in *Journal of Cognition and Culture*, 13, pp. 453-476.

Evans, J.St.B.T. (2002), "Logic and Human Reasoning: An Assessment of the Deduction Paradigm", in *Psychological Bulletin*, 128, pp. 978-96.

Evans, J.St.B.T. (2007), *Hypothetical Thinking: Dual Processes in Reasoning and Judgement*, Hove (UK), Psychology Press.

Evans, J.St.B.T. (2010), *Thinking twice: Two minds in one brain*, Oxford, Oxford University Press.

Evans, J.St.B.T., & Over, D.E. (1996), *Rationality and Reasoning*, East Sussex (UK), Psychology Press.

Evans, J.St.B.T., Newstead, S.E., & Byrne R.M. (1993), *Human Reasoning: The Psychology of Deduction*, Hillsdale (USA), Erlbaum.

Fiddick, L., Cosmides, L., & Tooby, J. (2000), "No Interpretation without Representation: The Role of Domain-specific Representations in the Wason Selection Task", in *Cognition*, 77, pp. 1-79.

Gigerenzer, G. (2000), *Adaptive Thinking. Rationality in the Real World*, New York, Oxford University Press.

Gigerenzer, G. (2007), *Gut Feelings: The Intelligence of the Unconscious*, New York, Viking Press.

Gilovich, T., Griffin, D., and Kahneman D. (eds.) (2002), *Heuristics and Biases: The Psychology of Intuitive Judgment*, New York, Cambridge University Press.

Grice, P.H. (1991), *The Conception of Value*, Oxford, Oxford University Press.

Grice, P.H. (2001), *Aspects of Reason*, Oxford, Oxford University Press.

Hahn, U., & Oaksford, M. (2006), "A Bayesian Approach to Informal Reasoning Fallacies", in *Synthese*, 152, pp. 207-23.

Hahn, U., & Oaksford, M. (2007), "The Rationality of Informal Argumentation: A Bayesian Approach to Reasoning Fallacies", in *Psychological Review*, 114, pp. 704-732.

Haidt, J. (2001), "The Emotional Fog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment", in *Psychological Review*, 108, pp. 814-834.

Hamilton, W.D. (1963), "The Evolution of Altruistic Behaviour", in *American Naturalist*, 97, pp. 354-356.

Hilton, D.J. (1995), "The Social Context of Reasoning: Conversational Inference and Rational Judgment", in *Psychological Bulletin*, 118, pp. 248–271.

Kahneman, D., & Frederick, S, (2002), "Representativeness Revisited: Attribute Substitution in Intuitive Judgement", in *Heuristics and Biases: The Psychology of Intuitive Judgment*, ed. by T. Gilcovich, D. Griffin & D. Kahneman, New York, Cambridge University Press, pp. 49-81.

Kahneman, D., & Tversky, A. (1982), "On the Study of Statistical Intuitions", in *Cognition*, 11, pp. 123-41; reprinted in *Judgment under Uncertainty: Heuristics and Biases*, ed. by D. Kahneman, P. Slovic, & A. Tversky, New York, Cambridge University Press, pp. 493-508.

Kahneman, D., Slovic, P., & Tversky, A. (eds.) (1982), *Judgment under Uncertainty: Heuristics and Biases*, New York, Cambridge University Press.

Labinaz, P. (2014), "Reasoning, Argumentation and Rationality", in *Etica e Politica*, 16, pp. 576-594; reprinted in *Moral Realism and Political Decisions. Practical Rationality in Contemporary Public Contexts*, ed. by G. De Anna & R. Martinelli, Bamberg, University of Bamberg Press, pp. 81-103.

Lerner, J.S. & Tetlock, P.E. (2003), "Bridging Individual, Interpersonal, and Institutional Approaches to Judgment and Choice: The Impact of Accountability on Cognitive Bias", in *Emerging Perspectives on Judgment and Decision Research*, ed. by S. Schneider & J. Shanteau, Cambridge, Cambridge University Press, pp. 431-457.

Mercier, H. (2011a) "On the Universality of Argumentative Reasoning", in *Journal of Cognition and Culture*, 11, pp. 85-113.

Mercier, H. (2011b), "Reasoning Serves Argumentation in Children", in *Cognitive Development*, 26, pp. 177-191.

Mercier, H., & Sperber, D. (2009), "Intuitive and Reflective Inferences", in *In Two Minds*, ed. by J.St.B.T. Evans & K. Frankish, New York, Oxford University Press, pp. 149-170.

Mercier, H., & Sperber, D. (2011a), "Why Do Humans Reason? Arguments for an Argumentative Theory", in *Behavioral and Brain Sciences*, 34, pp. 57-74.

Mercier, H., & Sperber, D. (2011b), "Argumentation: Its Adaptiveness and Efficacy", in *Behavioral and Brain Sciences*, 34, pp. 94-111.

Mercier, H., & Sperber, D. (2017), *The Enigma of Reason*, Cambridge (MA), Harvard University Press.

Moshman, D., & Geil, M. (1998), "Collaborative Reasoning: Evidence for Collective Rationality", in *Thinking & Reasoning*, 4, pp. 231–248.

Nisbett, R. (2003), *The Geography of Thought: Why We Think the Way We Do*, New York, The Free Press.

Nisbett, R.E., & Borgida, E. (1975), "Attribution and the Psychology of Prediction", in *Journal of Personality and Social Psychology*, 32, pp. 932-943.

Nisbett, R.E., Peng, K., Choi, I. & Norenzayan, A. (2001), "Culture and Systems of Thought: Holistic vs. Analytic Cognition", in *Psychological Review*, 108, pp. 291-310.

Nozick, R. (1993), *The Nature of Rationality*, Princeton, Princeton University Press.

Oaksford, M., & Chater, N. (1994), "A Rational Analysis of the Selection Task as Optimal Data Selection", in *Psychological Review*, 101, pp. 608-631.

Rysiew, P. (2008), "Rationality Disputes — Psychology and Epistemology", in *Philosophy Compass*, 3, pp. 1153-1176.

Samuels, R., & Stich, S.P. (2004), "Rationality and Psychology", in *The Oxford Handbook of Rationality*, ed. by A. Mele & P. Rawling, Oxford, Oxford University Press, pp. 279-300.

Samuels, R., Stich, S.P., & Faucher, L. (2004), "Reason and Rationality", in *Handbook of Epistemology*, ed. by I. Niiniluoto, M. Sintonen, & J. Wolenski, Dordrecht, Kluwer, pp. 131-179.

Sbisà, M. (2006), "Two Conceptions of Rationality in Grice's Theory of Implicature", in *Rationality of Belief and Action*, ed. by E. Baccarini & S. Prijić-Samaržija, Rijeka, University of Rijeka, pp. 233-247.

Sbisà, M. (2007), "On Argumentative Rationality", in *Anthropology & Philosophy*, 8, pp. 89-99.

Slovic, P., Fischhoff, B., & Lichtenstein, S. (1976), "Cognitive Processes and Societal Risk Taking", in *Cognition and Social Behavior*, ed. by J. Carroll & J. Payne, Hillsdale (NJ), Erlbaum, pp. 165-184.

Sperber, D. & Mercier, H. (2012), "Reasoning as a Social Competence", in *Collective Wisdom. Principles and Mechanisms*, ed. by H. Landemore & J. Elster, Cambridge (MA), Cambridge University Press, pp. 368-392.

Stanovich, K.E. (1999), *Who is Rational? Studies of Individual Differences in Reasoning*, Mahwah (NJ), Erlbaum.

Stanovich, K.E. (2004), *The Robot's Rebellion: Finding Meaning in the Age of Darwing*, Chicago, The University of Chicago Press.

Stanovich, K.E. (2011), *Rationality and the Reflective Mind*, Oxford, Oxford University Press.

Stanovich, K.E., & West, R.F. (1998), "Individual Differences in Rational Thought", in *Journal of Experimental Psychology: General*, 127, pp. 161-188.

Stanovich, K.E., West, R.F., & Toplak, M.E. (2011), "Individual Differences as Essential Components of Heuristics and Biases Research", in *The Science of Reason: A Festschrift for Jonathan St. B. T. Evans*, ed. by K. Manktelow, D. Over, & S. Elqayam, New York, Psychology Press, pp. 335-396.

Stein, E. (1996), *Without Good Reason: The Rationality Debate in Philosophy and Cognitive Science*, Oxford, Clarendon Press.

Thompson, V.A., Evans, J.St.B.T., Handley, S.J. (2005), "Persuading and Dissuading by Conditional Argument", in *Journal of Memory and Language*, 53, pp. 238–257.

Tooby, J., & Cosmides, L. (1990), "On the Universality of Human Nature and the Uniqueness of the Individual: The Role of Genetics and Adaptation", in *Journal of Personality*, 58, pp. 17-67.

Tooby, J., & Cosmides, L. (1992), "The Psychological Foundations of Culture", in *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*, ed. by J.H. Barkow, L. Cosmides, & J. Tooby, New York, Oxford University Press, pp. 19-136.

Wason, P.C. (1968), "Reasoning about a Rule", in *The Quarterly Journal of Experimental Psychology*, 20, pp. 273-281.

Wason, P.C. (1983), "Realism and Rationality in the Selection Task", in *Thinking and Reasoning*, ed. by J.St.B.T. Evans, London, Routledge & Kegan Paul, pp. 45-75.

Wason, P.C., & Evans, J.St.B.T. (1975), "Dual Processes in Reasoning?", in *Cognition*, 3, pp. 141-154.

Reason and Building Art: Mies van der Rohe's Lesson

Giorgio Grimaldi
University of Urbino Carlo Bo
giorgio.grimaldi@yahoo.it

1. The "battle" *for* "the new need" as a "battle" *for* "reason"

In one of the first notes Ludwig Mies van der Rohe jotted down in his notebook between 1927 and 1928 we find these words he thought of to base a lesson on: "The battle against the new need not necessarily arise out of an attachment to the old, and we believe freely that at the bottom, it is a battle against reason"[1]. In order to understand this sentence correctly, we must first of all clarify what Mies means by "the new need", then we proceed to correlate it with "reason", a correlation put forward by Mies. What can the relation between "against the new need" and "against reason" be?

It is immediately obvious that "the new need" is a request for "reason". But why is this so? "Not necessarily" because it is "new": novelty is not enough, is not sufficient to reach the rank of "reason". The correlation between "the new need" and "reason" is tighter, or rather more profound.

Let us understand exactly what Mies means by "the new need" by reading some passages from a primary text of his from 1923, which is a sort of manifesto whose guidelines Mies will never take distance from throughout his evolutionary process. The passages are contained in *Office Building*, an article in which he affirms: "Building art is the spatially apprehended will of the epoch. Alive. Changing. New. Not the yesterday, not the tomorrow, only the today is formable. […] Create form out of the nature of the task with the

[1] Mies van der Rohe (1986: 269). For Mies' works I refer to Neumeyer 1986, that is a volume containing almost all of Mies' writings in the appendix *Manifestos, Texts, and Lectures*. To avoid confusion in the footnotes that refer to cited texts, I use "Mies van der Rohe", indicating in the footnote and bibliography the source of the cited text. For a complete biography of Mies van der Rohe, see Schulze & Windhorst (2012).

Giorgio Grimaldi, "Reason and Building Art: Mies van der Rohe's Lesson", in Giorgio Grimaldi and Marialuisa Parise, Rational *Animals. Seven essays on Reason's fields of application*, pp. 105-119

means of our time. *That is our work*"[2]. This sequence is very dense. It expresses a polarity that is aptly summarised by Fritz Neumeyer with the question: "Hegel or Nietzsche?"[3] According to us, it is a polarity that can be reduced to a single pole – Hegel. Although Mies' language could even have been influenced by Nietzschean culture, its conceptual contents are extraneous to Nietzsche. Rather it can be traced back to a structure that reveals a Hegelian heritage, even though neither Hegel nor Nietzsche were ever cited by Mies in his writings.

So, we should consider "alive" to be different from vitalism and in fact connected to "changing", an historical "changing" in a very precise sense of the word, given Mies' interest in history and historical changes. "Alive" is "the spirit" that becomes concrete historical praxis as architecture explicitly does:

> Architecture is the real battleground of the spirit. Architecture wrote the history of the epochs and gave them their names. Architecture depends on its time. It is the crystallization of its inner structure, the slow unfolding of its form. That is the reason why technology and architecture are so closely related. Our real hope is that they grow together, that someday the one be the expression of the other. Only then will we have an architecture worthy of its name: Architecture as a true symbol of our time.[4]

In his 1950 text relevantly entitled *Architecture and Technology*, Mies sets a structural nexus between "architecture" and "spirit" (that is, it is not something absolutely abstract, rather it becomes concrete –it becomes *true*– as praxis, activity and transformation of the present) and another structural nexus between "our time" and "technology": "our time" is the time of "technology". Therefore, architecture, being an essential expression of its own time, must become an "expression" of "technology". "Technology", in turn, in this way takes on an architectural form and becomes an "expression" of it.

Ten years later (*Where Do We Go from Here?*) Mies adds:

> We are not at the end but at the beginning of a new epoch. This epoch is being determined by a new spirit and driven by new technological, sociological and economic forces, and they will come up with tools and new materials. For this

[2] Mies van der Rohe (1986: 241).
[3] Neumeyer (1986: 63).
[4] Mies van der Rohe (1986: 324).

> reason we will also have a new architecture. But the future does not come about by itself.[5]

In order to understand the specific sense in which Mies conceives the various elements of this passage – in particular the final sentence – we must refer back to a manuscript written in 1926 for a conference. Mies writes: "Each culture arises out of the landscape and its economic givens. Only in this connection can one even understand the term culture"[6]. He continues: "Transformations within cultures, too, depend on transformations of their economic structures. These can succumb to political influences, much as they in turn can influence them"[7]. Here, Mies pushes the dialectic of structure/superstructure beyond its usual form[8]; it is in line with the potential and the power of dialectics. He sets in motion a fundamental element of historical materialism, dealing with a question previously taken on by Engels[9]. Mies sets up a relationship between structure/superstructure, both without an absolute *primum*; he recognises a *primum* (that is, structure) *from* which and – dialectically – *on* which superstructure moves and acts, without automatism:

> It is, of course, wrong to assume that an economic change in a society is automatically followed by a change in ideology. The change of the ideological superstructure takes place often very much later and much slower than the changes in the societal ground. The exterior shell of things, the crystallization of life processes, remains standing even then and exerts its influence long after its kernel has been hollowed out. Even when one no longer comprehends their meaning, one still adheres to the forms; one continues to devalue them, one alienates them even further from their original meaning. What once was highest expression of vital forces gradually deteriorates into a senseless banality unless […] power and grandeur can be summoned to impart new expression to the changed life content.[10]

---

[5] Mies van der Rohe (1986: 332).
[6] Mies van der Rohe (1986: 252).
[7] Mies van der Rohe (1986: 252).
[8] Discussion of the potential of dialectic and the limitations of historical materialism is not relevant to this paper.
[9] Engels defended historical materialism from its mechanistic and deterministic reduction. A letter dated September 21th 1890 that Engels wrote to J. Bloch reads: "According to the materialistic conception of history, the factor that *ultimately* is determining in history is the production and reproduction of real life. Nor Marx, nor myself have not claimed anything else. However, if anyone misinterprets things, claiming that the economic factor is the *only* determining factor, they transform that claim into an empty, abstract, absurd sentence" – in Marx & Engels (1959: 43, footnote 1).
[10] Mies van der Rohe (1986: 252).

Therefore, it could be a *persistence* of elements that have lost their historical function and centrality; nevertheless their sunset delays; hence, we begin to perceive why Mies does not have "the old" as his polemical objective. The problem is not so much "an attachment to the old", rather it is hostility towards "reason". The opposition is not "the new need"/"the old", but – the opposition conceived "at the bottom" – "the new need"/"against reason".

We must proceed to a better in-depth analysis of the terms of this opposition in order to capture their implied configuration. We have made progress by recognising that "the old" – in itself, that is "old" because it does not correspond with current times – sets. It could be a long-lasting sunset, even insidious in its declining permanence[11]. But every colour will become diluted in the night, in its own night that will consign it to oblivion, or already keeps it in remembrance.

2. Which epoch?

We have come to learn that "the new need" is not hindered by "the old" but rather by another element, that we deduce as hostility towards "reason". We might therefore think of the irrational and we might probably be on the right track. This in fact is the point that must be explained, that is, what do we precisely mean by "irrational"? Above all, we must take into consideration *logos*.

We must analyse the texts that have been aforementioned to unravel the issue, whose entirety of its central contents we are far from having comprehended.

According to Mies van der Rohe, "the old", which is such because it is no longer capable of meeting the "economic structures" of the present, could well be "what once was highest expression of vital forces", and that now instead slowly sets and decays until "power and grandeur can be summoned to impart new expression" to "the new need", "to the changed life content".

How can this "new need" be born? From which soil does it rise? It is the economic structure, the organisation of production from time past to the current time; it changes configuration in accordance with the development of productive forces. This is the background – at an implicit level – of Mies' argument. This development does not automatically correspond with a

[11] Ernest Labrousse observed: "A dominant mentality that resists is one of the major restraining factors of History", in Bouloiseau (1972).

change in "ideological superstructure"; in fact, a mere subsisting of it, that no longer has full historical meaning.

Hence, we can reformulate the question "what does Mies means with 'the new need'" in the light of what we have plumbed up until now. "The new need" rises from the epoch of the industrialisation of production, of advanced capitalism, of the standardisation of the production processes, and of the progressive automatisation of the production processes. In other words, from the technological-productive system, instituted and implemented by capitalism. At present, there is no "escape route" from this system, no alternative way to develop the productive forces. The countries of the former Soviet Union well know it, Lenin had understood it[12], and China knows it well[13]. "The new need" is borne from the historical constellation of this technological-productive system, instituted by capitalism but not necessarily and inexorably bound to it (maybe not even truly corresponding to it, as Marx prefigured)[14].

Mies van der Rohe's strength lies in taking action on the present, placing attention on the present, and focussing exclusively on it: "only the today is formable". Beyond utopia, the present has to be realised in all its possibility that carries out its content full of progress. Beyond any waiting for a future that fulfils the promises of the present, it is necessary to act in the current times. However, they are not current times as ends in themselves, as they consume themselves in the immediate. "The future does not come about by itself"; in view of the future, we must act *in* the present.

"The new need" is made possible and is generated by the present, that is, the epoch of technology[15]: so, "the battle against the new need" is "the battle" against the epoch of technology, and at the same time "against reason". However, what is the nexus between the epoch of technology and "reason"? Is it correct to act obliging this nexus?

We have further unravelled the opposition "the new need"/"against reason" but the sense of reason that has not yet been clarified impedes us from fathoming the correspondence "the new need"/epoch of technology/"reason". It could be said that the epoch of technology is the epoch of unfolded reason? However, this view of correspondence could be taken as simplistic, superficial, triumphalist, and at the same time consolatory.

---

[12] Cfr. Lenin (2017).

[13] Refer to Grimaldi (2017) for the technical-industrial apparatus that is necessary for the development of productive forces.

[14] Cfr. Marx (1939: 710-711) and Marx (1867: 583-584).

[15] Refer to Grimaldi (2015: 102-104) for this decisive, however not exhaustive, formula for determining our epoch.

To access a conceptually profound meaning of reason, we must take a step back and return to what seems to oppose it: those passed *forms* ("the old") that linger in their long-standing, dangerous sunset.

3. On historical movement

Following Labrousse[16], why should and could the sunset of passed forms be dangerous? Why should their persistence be insidious? What [are] they "restrain[ing]"? Returning to Mies, how could it be that "what once was highest expression of vital forces gradually deteriorates into a senseless banality"?

Mies' attitude to thought – and our discussion that derives from it – becomes clear if we examine Hegel, who is never mentioned by Mies, but whose philosophical approach is implicitly present.

In § 3 of *Elements of the Philosophy of Right* we encounter a decisive conceptual "distinction"[17] that helps us to understand and interpret Hegelian thought and disclose the deepest meaning of Mies' analysis. That is, the "distinction" – that is not well received nor well-known to critics of Hegel that identify him as a philosopher of historic justificationism – between a "historical explanation and justification [*die geschichtliche Erklärung und Rechtfertigung*]"[18] and a "justification which is *valid in and for itself* [an und für sich gültigen *Rechtfertigung*]"[19]. The distinction follows:

> This distinction, which is very important and should be firmly borne in mind, is at the same time a very obvious one; a determination of right may be shown to be entirely *grounded in* and *consistent with* the prevailing *circumstances* and *existing* legal institutions, yet it may be contrary to right […] and irrational in and for itself [*an und für sich unrechtlich und unvernünftig sein*], like numerous determinations of Roman civil law […] which followed quite consistently from such institutions as Roman paternal authority and Roman matrimony. But even if the determinations of right are rightful and rational, it is one thing to demonstrate that this is so – and this cannot truly be done except by means of the concept [*durch den Begriff wahrhaftig geschehen kann*] – and another to depict their historical emergence and the circumstances, eventualities, needs and incidents which led to their introduction.[20]

[16] See footnote 11.
[17] Hegel (1821b: § 3).
[18] Hegel (1821b: § 3) = Hegel (1821a: § 3).
[19] Hegel (1821b: § 3) = Hegel (1821a: § 3).
[20] Hegel (1821b: § 3) = Hegel (1821a: § 3).

That something is justifiably tied to a given historical content does not at all or automatically mean that it is «rightful and rational». To be considered so, historical existence and not even prestige or success are sufficient[21]; instead, it is necessary to examine the judgement of concept. This judgement, rather than ratifying the complete rationalisation of the existing, queries it and evidences its caducity. In other words, it evidences its sunset lights:

> If it can be shown that the origin of an institution was entirely expedient and necessary under the specific circumstances of the time, the requirements of the historical viewpoint are fulfilled. But if this is supposed to amount to a general justification of the thing itself [*eine allgemeine Rechtfertigung der Sache selbst*], the result is precisely the opposite; for since the original circumstances are no longer present, the institution has thereby lost its meaning and its right [*ihren Sinn und ihr Recht*].[22]

After reading the above cited texts by Hegel, what Mies means when he affirms that what is "deteriorat[ing]" was "what once was highest expression of vital forces" becomes more comprehensible. After Hegel's *Elements* and before Mies' 1926 manuscript there are the Hegelian left and the historical materialism; however, even in the contrast, in the overturning, the conceptual lineage of Marx from Hegel is clear, as it is clear from the text we have just examined.

The materialistic approach that develops from the Hegelian thought structure allows Mies, in his 1926 manuscript, to proceed to a reconstruction of fundamental historical phases in the history of architecture and of the techniques employed in those phases from a historical point of view; more precisely from a historic-economic point of view. After having aligned (as Weber did) "Protestantism" and "capitalism"[23], Mies writes in a series of concise sequences that:

> The mechanization of the country progressed further. […] International traffic and international trade determine life from now on. Metropolises of immense proportions develop. […] Gigantic industrial complexes arise, yes, entire indus-

[21] This aspect is particularly relevant because it annuls the foundations of a banal criticism, which occurs more often than one may imagine. This criticism against Hegelian philosophy accuses it of proceeding to a celebration of the rational supremacy of any element that gains success and manages to impose itself on reality. Instead, the "justification which is *valid in and for itself*" shows that the rational/actual relation, as we will see, is a correspondence that is not in the form of a superficial immediacy with merely historical data.

[22] Hegel (1821b: § 3) = Hegel (1821a: § 3).

[23] Mies van der Rohe (1986: 254).

> trial cities. […] The machine has long since become master of production. […] Now there is only economy. It dominates everything, politics and life.[24]

This text is an analysis, and also at the same time an appraisal of the situation in Germany, Europe and the West up until the end of World War I ("Now there is only economy. It dominates everything, politics and life"), but as time will tell, its application goes well beyond its own time and geopolitical context.

What is Mies' answer to the domination of "politics" and "life"[25] by the economy? It is clearly articulated in a conference he held in Berlin in 1928, which recalls his 1926 manuscript in which the reconstruction of historical periods emphasises consciousness over the economy. It is an answer that is not prone to adapting to that dominion, but rather it invites to utilise power in praxis:

> the building art is always bound to its own time […] Our time […] It has been given to us as a task that we have to master. […] We have to become master of the unleashed forces and build them into a new order, an order that permits free play for the unfolding of life. […] No matter how immense knowledge may be, how huge the economic apparatus, how powerful technology, all that is only the raw material compared to life.[26]

History, therefore, is reconstructed by Mies following rational logic: it is not a series of events that cannot be depicted in a coherent whole thing. This does not mean (following Hegel, or rather in line with Hegel) that every element of any historic phase coincides with rationality in and for itself. The "task" that Mies intends to "master" is that of ordering "its own time", to which "the building art is always bound to"[27]. What is the criteria for ordering the reality of its time? To locate it we need to consult the notes that are guiding our reflection, that is, that "battle" for "the new need", for "reason" and for technology.

A triple correspondence that is fully confirmed in Mies' texts, as we read them juxtaposing one against the other. In his Berlin conference in

---

[24] Mies van der Rohe (1986: 254-255).

[25] Mies' framework can be variously developed following a biopolitical analysis. For reasons already referred to, this approach does not allow us to locate Mies in Nietzsche, for the very sense that "life" here has, in the *form of Geist*; which does not exclude the possibility of undertaking a non-Nietzschean biopolitical analytical line.

[26] Mies van der Rohe (1986: 301).

[27] It is difficult not to think – given the importance of "its own time" in Mies – of the famous Hegelian view according to which "each individual is […] a *child of his time* [*ein* Sohn seiner Zeit]" (Hegel 1821b: 21 = Hegel 1821a: 26).

1928 he aligned himself with the "group" that does not intend to "hold […] in low respect" the "achievements" of those who are bound to the past, that is, to "the old order". This "old order" is fading and has lost "vitality" because it does not know how to respond/correspond to "the new need" imposed by the new historical phase, by "technology" and by the domination of "economy". The "group" to which Mies affirms he belongs does not criticise the "old" because it has passed and been consolidated, but because it is unable to respond to the times, to be a "*child of his time*". According to the "group, the old habitual order has lost its meaning and vitality. One cannot suffer it any more. This group affirms the new world and fights with its means"[28], which is the underlying notion of the battle to which the 1927/1928 notes refer, contemporary to period in which Mies writes his Berlin conference text. If we put the two texts alongside one another, we see a correspondence between the "battle" for "the new need", for "reason", and the "affirm[ation]" of "the new world" and the "fight[ing] with its means". Returning to the conference text, we meet the "battle" for technology and its own epoch: "We do not need less but more technology"[29]. "The new need", "reason" and "technology" are therefore, in Mies' texts, set in a correspondence, in one constellation. "Technology" is the driving force for emancipation on a large scale, on mass, and which is accosted by the powers of "science"[30] and "economy":

> We see in technology the possibility of freeing ourselves, the opportunity to help the masses. We do not need less science, but a science that is more spiritual; not less, but a more mature economic energy. […] It must be possible to solve the task of controlling nature and yet create simultaneously a new freedom. It must be possible to see elitism disappear and admit the fact of the masses, the fact that each of the many has a right to life and goods. Mass must not be a cliché for us. It must come to articulation from within itself, for only that way can the forces residing in it be made serviceable for all.[31]

In the aforementioned citations we have come across the notion of "acceptance" of one's own current epoch: "affirm[s] the new world", "admit the fact of the masses". In other words, modernity has been accepted[32]. This acceptance is the fruit of the crux that makes sense of the *new need-reason-*

[28] Mies van der Rohe (1986: 301).
[29] Mies van der Rohe (1986: 301).
[30] Mies van der Rohe (1986: 301).
[31] Mies van der Rohe (1986: 301).
[32] Which is what Heidegger does not do (concerning technology, Heidegger, 1957a and Heidegger, 1994; concerning science, Heidegger, 1957b and Heidegger, 1950). At times Mies van der Rohe takes on a truly anti-Heidegger viewpoints.

technology sequence that we must reorder as *reason-technology-new need,* where no term is fixed, but rather exists in a dialectic relation.

Nevertheless, we must go one step further, the step that finally sheds light on the relationship between reason and reality and that allows such brightening that consequentially extend "the new need" to its generation in the epoch of technology. Therefore, again, to understand Mies' conceptual framework we must return to Hegel.

4. Architecture as historical-concrete form of *logos*

In § 6 of the *Encyclopedia of the Philosophical Sciences* Hegel returns to the meaning of his well-known identification of the rational and the actual that is written in the *Preface* of his *Elements of the Philosophy of Right*, according to which "What is rational is actual; and what is actual is rational"[33]. Over time, this guiding formula has raised quite a few accusations of historical "justificationism" that, by starting with a less hasty reading of the § 3 that we well know, could have been avoided. In the *Encyclopedia*, Hegel returns to the guiding formula explaining its sense, which is particularly useful, even for a non-hostile reader.

We read, in the guiding formula that we meet in the *Elements*, that "in ordinary life, we may accidentally call every idea, error, evil, and the like, *actual*, as well as every concrete existence […], crippled and transitory though it may be"[34]. Therefore, it is specified that not everything that happens as a phenomenon has the *status* of actual. Contesting the untiring supporters of ought – that is those who "sever[ance]" the "actuality from the idea"[35], and that in the name of the latter take distance from the former, and so critique the actual because, according to themselves, it is unable to realise ideas in their noblest and highest form – Hegel renders even more clearly the notion of that identification:

> When its [of the understanding] "ought" is directed against trivial, superficial and transitory objects, arrangements, situations, and so forth (that is to say, what may perhaps be of relative importance [*größe relative Wichtigkeit*] to certain circles for a period of time), then this understanding may indeed be right to find many things that are not in accord with universal and correct standards [*allge-*

[33] Hegel (1821b: 20).
[34] Hegel (1830b: § 6).
[35] Hegel (1830b: § 6).

> *meinen richtigen Bestimmungen*]. Who would not have enough good sense to see much around him that is indeed not as it should be [*wie es sein soll*]?[36]

It is highly singular that an author who recognises the possibility of a "*größe* […] *Wichtigkeit*" to given historical elements, and at the same time attributes to it "relative importance", is still today accused of – in certain circles – justifying and celebrating any historical element that has gained "success" by becoming dominant in a particular historical point in time. Yet, by not acknowledging superficial readings of some interpretations, we can affirm by paraphrasing Mies that *the battle against the Hegelian guiding formula does not have as its critical object the negativity of actual, but rather the critique of rational.* This critique is totally absent in Mies, who instead considers reason as a guide in the art of building, and in building according to the requirements of the epoch. Finally, Hegel, who seeks to identify the actual, not in any occurrence but rather in the presence of a rational substratum (that is the structure of the world, of the historical world), does not invite to acritical celebration of "*his time"* because it is an expression of the rational. Nor does he dismiss it because it has become deformed in "many things", rather he invites locating the rational substratum, that is, the core and the foundation in what is current every single time. Hegel's indication allows a critical but active approach towards "*his time"*, and therefore a full acceptance of modernity[37] (that we see as a central element in Mies' understanding of the "task" the architect takes on) that is indeed dissimilar to a full justification of its every aspect. An acceptance that makes it possible to assume the task – in the name of a rational that moves history as a power subtended to it – of the realisation of the concrete potential of modernity.

Hence, we are now in the position to read the 1927/1928 notes at a different level than before, unravelling its profound nodes and reorganising the complexity of its meaning starting from the correlation between "the new need" and "reason". Let us read the note again: "The battle against the new need not necessarily arise out of an attachment to the old, and we believe freely that at the bottom, it is a battle against reason". Now we understand why Mies' polemic objective is not "the old" as "old", but instead "the old" that no longer conforms to process of reason. Which reason? And why is it an expression of «the new need» whose way is paved and implemented by technology?

[36] Hegel (1830b: § 6) = Hegel (1830a: § 6).
[37] Concerning Hegel and modernity, see Losurdo 1992.

Reason and the new need arise in various phases, realising phases in history as history of reason, as reason gradually becomes true through new forms. Going "against the new need" and "today", in the name of a "yesterday" that no longer exists, is merely an "attachment to the old", or to a "tomorrow" that is yet to come (which is quite similar, if taken solely, to that "ought" that was much criticised by Hegel, because in the name of something that was absolutely *distant*, it denied the – even critical – understanding of the present). It means going "against reason", ultimately against the structure that forms, holds and moves the concrete, actual, historical movement, that is not chaos made of fragmentary elements in an unceasing and contingent configuration, but an immanence – as a *continuum* – that continuously transcends itself according to the way of *logos*. This is why "the new need" and "reason" correlate, and the fundamental opposition is not "the new need"/ "the old" but rather "the new need"/ "against reason". "The old", lingering its sunset, is dangerous not because it is "old" and has reached its sunset but because restraining the new, that is, a further phase of history of reason – "old" is what does not hold the stroke of the "justification which is *valid in and for itself"*, the stroke of *Aufhebung*, that takes away and keeps according to *logos*.

"The new need" is an expression of the new phase – the phase of the epoch of technology – of the history of reason, of history as a field of realisation of *logos*, that works continuously on itself, in itself.

Hence, it is not the epoch of technology, as a technological one, that is the epoch of reason unfolded; on the contrary, it is the reason, as reason, that unfolds itself in the epoch of technology. In other words, it is not technology that unfolds itself *via logos*, but on the contrary, it is *logos* that unfolds itself *via* technology. If we study Mies' notes, we can reach these conclusions, that are beyond the very text and a development of the concepts that act in it.

But there is a last, further step: history as history of *logos* is at the same time history of freedom as fulfilment of its universalistic movement. Once again, we refer to Hegel; in *Lessons on the Philosophy of History* he illustrates this movement: "Orient knew and only knows that *One* is free [*daß* Einer *frei ist*], the Greek and Roman world knew that *Some* are free [*daß* Einige *frei seien*], the Germanic world knows that *All* are free [*daß* Alle *frei sind*]"[38]. Universalism as key-content of the historical movement of *logos* (that has a universalistic character) is present in Mies' connection between "technology" and "masses" that we saw in his Berlin conference text. The

[38] Hegel (1837: 134).

objective is "controlling" immediate nature (that, in itself, threatens existence through constant danger) and the "creat[ion]" of "a new freedom". The two moments are "simultaneous[ly]". Freedom is the overcoming of immediate nature (this is another Hegelian theme[39]); building art is a decisive element for this overcoming, as an overcoming of the immediate natural environment.

Architecture, as a historical-concrete form of *logos*, is a manifestation/ expression of its figures. The movement of architecture leads to Mies' "*Less is more*"[40] as work toward the essential of form through linearity. This is a moment of the work of *logos* on itself; in art it fulfils concept.

References

Bouloiseau, M. (1972), *La République jacobine. 10 août 1792 – 9 thermidor an II* (*La Francia rivoluzionaria. La Repubblica giacobina 1792-1794*, Italian translation by S. Brilli Cattarini, C. Patanè, Roma-Bari, Laterza, 1987).

Grimaldi, G. (2015), *Oltre le tempeste d'acciaio. Tecnica e modernità in Heidegger, Jünger, Schmitt*, Roma, Carocci.

Grimaldi, G. (2017), "Industrializzazione e progresso: la lezione della Rivoluzione d'ottobre", in "Materialismo storico. Rivista di filosofia, storia e scienze umane", "Rivoluzioni e restaurazioni, guerre e grandi crisi storiche: cento anni dall'Ottobre russo. Parte prima", edited by S.G. Azzarà, 2017/2, 108-122.

[39] Cfr. Hegel (1821a: § 187).

[40] This is a sentence that Mies attributes to Peter Behrens (cfr. Mies van der Rohe, 2010: 220), but that has become famous as Mies' own.

Heidegger, M. (1950), *Die Zeit des Weltbildes* (*L'epoca dell'immagine del mondo*, in *Holzwege. Sentieri erranti nella selva*, Italian translation by V. Cicero, Milano, Bompiani, 2014, 175-265).

Heidegger, M. (1957a), *Die Frage nach der Technik* (*La questione della tecnica*, in *Saggi e discorsi*, Italian translation by G. Vattimo, Milano, Mursia, 1976, 5-27).

Heidegger, M. (1957b), *Wissenschaft und Besinnung* (*Scienza e meditazione*, in *Saggi e discorsi*, Italian translation by G. Vattimo, Milano, Mursia, 1976, 28-44).

Heidegger, M. (1994), *Einblick in das was ist*, in *Gesamtausgabe. III. Abteilung: Unveröffentlichte Abhandlungen. Band 79. Bremer und Freiburger Vorträge*, Frankfurt am Main, Vittorio Klostermann, 2005, 3-77.

Hegel, G.W.F. (1821a), *Grundlinien der Philosophie des Rechts. Naturrecht und Staatwissenschaft im Grundrisse*, Frankfurt am Main, Suhrkamp, 1986.

Hegel, G.W.F. (1821b), *Grundlinien der Philosophie des Rechts. Naturrecht und Staatwissenschaft im Grundrisse* (*Elements of the Philosophy of Right Or Natural Law and Political Science in Outline*, English translation by H. B. Nisbet, Cambridge, Cambridge University Press, 1991).

Hegel, G.W.F. (1830a), *Enzyklopädie der philosophischen Wissenschaften im Grundrisse* (*Enciclopedia delle science filosofiche in compendio*, parallel text, Italian translation by V. Cicero, Milano, Bompiani, 2000).

Hegel, G.W.F. (1830b), *Enzyklopädie der philosophischen Wissenschaften im Grundrisse* (*Encyclopedia of the Philosophical Sciences in Basic Outline. Part I: Science of Logic*, English translation by K. Brinkmann, D.O. Dahlstrom, Cambridge, Cambridge University Press, 2010).

Hegel, G.W.F. (1837), *Vorlesungen über die Philosophie der Geschichte*, Frankfurt am Main, Suhrkamp, 1986.

Lenin (2017), *Economia della rivoluzione*, Milano, il Saggiatore, edited by V. Giacché.

Losurdo, D. (1992), *Hegel e la libertà dei moderni*, 2 vols., Napoli, La scuola di Pitagora, 2011.

Marx, K. (1867), *Das Kapital. Kritik der politischen Oekonomie. Erstes Buch* (*Il capitale. Libro primo*, edited by A. Macchioro, B. Maffi, Torino, UTET, 2009).

Marx, K. (1939), *Grundrisse der Kritik der politischen Oekonomie. 1857-1858* (*Lineamenti fondamentali di critica dell'economia politica («Grundrisse»)*, 2 vols., edited by G. Backhaus, Torino, Einaudi, 1976).

Marx, K., & Engels, F. (1959), *La concezione materialistica della storia*, Roma, Editori Riuniti, 1981. It is an excerpt – edited by F. Codino – from *The German Ideology*.

Mies van der Rohe, L. (1986), *Manifestos, Texts, and Lectures*, in Neumeyer, F., *Mies van der Rohe. Das kunstlose Wort. Gedanken zur Baukunst* (*The Artless Word. Mies van der Rohe on the Building Art*, English translation by M. Jarzombek, Cambridge – Massachusetts, The MIT Press, 1991, 240-339).

Mies van der Rohe, L. (2010), *Ausgewählte Schriften* (*Gli scritti e le parole*, Italian translation by V. Pizzigoni, Torino, Einaudi, 2010).

Neumeyer, F. (1986), F., *Mies van der Rohe. Das kunstlose Wort. Gedanken zur Baukunst* (*The Artless Word. Mies van der Rohe on the Building Art*, English translation by M. Jarzombek, Cambridge – Massachusetts, The MIT Press, 1991).

Schulze, F., & Windhorst, E. (2012), *Mies van der Rohe*, Chicago and London, The University of Chicago Press, 2014.

Can We Rationally Talk about Pictorial Action Possibilities?

Gabriele Ferretti
University of Basel, University of Florence
gabriele.ferretti88@gmail.com

Silvano Zipoli Caiani
University of Florence
silvano.zipolicaiani@unifi.it

1. Introduction[1]

Depicted objects do not occupy a tridimensional portion of the external space. Indeed, the pictorial space does not overlap with all the three dimensions of the space around us. To this extent, depicted objects cannot afford action possibilities: we cannot act upon them. We can try, but all we can do is to hit the picture's surface. And, in the case the depicted object is presented on a monitor, we risk clashing against it.

For this reason, it seems rational to believe that depicted objects cannot even recall any action possibility. Since you cannot grasp a depicted candy, it seems natural to argue that you will not represent the possibility to grasp it. According to a classical view, indeed, representing an action possibility depends on the fact that you can really act on the object you represent the action possibility as having[2]. Hence, an agent cannot intend to act on targets that he/she believes are depicted objects, without being guilty of irrationality.

[1] This project was funded by 'Fondazione Franco e Marilisa Caligara per l'alta formazione interdisciplinare'. We would also like to thank Marialuisa Parise, Giorgio Grimaldi and Pierluigi Graziani.
[2] Matthen (2005), Nanay (2011), Nanay (2013), Jacob & Jeannerod (2003).

Gabriele Ferretti and Silvano Zipoli Caiani, "Can We Rationally Talk about Pictorial Action Possibilities?", in Giorgio Grimaldi and Marialuisa Parise, Rational *Animals. Seven essays on Reason's fields of application*, pp. 121-148

Against this natural and common idea, in this paper we defend the controversial idea that well-known depicted objects can, indeed, recall action possibilities without that the agent is liable of irrationality.

Here's the plan for the paper. We first distinguish between two different classes of pictures: *normal* pictures and *trompe l'oeil* pictures. With this distinction at hand, we address two different theses about pictorial perception of action possibilities (§2). Second, we describe the *dispositional view of action possibilities* (§3). Then, we report empirical evidence from visual neuroscience showing that perceiving depicted objects elicits the same cortical visuomotor areas of the brain that are functionally involved in the ascription of action possibilities in the case of real objects (§4). Thus, we discuss the implication of this evidence, by suggesting that it allows to defend the philosophical claim that we can ascribe action possibilities to depicted objects (§5). At this point, we suggest a possible way to renew our classical conception of the perception of action possibilities by extending the *requirement*s at the basis of this conception (§6): the *identification requirement*, the *specification requirement* and the *actualization requirement*. We conclude by sketching an analysis of the relations between the ascription of presence and the ascription of action possibilities in the case of pictorial objects (§7).

2. Pictorial Action Possibilities

Before proposing our claim, we need to draw a distinction between two different classes of pictures.

On the one hand, we are very familiar with *normal* pictures like paintings, photographs, etc. With *normal* pictures, the experience we undergo is pictorial. The depicted object looks as such. We can perceive the presence of a surface, which is the vehicle of the pictorial content. We can perceive the pictorial content itself. And these two components of the picture are simultaneously represented, consciously or unconsciously[3].

On the other hand, there is a kind of pictures, *trompe l'oeil* pictures, which are able to fool our visual system that we are dealing with a real object. The pictorial content looks like a real object and we cannot represent the presence of a surface[4].

[3] Lopes (2005), Hopkins (2012), Nanay (2011), Nanay (2017).

[4] Lopes (2005), Hopkins (2012), Nanay (2015), Nanay (2017), Ferretti (2016d), Ferretti (forthcoming a), Ferretti (forthcoming b), Ferretti (2018a), Ferretti (2018b), Ferretti (2019a), Ferretti (2017c), Zipoli Caiani (2016).

We saw that our main claim wants to establish that we can ascribe action possibilities to depicted objects. But since we have, at least, two different kinds of pictures, whose pictorial experience is very different, we should sharpen our claim with respect to this difference. So, we can distinguish two theses: *the weak thesis of pictorial action possibilities* (henceforth: WPA) and *the strong thesis of pictorial action possibilities* (henceforth: SPA).

WPA: *Trompe l'oeil* pictures foster in the subject the illusion, at the *personal* level, of affording suitable possibilities of action (and elicit the same subpersonal resources involved in representing action possibilities pertaining to real objects). That is, the subject perceives an action possibility both consciously and unconsciously.

SPA: *Normal* depicted objects do not foster in the subject the illusion, at the *personal* level, of affording suitable possibilities of action, but they can nevertheless elicit the same subpersonal responses involved in representing action possibilities pertaining to real objects. That is, though the subject does not perceive the action possibility consciously, she/he can ascribe action possibilities to them unconsciously.

It is important to note that, differently from *trompe l'oeil pictures*, normal pictures cannot foster in the subject the illusion, at the *personal* level of visual processing of affording suitable possibilities of action - and, as for *trompe l'oeil*, in some cases, of recalling intentional action.

On the one hand, the WPA is not very controversial: *trompe l'oeil* pictures are expressly designed in order to fool the subject, who consciously perceives the presence of action possibilities related to the object visually encoded in the perceived surface, which does not look as a depiction anymore. On the other hand, SPA seems to be more interesting: *normal* pictures are not designed to foster such an illusion. Indeed, rationally speaking, *normal* pictures should not recall any form of action to the subject, neither at the *personal* level of visual processing, nor at the *subpersonal* level of visual processing.

Thus, when dealing with *normal* depicted objects an agent cannot intend to act on targets that he/she believes are depicted objects, without being guilty of irrationality. Differently, since *trompe l'oeil* pictures trigger a visual illusion at the personal level, one can intend to act on them without being accused of irrationality although, in quality of pictures, they do not really offer real possibilities of action: they only fool our visual system that they do.

Against this intuition, we will defend SPA on both philosophical and empirical grounds. We'll suggest that we can ascribe action possibilities to *normal* depicted objects at the subpersonal level of visual processing. This happens without that the agent visually perceives this possibility at the conscious level, and that it visually represents, consciously or unconsciously, the object as present. Nor she needs to believe that the object is present - as in the case of *trompe l'oeil* pictures.

3. The Dispositional View of Action Possibilities

Visual perception of the environment quickly and automatically evokes possibilities of action. This processing is the natural mental antecedent of the planning and execution of motor actions. If this were not so, and visual perception were silent about how to interact with the environment, we would have serious difficulties in deciding how to plan and execute our actions. It is a commonly held conception that action possibilities are properties of the environment showing a specific salience for those agents whose body and motor skills make them able to interact with them. Thus, action possibilities are properties emerging from the encounter of a specific object and a specific individual. In other words, according to this view, some properties of the environment are possibilities of action if there are agents that can use it for executing goal-directed motor behaviors.

To this extent, over the last few decades it has become popular to conceive possibilities of actions as *dispositional properties* pertaining to objects, which can be complemented by properties of the agent's body. In this vein, possibilities of actions are relational properties of things that, given suitable circumstances, actualize[5]. More precisely, a possibility of action is a property pertaining to an object that, given background conditions, allows an agent endowed with suitable bodily and motor properties to execute a specific goal-directed motor behavior.

Following this dispositional approach, Turvey *et al*. (1981) have introduced a definition of action possibility according to which an object X affords the goal directed motor behavior B for an agent Z, if and only if a property of X is dually complemented by a property pertaining to the body of Z. Scarantino (2003) has famously discussed the concept of possibility of action in terms of the probability of an action manifestation, stating that something is a possibility of action for a certain agent if there is a positive probability that such a property can be engaged in an event that qualifies as

[5] Turvey *et al*. (1981).

an agent's action. In other words, according to Scarantino, possibilities of action are environmental properties complemented by properties of an agent's body, whose actualization might give rise to a case of action execution.

To clarify the meaning of a dispositional approach to action possibilities, it can be helpful to rely on a familiar case. Consider, for example, the handle of a mug. According to the dispositional view, the handle can be conceived as the bearer of a property whose actualization is made possible by the occurrence of suitable circumstances. Notably, the handle offers a possibility of action which actualizes just in case there is an agent whose body is construed in such a way that there can be a positive probability that she/he interacts with it giving rise to a goal-directed motor behavior.

However, for an agent to be able to interact with a property of the environment, she/he must be able to perceive it *as* a possibility of action. Accordingly, one may be tempted to assume that visually perceiving action possibilities consists in a specifying relation between the optical stimulus and the actual environmental property. Therefore, what is required to perceive the opportunities for motor interaction offered by the environment is nothing but the instantiation of such a relation[6]. Notably, this view has traditionally assumed that the laws governing the optical stimulation specify a *unique* property in the environment[7], allowing the light arrays to carry information about the presence of action-related dispositional properties[8]. This means that a possibility of action can be univocally related to a pattern of visual stimulation, and that detecting the latter is the same as perceiving the former.

To resume, the dispositional approach to the possibilities of action consists of three main *requirement*s:

(1) *The Actualization Requirement*: since action possibilities are dispositional properties of the environment, there should be circumstances in which they actualize with a positive probability;

(2) *The Identification Requirement*: to actualize, action possibilities necessarily involve the identification of action related properties bearers suitable for action execution;

[6] Gibson (2014).
[7] Gibson (1967), Turvey *et al*. (1981).
[8] Michaels & Carello (1981).

(3) The *Specification Requirement*: action related property bearers can be univocally specified by the structure of the optical stimulation.

It is worth noting that the actualization of an action depends on the presence of a property bearer suitable for action; therefore (1) depends on (2). Indeed, it is impossible to obtain the actualization of a certain action without postulating the existence in the environment of a referent suitable for that action. Moreover, since the execution of an action depends on the agents' ability to perceive an action-related property bearer as a possibility of action,(2) depends on (3).

However, as we shall see, in order for an agent to perceive something as an action-related property, it is not necessary that the bearer of such an action-related property is a real object. In what follows, we report empirical evidence from visual neuroscience showing that perceiving *normal* depicted objects, that is, *non-trompe l'oeil* depicted objects mentioned by SPA, elicits the same cortical visuomotor areas of the brain that are functionally involved in the ascription of action possibilities in the case of normal objects. This will be the empirical basis that allows us to defend the philosophical claim that we can ascribe action possibilities to *normal* depicted objects. But that means that the ascription of action possibilities can occur regardless of the real tridimensional presence of action-related property bearers and positive conditions for actualization.

4.Visual Neuroscience, Pictures and Action

Here we report empirical evidence that is relevant for an argument supporting SPA. The empirical results we address how that the perception of *normal* depicted objects elicits the same cortical areas of the brain that are functionally involved in the ascription of action possibilities in the case of real objects. That means that both normal depicted objects and real objects recall, at the cortical level, the same functional response concerning the visuomotor processing that allows us to perceive action possibilities. This evidence will prepare the empirical ground for the philosophical argument in favor of SPA that we will offer in (§4). The philosophical analysis of such results will allow to suggest that it is reasonable to defend SPA.

Our best model of vision in visual neuroscience, the *Two Visual Systems Model* (TVSM)[9], suggests the presence of an anatomo-functional separation of our visual system into two main streams: a ventral stream for

[9] Milner & Goodale (1995/2006).

visual recognition and a dorsal stream for visually guided action. While ventral processing is conscious, the dorsal one is unconscious. Now, the detection of action possibilities, with the consequent elaboration of the motor acts we can use in order to act upon these possibilities are courtesy of several cortical sub-portions of the dorsal stream. In particular, the transformation of intrinsic object properties into motor acts relies on a well defined cortical network lying in between the parietal cortex and the ventral premotor cortex, which pertains to a sub-portion of the dorsal stream: the ventro-dorsal stream.

The main cortical components for this task are the anterior intraparietal (AIP) and the area F5 in the most rostral part of the ventral premotor cortex. The AIP extracts the geometrical information concerning action possibilities for grasping purposes and relays this information to neurons in F5. At this point, these neurons compute the information received by AIP in order to translate the geometrical attributes of the objects, read as action possibilities by the AIP, into motor commands we can use to manipulate the object. This process is known as visuomotor transformation[10].

Note that these areas are activated, and the visuomotor transformation performed at the cortical level, by merely looking at the object, without that the action is performed. And both AIP and F5 respond to mere fixation of objects. Observation is sufficient to evoke visuomotor responses[11].

At the same time, visuomotor transformation leads to the neurophysiological simulation of the congruent motor schema related to the action we want to perform. However, this simulation is usually blocked by inhibitory mechanisms[12]. The presence of a motor resonance evoked by objects without that the action is performed is confirmed also by behavioral evidence[13].

Imagine the following scenario. You are thirsty and you note the presence of a bottle of water in front of you. At this point, the ventro-dorsal circuit is crucial for you to grasping the bottle. First of all, the AIP detects the geometrical features of the bottle that exhibit precise motor characteris-

[10] For a review from the point of view of computational neuroscience see Chinellato & Del Pobil (2016), Raos *et al*. (2006), Romero *et al*. (2014), Shikata *et al*. (2003), Murata *et al*. (1997), Sakata *et al.*(1995); for a philosophical review concerning the representational functions of this circuit see Ferretti (2016a), Ferretti (2016b), Ferretti (2016c), Ferretti (2016d), Ferretti (2017a), Ferretti (2018b), Ferretti (2017b), Ferretti (2019b), Zipoli Caiani & Ferretti (2016), Zipoli Caiani (2013), Zipoli Caiani (2014), Ferretti & Chinellato (2019), Chinellato, Ferretti and Irving (2019), Ferretti & Alai (2016).

[11] Milner & Goodale (1995/2006).

[12] Jeannerod (2006), Gallese (2000), Gallese (2005), Gallese (2007), Gallese (2009), Borghi & Cimatti (2010), Decety & Grèzes (2006), Borghi *et al.* (2010).

[13] Tucker & Ellis (1998), Ellis & Tucker (2000).

tics that are salient for your motor skills. The information about these characteristics is then sent to F5, which, given the information received by the AIP, computes the most suitable motor act we can use to grasp the bottle. Then, your visuomotor brain triggers the motor simulation congruent with the act to be used, independently of the fact that execution will either be performed or not.

Now, there is compelling evidence that the process of visuomotor transformation with the consequent motor simulation is triggered even when we are looking at depicted objects[14]. Indeed, the AIP[15], F5[16] and, in general, the dorsal stream[17] and its ventro-dorsal component[18] and the premotor areas[19] are active during the fixation of depicted objects[20].

Single-pulse TMS methodology can be adopted to asses more carefully the time-point of activation of the dorsal stream during the observation of action-related objects, showing at what stage the premotor system is actually involved in perceptual processing. Interestingly, in an experiment performed by (Buccino *et al*., 2009), motor-evoked potentials (MEPs) were recorded on the right hand while participants were being presented either with common images of mugs endowed with an handle, or with images of mugs deprived of the handle. At the same time, their motor area in the brain was stimulated 200 ms after the presentation of the visual stimulus. Results showed that MEPs were larger when the mug's handle was located on the side ipsilateral to the monitored hand than when the apparent graspable part was broken or located on the contralateral side of the monitored hand. More interestingly, in a different TMS experiment, (Cardellicchio *et al*., 2011) showed that MEP can be modulated also by the apparent spatial location of the represented object within or without the peripersonal space of the observer. Notably, this effect can be measured by delivering the electromagnetic stimulus already after 50 ms from the visual presentation of the images of graspable objects.

To this extent, it should be also noted that this visuomotor resonance depends on two perceptual facts. First, it is fostered only when the depicted object is apparently located within the subject's peripersonal action space of

[14] For a review see Ferretti (2016b), Ferretti (2016c), Zipoli Caiani (2013).
[15] Romero *et al*. (2013), Romero *et al*. (2014), Pani *et al*. (2014).
[16] Chao & Martin (2000).
[17] Chao & Martin (2000).
[18] Grézes & Decety (2002).
[19] Gerlach *et al*. (2002), Craighero *et al*. (2002).
[20] See also Proverbio *et al*. (2001), Buccino *et al*. (2009), Anelli *et al.* (2012), Costantini *et al*. (2010).

the observer. This is the space *within-our-reach* in which objects can be manipulated with a grasping action without the subject moves (walks) toward the object, or the object is moved toward the subject (ter Horst *et al*., 2011).When an object is represented as falling within this space, it can be, *ipso facto*, represented by the dorsal stream of the observer and its components as being apparently reach able. This is what happens in most of the experimental settings concerning the evidence of visuomotor resonance with depicted objects[21].

Second, this is possible because the dorsal stream cannot distinguish between normal and depicted objects[22]. Thus, once an object, whether depicted or real is presented in the visual scene, the dorsal stream activates the visuomotor transformation, with the consequent motor simulation. The 2-D structure of an object is sufficient to evoke visuomotor processing[23]. To this extent, it is the ventral stream that is able to distinguish between depiction and real objects – this process of identification is called 'response selection'. Accordingly, it is the ventral stream that selects the objects toward which we can trigger reliable action planning[24]. Only after the ventral stream performs response selection – whose result is consciously accessed during object recognition - and establishes that the object is suitable for action planning the dorsal stream triggers, on the basis of both visuomotor transformation and motor simulation, the parameters to act on the object – thought the ventral stream also contributes to the semantic representation of action[25] – see the distinction between 'ventral action planning' and 'dorsal motor processing'[26].

Finally, the reader should consider that the visuomotor responses are congruent with the action that would be appropriate for the depicted object, which is not the same appropriate to interact with the picture surface[27]. This point is in line with the fact that, most of the time, the action possibility elicited by the depicted object is different from the one recalled by the surface (e.g. a *power* grip vs. a *precision* grip). This is also in line with the evidence that attempting to grasp normal depicted objects is different from

[21] Cardellicchio *et al*. (2011), Chao & Martin (2000), Costantini *et al*. (2010), Romero *et al*. (2012), Romero *et al*. (2014), Cardellicchio *et al*. (2011), Tucker & Ellis (1998).
[22] Westwood *et al*. (2002).
[23] See Ferretti (2016b: 4.2).
[24] Westwood *et al*. (2002).
[25] For a review see Briscoe & Schwenkler (2015), Zipoli Caiani & Ferretti (2016), Ferretti (2016b), Ferretti (2016c), Ferretti (2016d), Ferretti (2018b).
[26] Dijkerman *et al*. (2009), Milner & Goodale (1995/2006), Briscoe & Schwenkler (2015), Ferretti (2018b).
[27] See the case by Buccino *et al*. (2009).

grasping real objects[28].

Let's now see the philosophical implications of this evidence for SPA.

5. Philosophical Implications

According to the 'dispositional notion of action possibilities' (see § 2) the optical stimulation induced by the representation of an object contains information about the presence of possible motor interactions only if the target suitable for action is really present. For this reason, visually ascribing possibilities of action to common images of objects can be understood as a mere case of *perceptual error* – or, what is usually called 'misperception', or 'misrepresentation'[29].

However, recent findings in the cognitive science of vision suggest that the illusory character of our perceptual experience of bi-dimensional objects, such as *trompe l'oeil*, as well as the veridical perceptual content related to normal depicted objects, are only part of the story. Our visual apparatus, indeed, is *naturally organized* so that there is a functional portion of this system that is specialized in detecting patterns of stimulation that are commonly, but not necessarily, associated with the presence of action possibilities in the environment. This functional specialization of the dorsal stream is made possible by the reuse of motor representations that are also involved in action execution, irrespectively to the fact that the source of the stimulation is a three-dimensional or a bi-dimensional object, or that the action can or will be executed or not (see below). It should be noted that, without relying on this visuo-motor specialization of the dorsal stream, it would be *virtually impossible* to assign so quickly a motoric meaning to the objects in the environment and to pragmatically access the manifold systems of representations that populate our environment**.** To this extent, note that dorsal stream processing is way faster than ventral one. This is due to what is known in the literature as its magnocellular advantage[30]. This is the reason why the former performs visuomotor transformation and triggers motor simulation before ventral processing performs response selection and

[28] Freud & Ganel (2015).

[29] For this point see Jacob & Jeannerod (2003), Orlandi (2010), Orlandi (2011); for the relation between misrepresentations, action properties and pictures, see Ferretti (2016b: 5.3).

[30] Barrett & Bar (2009), Bar *et al*. (2006), Bullier (2001), Laycock *et al*. (2007), Milner & Goodale (1995/2006).

action planning, which than make use of dorsal computations already stored in the visuomotor quiver.

The evidence above reported suggests that, at least at the subpersonal level, our visuomotor system can respond to the geometrical properties, offered by normal depicted objects, which in normal objects would be related to an action possibility. Of course, saying that normal depicted objects offer action possibilities does not mean that we can really manipulate them. And it does not mean that we consciously represent them as offering action possibilities. Rather, it suggests that some part of our dorsal visuomotor system, whose processing is not consciously accessible, behaves as if the action possibilities recalled by the features of the depicted object pertained to a real tridimensional object.

Hence, while normal depicted objects cannot be consciously perceived as offering reliable action possibilities, nonetheless, they can lead our dorsal subpersonal visuomotor system to perform both the visuomotor transformation and the motor simulation. But these processes are at the basis of the ascription of action possibilities. Thus, our subpersonal visuomotor system can ascribe action possibilities to normal depicted objects and, thus, we can unconsciously ascribe action possibilities. This confirms SPA.

Thus, the everyday stance which assumes that truly believed depicted objects cannot be perceived as affording action – even without recalling intentional action at the personal level, as in the case of *trompe l'oeil* pictures - turns out to be untrue on empirical grounds: visually ascribing pictorial action possibilities seem to be, even if at the unconscious level of subpersonal processing, rationally and intuitively, very reasonable, without the agent being guilty of irrationality. This also suggests that in order to ascribe action possibilities at the subpersonal level, we do not need perceive, neither consciously nor unconsciously the action-related property bearer as present as a real object is – we'll get back to this point in (§6). Indeed, we can ascribe action possibilities to normal depicted objects even if we can consciously recognize the bearer as being present, as one might say, only in the realm of pictures[31]. But this pictorial presence can be only recognized, consciously, due to the processing performed by the ventral stream. Our dorsal stream does not make any discrimination of this kind. Thus, it unconsciously attributes what we might call *pictorial action possibilities*, pertaining to *pictorial bearers*.

Rationally speaking, *normal* pictures can recall action to the subject. Though at the *personal* level of visual processing, only *trompe l'oeil* pic-

[31] Aasen (2015), Ferretti (2017c), Ferretti (2018b).

tures can be consciously *mis*perceived as offering action possibilities, at the *subpersonal* level of visual processing we can ascribe action possibilities also to normal *depicted* objects. Accordingly, we maintain that only with the former an agent could intend to act on targets that she believes are *normal* depicted objects, without being guilty of irrationality, although, in quality of pictures, *trompe l'oeil* do not really offer reliable possibility of action: they only fool our visual system that they do.

Summing up, it is possible to defend SPA on both philosophical and empirical grounds. We can rationally talk about pictorial action properties.

6. Beyond the Dispositional Idea of Action Possibilities

In the previous section we suggested that we can unconsciously ascribe action possibilities to *normal* depicted objects, that is, at the subpersonal level of visual processing. But this happens without that the agent consciously (i.e. at the personal level of visual processing) visually perceives this possibility, and that she visually represents, neither consciously nor unconsciously, the object as present. Nor she needs to believe that the action-related property bearer is present as a real object is - as in the case of *trompe l'oeil* pictures. And, of course, this unconscious ascription does not entail that she intends to act on the depicted object: the subject knows she cannot act on it and she does not visually represent the depicted object as really fostering reliable interaction.

Here we suggest that this idea renews our conception about the perception of action possibilities. We'll do that by refining the *requirements* at the basis of this conception. Let us start from the bottom.

According to the *specification requirement*, perceiving an action possibility cannot be divorced from the specification in the environment of an actual property bearer suitable for action. In this vein, visually ascribing a possibility of action in the absence of a real and present target for actual interaction might suggest the presence of a sort of perceptual error concerning the constraint of actualization. Therefore, it is a consequence of this view that either we perceive a possibility of action that depends on an action-related property bearer that is truly present in the environment, or there is not any possibility of action in the environmental situation we deal with, and thus we do not, and cannot, perceive it[32].

[32] Note that this conclusion coincides with the classic *disjunctivist* claim: e.g., Austin & Warnock (1979), Brewer (2007), Haddock & Macpherson (2011).

Interestingly, this *requirement* admits some exceptions. It can be argued that the idea that perception involves specification is in contrast with evolutionary assumptions[33]. For example, it is plain that natural evolution has not selected only organisms that are able to perceptually specify actual properties in the environment. Indeed, the variability among individual agents when performing visual tasks[34] can be better explained if the visual information exploited by the agent is not the result of a specifying mechanism. Rather, *suboptimal solutions* in which organisms are selected according to their ability in minimizing errors when detecting environmental properties are commonly admitted by natural evolution[35]. Accordingly, it is possible to assume that optic variables are only *statistically* associated with the environmental properties, so that sometimes the perceptual stimulus conveys the same information of a certain condition of the environment even without being causally associated with the presence of this condition. Cases such as this are possible because the agent is sensitive to the visual stimulation pattern, without any access to the presence of the sources this stimulation comes from. In this vein, evidence reviewed in (§4) shows that the human visual system is sensitive to sensorimotor patterns in the stimulus independently from the presence of actual action-related properties bearers in the environment. That is, the attribution of action possibilities is detached by the possibility of actual execution upon a real tridimensional object. It is sufficient to have the two dimensional geometrical pattern that usually instantiates an action possibility in a real, present object.

By way of this view, it is also possible to argue that when agents attribute action possibilities to the objects they deal with in the environment, they do not identify a real and present property bearer suitable for action. And this is in contrast with the previously introduced *identification requirement*. Indeed, according to the *identification requirement*, attributing action possibilities to the environment involves the actual presence of a suitable target for action. This amounts to assume that visually ascribing action possibilities to the environment implies the detection of the presence of action-related properties bearers. However, on looking carefully at the evidence introduced in (§ 4), this requirement clashes with the facts. Indeed, experimental results show that visually attributing action possibilities is not necessarily related to the real presence of targets suitable for action, inasmuch as the image of an object is normally enough to elicit the activation of the

[33] Withagen & Chemero (2012).
[34] Jacobs *et al.* (2001).
[35] Darwin (2005), Mayr (2002); for recent developments see Hoffman *et al*. (2012), Hoffman *et al*. (2013), Hoffman *et al*. (2015).

visual dorsal stream even if there is not any actual target for action. This is because the visual system is sensitive to the sensorimotor patterns that are common between both the real and the depicted object, and does not distinguish among the two conditions. Of course, at the personal level the *identification requirement* holds: with normal objects we veridically perceive that the action possibility is attached to a real object; with *trompe l'oeil* we are fooled that the action possibility is attached to a real object. But at the subpersonal level, this is not necessary: the visuomotor resource that detects action possibilities is not able to detect real presence. And this is due to the fact that humans rely on ventral/dorsal interactions in which ventral processing detects 'real presence for interaction' and dorsal processing computes the motor parameters. But, that said, the evidence about dorsal computational responses is sufficient to go beyond the *identification requirement*.

It has been shown, indeed, that the detection of specific patterns of visual stimulation of a bi-dimensional representation of an object might influence the planning and the execution of actions according to the possibilities that are evoked by the represented object. This is precisely the case illustrated by much of the evidence concerning the presence of compatibility effects occurring when agents are asked to execute a motor task that is congruent with one of the possible actions afforded by the visual target[36] showed that this type of compatibility effect is associated with the apparent possibility of interacting with a real or represented visual target. In order to confirm the hypothesis that what is relevant in the visual ascription of action possibilities to the environment is the detection of patterns of stimulation, instead the occurrence of real opportunities of actual possibility of interaction provided by real, present objects, the authors divided the apparent space into reachable and non-reachable spaces by showing the image of a mug either in front of or behind the image of a semi-transparent panel. Interestingly, the authors found a compatibility effect only when the image of the target was *apparently reachable* by the agent, that is, when it was apparently located in front of the panel and within the agent's space of action. Similarly, Vainio *et al*. (2011) have revealed the presence of a compatibility effect when an agent executed actions while observing the normal image of a target of action, but not when he or she observed the orientation of a depicted arrow.

Furthermore, the visual attribution of action possibilities to images of possible targets of action influences spatial attention even when the agent is not required to recognize the depicted object. Handy *et al*. (2003), for ex-

[36] Costantini *et al*. (2010), Ellis & Tucker (2000).

ample, showed that agents who are presented with images of action-related targets direct their attention in the visual hemifield that is dominant for the execution of the suggested action, re-allocating their attention in the direction of the suggested motor interaction. This evidence advocates the hypothesis that information concerning possibilities of action can be extracted even when the agent is not able to semantically recognize the identity of the visual target, suggesting that conceptual abilities are not required for the visual attribution of action possibilities[37].

Experiments performed on patients who suffer from extinction can be considered of particular relevance to illustrate this phenomenon. Extinction patients, indeed, show a deficit in visual attention to stimuli that are presented toward the contralesional side of the visual field, while competing stimuli are presented to the other side of the visual field. Interestingly, Di Pellegrino *et al.* (2005) showed that extinction was reduced in subjects who were presented with images of pairs of cups that were endowed with handles in both of the visual fields, but no reduction was registered when the cups were deprived of their functional part. This means that it is easier for the normal image of a cup endowed with an apparent functional part to elicit apparent motor interactions, enabling patients suffering from extinction to report their presence in the visual field.

To sum up, it is a matter of fact that merely viewing the image of an object elicits the related motor plans required for the execution of real actions[38]. This explains why we observe a facilitation effect when the action-related stimulus is compatible with the effect or we habitually use to interact with it. The previous conclusion based on empirical evidence makes it disputable the assumption that perceiving action possibilities necessarily involves the actual presence of action-related dispositional properties in the environment. Rather, what the visual system is sensible to are sensorimotor patterns of stimulation, instead of actual environmental properties.

Interestingly, as noted in (§3), the actualization of an action depends on the possibility of identifying a property bearer suitable for action, so that the *actualization requirement* depends on the *identification requirement*. Indeed, since the activation of the dorsal visual stream is possible even without involving the actual presence of any object suitable for action (see §4), we are led to exclude that visually attributing action possibilities to the environment necessarily involves the actualization of an action. After all, it

[37] This is in line with our premotor theories of attention Craighero *et al*. (1999), Adamo & Ferber (2009), see also Riddoch (2003).
[38] McBride *et al.* (2012), Roberts & Humphreys (2011).

cannot be necessary to foster a positive possibility to actualize an action if there is no necessity to postulate the presence in the environment of a suitable target for action. Otherwise, it would be like claiming that we should witness a solubility related effect, without needing to postulate the presence of any soluble object in the environment. Similarly, contrary to the *actualization requirement*, it would be meaningless to assume that when attributing action possibilities to the objects the environment offers to us there would necessarily be a positive probability of actualization even without any necessity to identify a suitable action-related property bearer in the nearby. And, as we saw, dorsal visuomotor processing does not respect this constraint: it cannot even rely on the computations necessary to respect the *actualization requirement*. This is courtesy of ventral processing.

To sum up, in light of the experimental literature, it may be argued that, contrary to the *specification requirement*, visually attributing action possibilities to the environment is not a process of relating visual stimulations to action-related properties bearers. Accordingly, since agents commonly attribute action possibilities to the environment independently from the actual presence of a suitable referent for action, there is no need to endorse the *identification requirement* as well. As a final consequence, since action possibilities can be visually ascribed without necessarily postulating the actual presence of a real action-related property bearer, there is no necessity to foster a positive probability to execute an action. This implies that also the *actualization requirement* should be rejected.

But there is also another point we want to suggest here. We saw that when we look at a depicted object, or at a normal object, our brain simulates the action that can be performed. Several cerebral and physiological processes related to action generation are triggered[39].This simulation is independent of whether over execution will follow. In the case of normal objects, we do not need to necessarily act. In the case of normal objects, we simply can't (Ibid.).To this extent, there is a crucial difference, addressed in the literature on motor neuroscience. On the one hand we have mental action, which is a sort of covert action processing (Ibid.) related to the trigger of motor resources above mentioned, without that we actually act. On the other hand we have overt action execution, or motor action, which is motor processing used until the action unfolds[40].

We also saw that (§2), according to Turvey, possibilities of actions are relational properties of environmental things that, given suitable circum-

[39] Jeannerod (2006).
[40] Jeannerod (2006).

stances, necessarily actualize[41]. Against Turvey, Scarantino (2003) suggested that action possibilities should be thought not in terms of necessary actualization, but in terms of the probability of an action manifestation.

This suggests the following distinction. Scarantino's insight about action possibilities captures the nature of overt action response, which in case of depicted objects cannot actualize, while in the case of real objects does not necessarily actualize. On the contrary, Turvey's insight about action possibilities captures the nature of covert action, which always actualizes, with both real and depicted objects: the ascription of action possibilities recalled by the visuomotor processing of the geometrical arrangement that recall them, along with the simulation of the motor act we might perform upon them – in the case of normal object, if we want to, and in the case of depictions, were the depicted object a normal one - necessarily actualizes.

But - and this is crucial - our account, which goes beyond the dispositional one, can also distinguish between this case action possibilities that activate a subpersonal motor response even if the subject does not really believe, at the personal level, that the object is present, and the situation in which the subject believes the action possibility can actualize, but it cannot, as in the case of *trompe l'oeil*. Thus, we can extend the theory of action possibilities even further, by addressing the following possible scenarios:

(a) We believe, or consciously perceive as if an action possibility can actualize with respect to overt action and, indeed, it really can. This is Scarantino's scenario pertaining real objects.

(b) We don't believe, or consciously perceive as if an action possibility can actualize with respect to overt action and, indeed, it really can't. But, however, it always actualizes with respect to covert action. This is the case of neurophysiological motor simulation captured by Turvey's scenario with depicted objects.

(c) We believe, or consciously perceive as if an action possibility can actualize with respect to overt action, but it cannot, while, however, it can actualize concerning neurophysiological motor simulation. This is the *trompe l'oeil* case.

[41] Turvey *et al*. (1981), Shaw *et al*. (1982).

7. Conclusion: Pictures, Presence and Pictorial Action Possibilities

We saw (§2) that *trompe l'oeil* pictures foster in the subject the illusion, at the *personal* level of visual processing, of affording suitable possibility of action, while *normal* pictures foster in the subject the illusion, at the *subpersonal* level of visual processing, of affording suitable possibility of action.

Accordingly, the common stance suggests that the ascription of action possibility is related to the feeling of presence. That's the reason why, following this stance, only *trompe l'oeil* pictures, but not *normal* pictures, allow the ascription of action possibilities to the depiction. However, we saw that it is reasonable to suggest, on both empirical (§3) and philosophical grounds (§4) that even *normal* pictures allow the ascription of action possibilities. Against the common stance, it is possible to defend the SPA. And while we maintain that with *trompe l'oeil* pictures the ascription of action possibilities is about at the *personal* level of visual processing, with *normal* pictures this ascription only concerns the *subpersonal* level of visual processing. But only the former can foster in us the feeling of presence. This means that, while along with the ascription of action possibilities at the *personal* level, we also visually represent, neither consciously nor unconsciously, the depicted object as present, and, usually, we also believe, even for a moment, it is present, with the ascription of action possibilities at the *subpersonal* level, we do not need to visually represent, neither consciously nor unconsciously, the depicted object as present as well as we do not believe that it is.

Thus, as a matter of fact, and against the more rational of the intuitions, not all the forms of perception of action possibilities require the so-called *visual feeling of presence*. Indeed, while only normal objects and *trompe l'oeil* pictures foster this feeling and, also, recall intentional action at the personal level, depicted objects can, without neither fostering this feeling, nor recalling intentional action, recall the perception of action possibilities at the subpersonal level of visuomotor processing. And even so, we can rationally talk about pictorial action possibilities.

To this extent, our proposal is also very important in relation to the notorious ecological theory of vision by the father of the notion of 'action possibility', namely J.J. Gibson, who, however, usually used the term 'affordance'[42]. Gibson had a very deep interest for pictures[43]. Bearing in mind

[42] We do not use this term here because it often denotes an anti-representationalist stance concerning vision, which we would like to avoid in this paper, see Chemero (2009), Zipoli Caiani (2013), Tillas *et al*. (2016), Ferretti (2016d), Ferretti & Alai (2015), Ferretti (2019b).

his insights, we have shown that depicted objects are salient objects for action *tout-court*. *Normal* depicted objects are salient for our subpersonal motor processing. And this is due to the unconscious nature, as well as to the computational characteristics of dorsal vision. But *trompe l'oeil* are salient even for our conscious perception: we tend to perceive the depiction as being present and as offering suitable motor interaction because the *trompe l'oeil* is masterfully built in such a way that it is able to fool our ventral vision's response selection and to direct ventral action planning (§4) toward it, even if the object is depicted. Finally, detecting an actual action possibility as pertaining to the picture's surface, perceived as present, is what allows us not to be fooled that the depicted object is a present one we can act upon[44], as in the case of *trompe l'oeil* depicted objects. That means that a theory of pictorial action possibilities is not only important for our accounts of action possibilities, but also to deeply understand picture perception, and its relation with face-to-face perception. And, as already suggested, recognizing that also action can be present in picture perception makes it more similar to face-to-face perception than previously thought[45].

Funding: This project was supported by the 'Fondazione Franco e Marilisa Caligara per l'alta formazione interdisciplinare' and by a NOMIS Fellowship, awarded by the Eikones - Center for the Theory and History of the Image at the University of Basel, Switzerland.

References

Aasen, S. (2015), "Pictures, Presence and Visibility", in *Philosophical Studies*. doi: 10.1007/s11098-015-0475-4.

Anelli, F., Borghi, A. M., & Nicoletti, R. (2012), "Grasping the Pain: Motor

[43] Gibson (1947), Gibson (1954), Gibson (1960), Gibson (1971), Gibson (1978); see also Cutting (2000).
[44] Vishwanath (2014), Ferretti (2016d).
[45] Briscoe (2016), Ferretti (2016d).

Resonance with Dangerous Affordances", in *Consciousness and Cognition*, *21*, 1627–1639.

Austin, J. L., & Warnock, G. J. (1979), *Sense and Sensibilia*, London: OUP USA.

Bar, M., Kassam, K.S., Ghuman, A.S., Boshyan, J., Schmid, A.M., Dale, A.M., Hamalainen, M.S., Marinkovic, K., Schacter, D.L., Rosen, B.R., & Halgren, E. (2006), *Top-down Facilitation of Visual Recognition*, Proceedings of the National Academy of Sciences of the United States of America, 103: 449–454. doi:10.1073/pnas.0507062103.

Barrett, L.F., & Bar, L.F. (2009), "See It with Feeling: Affective Predictions During Object Perception", in *Philosophical Transactions of the Royal Society*, 364: 1325–1334. doi:10.1098/rstb.2008.0312.

Borghi, A. M., & Cimatti, F. (2010), "Embodied Cognition and Beyond: Acting and Sensing the Body", in *Neuropsychologia*, *48*, 763–773.

Borghi, A. M., Gianelli, C. & Scorolli, C. (2010), "Sentence Comprehension: Effectors and Goals, Self and Others. An Overview of Experiments and Implications for Robotics", in *Frontiers in Neurorobotics*, *4*(3). doi:10.3389/fnbot.2010.00003.

Brewer, B. (2007), "Perception and Its Objects", in *Philosophical Studies*, *132*(1), 87–97. https://doi.org/10.1007/s11098-006-9051-2

Briscoe, R., & Schwenkler, J. (2015), "Conscious Vision in Action", in *Cognitive Science*, 39(7), 1435–1467.

Buccino, G., Sato, M., Cattaneo, L., Rodà, F., & Riggio, L. (2009), "Broken Affordances, Broken Objects: A TMS Study", in *Neuropsychologia*, *47*, 3074–3078. doi:10.1016/j.neuropsychologia.2009.07.003.

Bullier, J. (2001) "Integrated Model of Visual Processing. Brain Research", in *Brain Research Reviews*, 36: 96–107.

Cardellicchio, P., Sinigaglia, C., & Costantini, M. (2011), "The Space of Affordances: A TMS Study", in *Neuropsychologia*, *49*(5), 1369–1372.

Chao, L. L., & Martin, A. (2000), "Representation of Manipulable

Manmade Objects in the Dorsal Stream", in *NeuroImage*, *12*, 478–484.

Chemero, A. (2009), *Radical Embodied Cognitive Science*, Cambridge, MIT Press.

Chinellato, E., & Del Pobil, A. P. (2016), *The Visual Neuroscience of Robotic Grasping. Achieving Sensorimotor Skills Through Dorsal-ventral Stream Integration*, Switzerland, Springer International Publishing.

Chinellato E., Ferretti G., Irving L. (2019), "Affective Visuomotor Interaction: A Functional Model for Socially Competent Robot Grasping", in Martinez-Hernandez U. *et al.* (eds), *Biomimetic and Biohybrid Systems. Living Machines 2019. Lecture Notes in Computer Science*, vol 11556. Springer, Cham.

Costantini, M., Ambrosini, E., Tieri, G., Sinigaglia, C., & Committeri, G. (2010), "Where Does an Object Trigger an Action? An Investigation about Affordances in Space", in *Experimental Brain Research*, *207*(1–2), 95–103. https://doi.org/10.1007/s00221-010-2435-8

Craighero, L., Bello, A., Fadiga, L., & Rizzolatti, G. (2002), "Hand Action Preparation Influences the Responses to Hand Pictures", in *Neuropsychologia*, *40*(5), 492–502.

Cutting, J. E. (2000), "Images, Imagination, and Movement: Pictorial Representations and Their Development in the Work of James Gibson", in *Perception*, 29, 635-648. doi: 10.1068/p2976.

Darwin, C. R. (2005). *The Origin of Species by Means of Natural Selection*. United States, Adamant Media Corporation.

Decety, J., & Grèzes, J. (2006), "The Power of Simulation: Imagining One's Own and Other's Behavior", in *Brain Research*, *1079*, 4–14.

Di Pellegrino, G., Rafal, R., & Tipper, S. P. (2005), "Implicitly Evoked Actions Modulate Visual Selection: Evidence from Parietal Extinction", in *Current Biology*, *15*(16), 1469–1472. https://doi.org/10.1016/j.cub.2005.06.068.

Dijkerman, H. C., McIntosh, R. D., Schindler, I., Nijboer, T. C. W., & Milner, A. D. (2009), "Choosing Between Alternative Wrist Postures:

Action Planning Needs Perception", in *Neuropsychologia*, 47(6), 1476–1482.

Ellis, R., & Tucker, M. (2000), "Micro-affordance: The Potentiation of Components of Action by Seen Objects", in *British Journal of Psychology*, *91*(4), 451–471. https://doi.org/10.1348/000712600161934

Ferretti, G. (2016a), "Neurophysiological States and Perceptual Representations: The case of Action Properties detected by the Ventro-Dorsal Visual Stream". In Magnani, L., & Casadio, C. (Eds.), *Model-Based Reasoning in Science and Technology. Models and Inferences: Logical, Epistemological, and Cognitive Issues*, Series "Sapere", Springer.

Ferretti, G. (2016b), "Pictures, Action Properties and Motor Related Effects", in *Synthese Special Issue Neuroscience and Its Philosophy*, 193(12): 3787–3817. doi:10.1007/s11229-016 1097-x.

Ferretti, G. (2016c), "Through the Forest of Motor Representations", in *Consciousness and Cognition*, 43: 177–196. doi:10.1016/j.concog.2016.05.013.

Ferretti, G. (2016d), "Visual Feeling of Presence", in *Pacific Philosophical Quarterly*. doi:10.1111/papq.12170.

Ferretti G. (2017a), "Pictures, Emotions, and the Dorsal/Ventral account of Picture Perception", in *Review of Philosophy and Psychology*. doi 10.1007/s13164-017-0330-y.

Ferretti, G. (2017b), "Two Visual Systems in Molyneux Subjects", in *Phenomenology and the Cognitive Sciences*, 17(4), pp. 643-679. doi: https://doi.org/10.1007/s11097-017-9533-z

Ferretti, G. (2017c), "Are Pictures Peculiar Objects of Perception?", in *Journal of the American Philosophical Association*, 3(3), pp. 372–393. doi: 10.1017/apa.2017.28.

Ferretti, G. (2018a), "The Nature of Pictorial Representations", in *Phenomenology and Mind*, 14, pp. 136-144. doi: 10.13128/Phe_Mi-23631.

Ferretti, G. (2018b), "The Neural Dynamics of Seeing", in *Erkenntnis*, doi: https://doi.org/10.1007/s10670-018-0060-2.

Ferretti, G. (2019a), "Perceiving Surfaces (and What They Depict)", in *The Senses and the History of Philosophy,* Edited by Glenney, B. and Silva, J.F., Routledge.

Ferretti, G. (2019b), "Visual Phenomenology versus Visuomotor Imagery: How Can we be Aware of Action Properties?" in *Synthese*, https://doi.org/10.1007/s11229-019-02282-x.

Ferretti, G. (Forthcoming a), "Why *Trompe l'oeils* deceive our Visual Experience", *The Journal of Aesthetics and Art Criticism*.

Ferretti, G. (Forthcoming b), "Do Trompe l'Oeils look right when viewed from the wrong place?", in *The Journal of Aesthetics and Art Criticism*.

Ferretti G., & Alai, M. (2016), "Enactivism, Representations and Canonical Neurons", in *Argumenta,* 1 (2).

Ferretti G. and Chinellato E. (2019), "Can Our Robots Rely on an Emotionally Charged Vision-for-Action? An Embodied Model for Neurorobotics", in Vallverdú J., Müller V. (eds), *Blended Cognition, The Robotic Challenge*. Springer Series in Cognitive and Neural Systems, vol 12. Springer, Cham.

Freud E., & Ganel T. (2015), "Visual Control of Action Directed Toward Two-dimensional Objects Relies on Holistic Processing of Object Shape", in *Psychon Bull Rev*, 22: 1377–1382. DOI 10.3758/s13423-015-0803-x.

Gallese, V. (2000), "The Inner Sense of Action. Agency and Motor Representations", in *Journal of Consciousness Studies*, *7*(10), 23–40.

Gallese, V. (2007). "The *Conscious* Dorsal Stream: Embodied Simulation and Its Role in Space and Action Conscious Awareness", in *PSYCHE, 13*(1), 1–20.

Gallese, V. (2009), "Motor Abstraction: A Neuroscientic Account of How Action Goals and Intentions are Mapped and Understood", in *Psychological Research*, *73*, 486–498. doi:10.1007/s00426-009-0232-4.

Gerlach, C., Law, I., & Paulson, O. B. (2002), "When Action Turns into Words. Activation of Motor-based Knowledge during Categorization of

Manipulable Objects", in *Journal of Cognitive Neuroscience*, *14*, 1230–1239.

Gibson, J. J. (Ed.) (1947), "Motion Picture Testing and Research", in *Report No7, Army Air Force Aviation Psychology Program Research Reports* (Washington, DC: US Government Printing Office).

Gibson, J. J. (1954), "A Theory of the Perception of Pictures", in *Audio-Visual Communication Review*, 1,3-23.

Gibson, J. J. (1960), "Pictures, Perspective, and Perception", in *Daedalus*, 89 216-277.

Gibson, J. J. (1967), "New Reasons for Realism", in *Synthese*, *17*(1), 162–172. https://doi.org/10.1007/BF00485025

Gibson, J. J. (1971), "The Information Available in Pictures", in *Leonardo*, 4, 27-35.

Gibson, J. J. (2014), *The Ecological Approach to Visual Perception: Classic Edition*, Psychology Press.

Grézes, J., & Decety, J. (2002), "Does Visual Perception of Object Afford Action? Evidence from a Neuroimaging Study", in *Neuropsychologia*, *40*, 212–222.

Haddock, A., & Macpherson, F. (2011), *Disjunctivism: Perception, Action, Knowledge*, Oxford - New York, OUP Oxford.

Handy, T. C., Grafton, S. T., Shroff, N. M., Ketay, S., & Gazzaniga, M. S. (2003), "Graspable Objects Grab Attention When the Potential for Action Is Recognized", in *Nature Neuroscience*, *6*(4), 421–427. https://doi.org/10.1038/nn1031

Hoffman, D.D., & Singh, M. (2012), "Computational Evolutionary Perception", in *Perception* 41, pp. 1073–1091.

Hoffman, D. D., Singh, M., & Prakash, C. (2015), "The Interface Theory of Perception", in *Psychon Bull Rev*, 22, pp. 1480-1506.

Hoffman, D.D., Singh, M., & Mark, J. T. (2013), "Does Natural Selection Favor True Perceptions?", in *Proceedings of the SPIE*, Bellingham,

WA.

Hopkins, R. (2012), "Seeing-in and Seeming to See", in *Analysis*, 72, 650–659.

Jacob, P., & Jeannerod, M. (2003), *Ways of Seeing. The Scope and Limits of Visual Cognition*, Oxford, Oxford University Press.

Jacobs, D. M., Runeson, S., & Michaels, C. F. (2001), "Learning to Visually Perceive the Relative Mass of Colliding Balls in Globally and Locally Constrained Task Ecologies", in *Journal of Experimental Psychology. Human Perception and Performance*, *27*(5), 1019–1038.

Jeannerod, M. (2006), *Motor Cognition: What Actions Tell the Self*, Oxford, Oxford University Press.

Laycock, R., Crewther, S.G., & Crewther, D.P. (2007), "A Role for the B 'Magnocellular Advantage' in Visual impairments in neurodevelopmental and psychiatric disorders", in *Neurosci Biobehav Rev*, 31(3):363-76. Epub 2006 Dec 1.

Lopes, D. M. (2005), *Sight and Sensibility*, Oxford, Oxford University Press.

Matthen, M. (2005), *Seeing, Doing and Knowing: A Philosophical Theory of Sense Perception*, Oxford, Oxford University Press.

Mayr, E. (2002), *What Evolution Is* (Reprint edition), New York, Basic Books.

McBride, J., Sumner, P., & Husain, M. (2012), "Conflict in Object Affordance Revealed by Grip Force", in *The Quarterly Journal of Experimental Psychology*, *65*(1), 13–24. https://doi.org/10.1080/17470218.2011.588336

Michaels, C. F., & Carello, C. (1981), *Direct perception*, Prentice-Hall Englewood Cliffs, NJ.

Milner, A. D., & Goodale, M. A. (1995/2006), *The visual brain in action*, Oxford, Oxford University Press, 2nd Edition.

Murata, A., Fadiga, L., Fogassi, L., Gallese, V., Raos, V., & Rizzolatti, G. (1997), "Object representation in the ventral premotor cortex (area F5) of the monkey", in *Journal of Neurophysiology*, 78, 2226–2230.

Nanay, B. (2011), "Perceiving Pictures", in *Phenomenology and the Cognitive Sciences*, 10, pp. 461–480.

Nanay, B. (2013), *Between perception and action*, Oxford, Oxford University Press.

Nanay, B. (2017). "Threefoldness", in *Philosophical Studies*, 1–20, doi: 10.1007/s11098-017-0860-2.

Orlandi, N. (2010), "Are sensory properties represented in perceptual experience?", in *Philosophical Psychology*, 23 (6), pp. 721-740.

Orlandi, N. (2011), "Embedded Seeing: Vision in the Natural World", in *NOÛS*, 00:0 1–23.

Pani, P., Theys, T., Romero, M. C., & Janssen, P. (2014), "Grasping execution and grasping observation activity of single neurons in the macaque anterior intraparietal area", in *Journal of Cognitive Neuroscience*, *26*(10), 2342–2355. doi:10.1162/jocn_a_00647.

Proverbio, M. A., Adorni, R., & D'Aniello, G. E. (2011), "250 ms to code for action affordance during observation of manipulable objects", in *Neuropsychologia*, *49*, 2711–2719.

Raos, V., Umiltà, M. A., Murata, A., Fogassi, L., & Gallese, V. (2006), "Functional properties of grasping-related neurons in the ventral premotor area F5 of the macaque monkey", in *Journal of Neurophysiology*, 95, 709–729.

Riddoch, M. J., Humphreys, G. W., Edwards, S., Baker, T., & Willson, K. (2003), "Seeing the action: neuropsychological evidence for action-based effects on object selection", in *Nature Neuroscience*, *6*(1), 82–89. https://doi.org/10.1038/nn984

Roberts, K. L., & Humphreys, G. W. (2011), "Action relations facilitate the identification of briefly-presented objects", in *Attention, Perception, & Psychophysics*, *73*(2), 597–612.

https://doi.org/10.3758/s13414-010-0043-0

Romero, M. C., Pani, P., & Janssen, P. (2014), "Coding of shape features in the macaque anterior intraparietal area", in *The Journal of Neuroscience*, 34(11), 4006–4021.

Romero, M. C., Van Dromme, I., & Janssen, P. (2013). "The role of binocular disparity in stereoscopic images of objects in the macaque anterior intraparietal area", in *PLOS ONE*. http://dx.doi.org/10.1371/journal.pone. 0055340.

Sakata, H., Taira, M., Murata, A., & Mine, S. (1995), " Neural mechanisms of visual guidance of hand action in the parietal cortex of the monkey", in *Cerebral Cortex*, 5, 429–438.

Scarantino, A. (2003), "Affordances Explained", in *Philosophy of Science*, *70*(5), 949–961. https://doi.org/10.1086/377380

Shikata, E., Hamzei, F., Glauche, V., Koch, M., Weiller, C., Binkofski, F., & Büchel, C. (2003), "Functional properties and interaction of the anterior and posterior intraparietal areas in humans", in *European Journal of Neuroscience*, 17(5), 1105–1110.

Shaw, R. E., Turvey, M., & Mace, W. M. (1982), "Ecological Psycology. The consequence of a commitment to realism", in Weimer, W., & Palermo, D. (Eds.), *Cognition and the symbolic processes* (pp. 159–226), Hillsdale, Lawrence Erlbaum Associates, Inc.

Gallese, V. (2005), "Embodied simulation: From neurons to phenomenal experience", in *Phenomenology and the Cognitive Sciences*, *4*, 23–48.

Tillas, A., Vosgerau, G. Seucher, T., & Zipoli Caiani, S. (2016), "Can Affordances Explain Behavior?", in *Review of Philosophy and Psychology* http://link.springer.com/article/10.1007%2Fs13164-016-0310-7

Tucker, M., & Ellis, R. (1998), "On the relations between seen objects and components of potential actions", in *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 830–846.

Turvey, M. T., Shaw, R. E., Reed, E. S., & Mace, W. M. (1981), "Ecological laws of perceiving and acting: in reply to Fodor and Pylyshyn", in *Cognition*, *9*(3), 237–304.

Vainio, L., Hammarén, L., Hausen, M., Rekolainen, E., & Riskilä, S. (2011), "Motor inhibition associated with the affordance of briefly displayed objects", in *The Quarterly Journal of Experimental Psychology*, *64*(6), 1094–1110. https://doi.org/10.1080/17470218.2010.538221

Westwood, D., Danckert, J., Servos, P., & Goodale, M. (2002), "Grasping two-dimensional images and three-dimensional objects in visual-form agnosia", in *Experimental Brain Research*, *144*, 262–267.

Withagen, R., & Chemero, A. (2012), "Affordances and classification: On the significance of a sidebar in James Gibson's last book", in *Philosophical Psychology*, *25*(4), 521–537. https://doi.org/10.1080/09515089.2011.579424

Zipoli Caiani, S. (2014), "Extending the notion of Affordance", in *Phenomenology and the Cognitive Sciences,*13*,* 275–293. ISSN:1568-7759.

Zipoli Caiani, S. (2014), "Framing Visual Perception in Terms of Sensorimotor Mapping", in Prinz, J., & Machery, E., *Perception and Concepts.* ISSN: 1944-3676.

Zipoli Caiani, S. (2016), " Through the flat Canvas. The Motor Meaning of Realistic Paintings", in *Aisthesis.* http://dx.doi.org/10.13128/Aisthesis-19425.

Zipoli Caiani, S., & Ferretti, G. (2017) "Semantic and Pragmatic Integration in Vision for Action", in *Consciousness and Cognition,* 48, 40–54, http://dx.doi.org/10.1016/j.concog.2016.10.009.

(Ir)Rational Listeners: Music and the Construction of the Self

Alessandro Giovannucci
University of Teramo
alessandro.giovannucci@gmail.com

Music always has had a strong connection with logic and reasoning. It is a common idea that proportions and form are the bases for an effective perception of music. This paradigm, in which rationality and music often overlap, undergoes several modifications when sound and body become increasingly relevant to the musical experience. How can rationality deal with the instability of sound and the irrational appetite of the body?

The rational aspect of music is based on the graphocentrism that animates the well-learned musical tradition. The emergence of new theories about listening and the inner life of sound propagated a series of interesting new perspectives on rationality and music. The works of Derrida, Nancy and Erlmann have pointed out that musical studies have neglected resonance for too long. While rationality is based on the separation of subject and object, resonance involves their connection. Rationality and resonance – the brain and the ear - are conflicting but necessary in music.

Nowadays our musical culture is rooted on this dichotomy, according to which there is "serious" music, based on the silenced experience that originates firstly from the musical score, and an "emotional" one that is a non-mediated reaction to sounds. Technologies and new musical practices reveal an inner life of sound that challenges traditional ideas about music perception and the Cartesian paradigm of musical experience.

The rationalist bases of music are based on a very old tradition that managed to impose themselves thanks to a series of discoveries and scientific epochal processes that took place at the end of the 1700s and the first half of the 1800s; more precisely, the tradition maintains that there is a connection between musical perception and the inner ear. This connection

Alessandro Giovannucci, "Ir(R)ational Listeners: Music and the Construction of the self", in Giorgio Grimaldi and Marialuisa Parise, Rational *Animals. Seven essays on Reason's fields of application*, pp. 149-157

finds its greatest expression in what was considered for some time as the most appropriate treaty explaining the scientific matrices of musical processes; namely, *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik.*

This treaty was published by Hermann von Helmholtz in 1863 and is undoubtedly unique because of its detail, rigour and research methodology. In the course of the time, it has been greatly critiqued, and in certain instances understandably so. It has encountered severe criticism for its reductionism as a founding principle for musical thinking towards acoustic physiology that it proposes. According to Helmholtz, music and nature generate a sort of *kalokagathìa,* that is*,* a "civilising" force, in that it is animated by a rational and proportionate force. In any case, Helmholtz's work is also sustained by others. It is a musical by-product of a rich series of studies on perceptions that are intended to provide a body of knowledge on cognitive processes. *Lehre* contains various approaches on musical issues, that mainly based on the relation between acoustic and physical approaches. Helmholtz explains in detail vibratory phenomena, the internal complexity of sound and its relation to mathematics, connecting them to the physiology of the listener and explaining the laws that govern auditory senses, more specifically in relation to those connected with consonance and dissonance.

Further, the treaty deals with approaches to musical aesthetics. Helmholtz evidences the logic of musical construction in various epochs and among various ethno-social groups, that take into consideration taste, choice and individual freedom. This aestheticism does not derive from physical or physiological laws; indeed, *Lehre* has among its objectives that of illustrating the scientific foundations of music, that is, as a form of artistic discipline.

Helmoltz is an authoritative promoter, if not the only promoter, of the union between harmony and acoustic science. It is important to recall that theories of this type have ancient origins, which, according to the German scientist are lacking in scientific rigour and fall into mysticism and numerology, since "even in very recent times, there are far too many friends of music that are inclined towards speculation, preferring to take delight in arithmetic mysticism rather than being concerned with experimenting harmonious sounds"[1].

Helmhotz prefers purely mechanical thinking to the linear thinking that started with Pythagoras and ended with D'Alembert; mechanical thinking being based exclusively on scientific principles. Even aesthetic sensations

[1] Helmholtz (1986: 375).

are considered to be the by-product of mechanical and physical processes, are concisely explained in mechanical and physical terms. According to Helmoltz, the ear could be considered as a measuring device:

> The ear does the same thing that mathematics does according to Fourier's theory […] - that is, it decomposes sound waves that do not correspond, like the sounds of diapason, into a simple wave that comprises the sum of simple waves, isolating and collecting the sounds that belong to each of these - independently of whether the wave was transmitted from a source or became such en route.[2]

In any case, the complexity of the musical experience is perceptible only to the careful and trained listener. But, if the ear – as long as it is not impeded in its functions – is capable of registering all the variables of sound, why do all listeners not have the same reactions to the same musical stimuli? This is because human, psychological and cultural factors are an integral part of the musical appreciation process. As Serravezza explains, it is all due to

> the very nature of perception, where certain factors influence in certain ways, and at the same time take distance from data that is received, becoming a purely simple sensation. […]. Attention and habit play a decisive role in sensory observation, and the requirements of everyday life guide the formation of perception schemes, dictating the means according to which attention is exercised and determined by consciousness.[3]

Further, according to Serravezza, Helmholtz who included these considerations in his writings subsequent to *Lehre*, was constantly undecided on the role of sensation and perception in the reception of music, because Helmholtz's experimental rigour did not allow him to negate given reactions that had not been clarified. In fact, the psychological aspects of musical reception remained secondary to physiological aspects. Helmholtz never completely resolved the fissure between the scientific dimension of sound and the subjectivity of the listener, rather he generated a sort of "undulatory motion" between the two that alternatingly prevailed. In more recent times, rationalism has lost ground to research that prefers to question the aforementioned concepts that had previously been discarded because they were not readily placeable or analysable; for example, the body and experiential phenomena relating to the body. As evidenced by Enrica Lisciani Petrini:

[2] Serravezza (1996: 23).
[3] *Ibidem*.

> A centuries old way of conceiving things, whose way was paved by Plato, and that had at its basis western culture – it being a way of conceiving reality as a web of immobile and unchangeable ideas – not only fixed the fissure between the noun - that is concept, intelligible thought - and the thing - that is the executed sensory and bodily reality – with the consequence of the latter being relegated to mere and irrelevant appearance. It, also, above all, fixed the fissure between the word, or its "meaning", and its phonetic "colour" – as well as the bodily and vocal timbre that inseparably accompanies it - sanctioning it.[4]

The notion of a "muted" intellect, that is, a unique instrument that is capable of perceiving the sense of the sound vibration, reaches a crisis because definable form and rational identity are not the only elements pertaining to musical perception, because the insurmountable limits between the subject and object become porous and unstable. Physical phenomena also become carriers of anomalies, and pure calculation – as in the case of harmonious non-proportional harmonies – that reveal the limits and critical issues that have yet to be overcome. According to recent studies, musical material is increasingly more difficult to grasp, producing confusing logical-rational matrices.

The structuralist musical tradition, that has contributed much to research, was essentially founded on the concept of resonance. But, this resonance is nowadays no longer considered as a mere process of propagation of waves, rather it is considered as a fusion between the mind and the body. At the epistemological level reason supposes a clear distinction between subject and object, however the new notion of resonance is characterised by fusion. If the rational reason requires the separation of the two in the musical discourse, then resonance functions by attraction and interpenetration. This occurs as a result of mutual vibration, without purely mechanical processes. As a consequence, musical phenomena would be active at almost particle and viral levels, leaving trace of their transit. Some scholars have attempted to marry this notion with Derrida's concepts of *trace* and *hantologie*. That is, imaging sounds as a continuous presence, that is not limited in space and time, but is mirrored in contemporary sound landscape in which listening is certainly mostly involuntary.

The contemporary musical individual, following Derrida, is constantly in search of traces. The musical past and present live alongside one another, in virtue of the many technological means that we have at our disposal, and that we never had in the past. In this way, fruition becomes increasingly stratified and inhabited by ghosts and spectres of other epochs or

[4] Lisciani-Petrini (2007: 92).

geographical areas. *Hantologie*, as explained by Derrida in the field of political ideas, finds implications in the field of music. In fact, re-appropriation and re-elaboration are seen as processes that composers often use. These processed can take place at a conscious or unconscious level, because the circulation of sensory impulses, and their part in general *hantologie*, in any case produces these same effects:

> As an identifiable, bordered, posed subject, the one who writes and his or her writing never give anything without calculating, consciously or unconsciously, its re-appropriation, its exchange, or its circular return – and by definition this means reappropriation with surplus values, a certain capitalization. We will even venture to say that this is the very definition of *subject as such*. One cannot discern the subject except at the subject of this operation of capital.[5]

The crisis of the subject in music appears amplified – if not provoked – by the continuous mixing of subject and object, composition and composer, composers and listeners, the market and technological devices. In essence, we could say that it is a mixing of performance and fruition. As explained by David Toop, the embracing sensory world, that partially questions the rationalist paradigm, is connected to archaic and pre-psychological processes:

> All of us, or should I say those of us equipped from the beginning with the faculty of hearing, begin as eavesdroppers in darkness, hearing muffled sounds from an external world into which we have yet to be born. […] four and half months after conception we begin to hear – This is the first of our senses to function: hearing dominates amniotic life and yet after birth its importance is overtaken by seeing.[6]

However, sound is also a historically situated phenomenon; therefore is produced and perceived differently, according to its situatedness in time. Sound changes the listener, modifying his/her behaviour. As David Hendy explains, the industrial revolution produced an epoch-making change:

> Before, the rhythms of labour – tilling the field, hauling in fish, spinning cloth – had often been matched to the rhythms of the human body, the rhythms of breathing, of bending, of hands and feet moving, by people singing work-songs. These songs made the work more bearable […]. The demands of efficiency required workers to synchronise themselves with the machine.[7]

---

[5] Derrida (1992: 101).
[6] Toop (2010: IX).
[7] Hendy (2013: 221).

The new century is in fact ushered in by machines and the noise they make. This paradigm shift impacts significantly on music and its listeners. Machines start to impose their very own sensory existence, it appearing increasingly independent. It becomes increasing difficult to decide what to listen to, because most of the sounds that are produced by machines often do not require the presence of a human. Electricity, modes of transport and heavy industry produce a gamma of new sounds that instigate a crisis for *homo audiens*.

This crisis is augmented by the fact that this sound acquires aestheticism and acquires pleasant sounds. At this point in time, the pre-Helmhotz paradigm, and that of his inflexible followers, begins to crumble, because - from their point of view – it possible to retrieve pleasant sounds from acoustically non-consonant harmonies. Further, their point of view, questions whether these harmonies can be considered music.

Cristopher Cox and Daniel Warner identify three elements that contributed to this sound-noise paradigm shift:

> The reasons for this are many. Already at the turn of the nineteenth century, the music of Debussy, Schönberg, and Stravinsky challenged tonality on a number of fronts. Not longer after, Cowell, Varèse, and Cage began to explore non-pitched sounds. Ethnomusicological research into the nature of music outside of Europe began to suggest a need to expand the concept of music beyond the narrow and specialized domain it demarcated in the West.[8]

Therefore, the concerted stances of composers, researchers and technology gave way to a process of erosion that lasted a century and from which emerged a by-product of nature: noise. Even though there is much to say on noise in music, and much has been said, in this article it is analysed as a destabilising element in musical experience. This is because noise, in addition to it being an expression of creativity for composers, "obliges" listeners to reflect. In fact, noise being logical, ordered and rational, counters and undermines the entire musical paradigm. According to the composer Morton Feldman:

> Sound is all our dreams of music. Noise is music's dream of us. And those moments when one loses control, and sound like crystals forms its own planes, and with a thrust, there is no sound, no tone, no sentiment, nothing left but the significance of our first breath […].[9]

[8] Cox & Warner (2004: 5).
[9] Varèse (1966: 17).

Edgar Varèse is undoubtedly one of the musicians that systematically worked at emancipating of noise in music. His view that a composer is the organiser of sounds is a significant break through in resolving the music/noise juxtaposition. By placing emphasis on the act of organisation the principal focus on the action is shifted. It is no longer on the content that can be represented by any type of sensory impulse, irrespective of the harmonious nature of the sound. Yet, if the creativity of the individual artist can easily accept the use of any sound, inserting it in a rationally organised formal context, the same may not apply to the public. The reception of composition practices that are on the edge of tradition, even when they are organised in a formally correct manner, has always been problematic. The lack of familiarity with timbre has always raised issues relating to whether the use of noise is legitimate or not. As Varèse comments:

> Although this new music is being gradually accepted, there are still people who, while admitting that it is "interesting", say: "but is it music?"It is a question I am only too familiar with. Until quite recently I used to hear it so often in regard to my own works that, as far back as the twenties, I decided to call my music "organized sound" and myself, not a musician, but a "worker in rhythms, frequencies, and intensities". Indeed, to stubbornly conditioned ears, anything new in music has always been called noise. But after all, what is music but organized noises? And a composer, like all artists, is an organizer of disparate elements. Subjectively, *noise* is any sound one doesn't like. Our new medium has brought to composers almost endless possibilities of expression, and opened up for them the whole mysterious world of sound.[10]

The theoretical assumption that which is not measureable or logically classifiable is not music was also debated by other composers in the previous century. The concepts of "natural" or "regular" sounds are increasingly historicised and ascribed to the world of physical acoustics, rather than that of aesthetics.

The union between the laws of physics and aesthetics are at the heart of rationalist conceptions of music. In the *Joys of Noise*, Henry Cowell affirms:

> Music and noise, according to a time-honored axiom, are opposites. If a reviewer writes "It is not music, but noise", he feels that all necessary comment has been made. Within recent times it has been discovered that the geometrical axioms of Euclid could not be taken for granted, and the explorations outside them have given us non-Euclideian geometry and Einstein's physically demonstrable theories. Might not a closer scrutiny of musical axioms break down some of the hard-and-fast notions still current in musical theory, and build up a non-Bachina

[10] Varèse (1966: 19).

> counterpoint, a non-Beethovenian harmony, or even a non-Debussian atmosphere, and a non-Schoenbergian atonality? […].[11]

A galaxy of possible universes would in actual fact become possible thanks to the extensions of creative paradigms. This would open up new potential realities that would enrich – and at times overturn – musical tradition.

> As musical sound grows louder, the noise in it is accentuated and the tone element reduced. Thus a loud sound is literally noisier than a soft one; yet music does not touch our emotional depths if it does not rise to a dynamic climax. Under the best circumstances, the emotions are aroused by musical noise and lulled musical tone. Since the "disease" of noise permeates all music, the only hopeful course is to consider that the noise-germ, like the bacteria of cheese, is a good microbe, which may provide previously hidden delights to the listener, instead of producing musical oblivion. Although existing in all music, the noise element has been to music as sex to humanity, essential to its existence, but impolite to mention, something to becloaked by ignorance and silence[12].

Noise as an anomaly, as an *agent provocateur*, can surely re-balance musical narrative that accentuates the difference between rational listening and purely emotive listening. If, as we have seen, these two aspects are inextricably bound, then insisting on how to divide the two could possibly be a way to resolve the issue.

Musical experience is both rigorous calculation and vortex, confusion and a well-defined path. Every listener is a machine that can feel emotion. By accepting this way of seeing things, music can be conceived as something that is constructed by the listener. Music possibly harvests its most precious fruits by alternating rationality with its exact opposite.

[11] Cowell (1929: 287).
[12] Cowell (1929: 288).

References

Cowell, H. (1929) "The Joys of Noise", in *New Republic No. 59* (31 July 1929), pp. 287-288.

Cox, C., & Warner, D., (eds) (2004), *Audio Culture. Readings in Modern Music*, New York/London, Continuum Books.

Derrida, J. (1992), *Given time: Counterfeit Money*, Chicago, The University of Chicago press.

Helmholtz von, E. (1896), *Die Lehre von den Tonemfindungen als physiologische Grundlage für die Theorie der Musik*, Braunschweig, Vieweg.

Hendy, D. (2013), *Noise. A Human History of Sound and Listening*, New York, HarperCollins.

Lisciani-Petrini, E. (2007) *Risonanza. Ascolto, corpo, mondo*, Milano, Mimesis.

Serravezza, A. (1996), *Musica e scienza nell'età del Positivismo*, Bologna, Il Mulino.

Toop, D. (2010), *Sinister Resonance. The Mediumship of the Listener*, NewYork/London, Continuum Books.

Varèse, E. (1966), "The Liberation of Sound", in *Perspectives of New Music*, Vol. 5, No. 1 (Autumn - Winter, 1966), pp. 11-19.

This volume is the ideal and concrete sequel of another special issue of "Isonomia": *Animali razionali. Studi sui confini e sulle possibilità della razionalità*, edited by Pierluigi Graziani, Giorgio Grimaldi and Massimo Sangoi. The aim – then and today – is to overcome the barriers among different ways of thinking the status, the role, the function and the tools of philosophy. The theme of this issue is – once again – Reason, one of the most important things we have in common, but that is at the same time so enigmatic, complicated and problematic to think about. The authors of these seven essays reflect on some of Reason's fields of application, that is, history of thought, economics, artificial intelligence, theory of argumentation, architecture, neuroscience and aesthetics.

ISSN 2037-4348

This volume is the ideal and concrete sequel of another special issue of "Isonomia": *Animali razionali. Studi sui confini e sulle possibilità della razionalità*, edited by Pierluigi Graziani, Giorgio Grimaldi and Massimo Sangoi. The aim – then and today – is to overcome the barriers among different ways of thinking the status, the role, the function and the tools of philosophy. The theme of this issue is – once again – Reason, one of the most important things we have in common, but that is at the same time so enigmatic, complicated and problematic to think about. The authors of these seven essays reflect on some of Reason's fields of application, that is, history of thought, economics, artificial intelligence, theory of argumentation, architecture, neuroscience and aesthetics.

ISSN 2037-4348

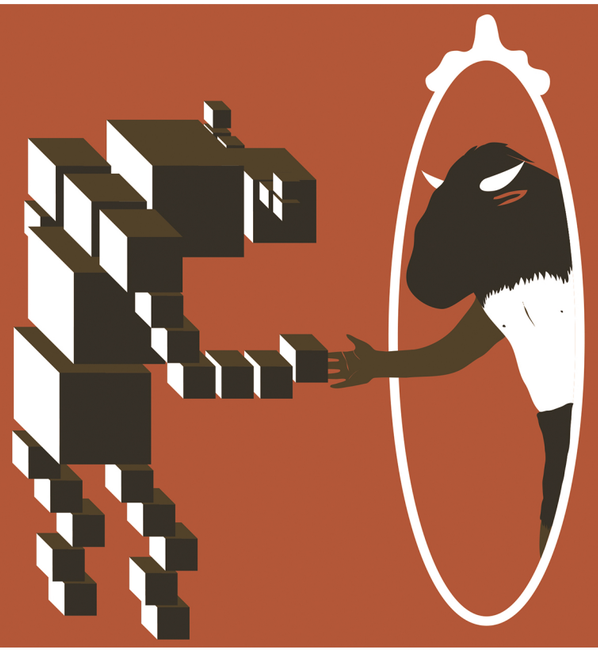
Giorgio Grimaldi, Marialuisa Parise, RATIONAL ANIMALS

Rational Animals

SEVEN ESSAYS ON REASON'S FIELDS OF APPLICATION

edited by

Giorgio Grimaldi
Marialuisa Parise



Isonomia *Epistemologica*