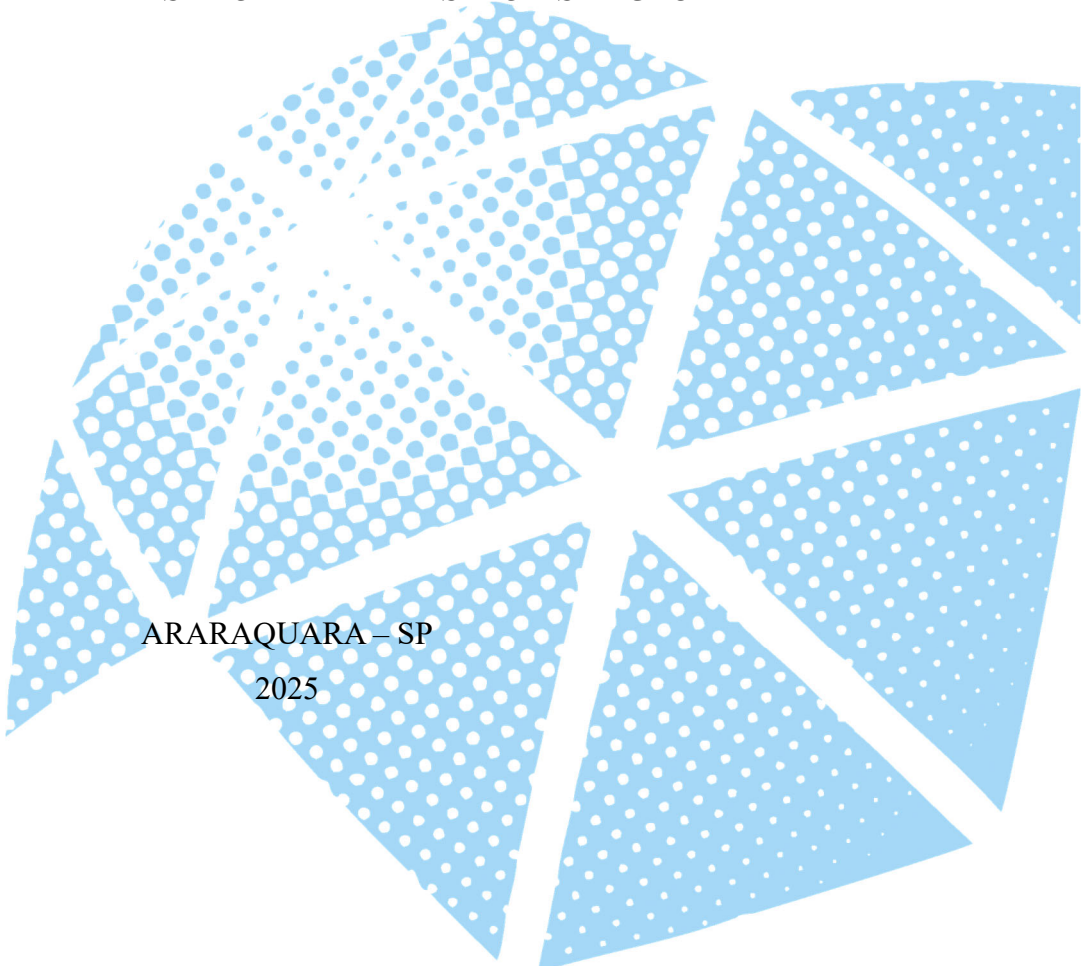


UNIVERSIDADE ESTADUAL PAULISTA – UNESP
Faculdade de Ciências e Letras – Campus de Araraquara

BRUNO CUNHA MARCHETTI

AFTER ALL, WHAT IS BOUNDED RATIONALITY?
A CRITICAL APPRAISAL OF HERBERT SIMON'S LEGACY

ARARAQUARA – SP
2025

An abstract geometric design in the bottom right corner of the page. It features a large, irregular shape composed of several triangular and quadrilateral sections. These sections are filled with a light blue color and a pattern of small white dots. The sections are separated by white lines, creating a complex, crystalline or architectural appearance. The design is partially cut off by the right edge of the page.

BRUNO CUNHA MARCHETTI

AFTER ALL, WHAT IS BOUNDED RATIONALITY?
A CRITICAL APPRAISAL OF HERBERT SIMON'S LEGACY

Tese apresentada à Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Letras, Araraquara, para obtenção do título de Mestre em Economia.

Área de Concentração: Economia

Orientador(a): Prof. Dr. Sebastião Neto Ribeiro Guedes

Araraquara – SP

2025

M317a

Marchetti, Bruno Cunha

AFTER ALL, WHAT IS BOUNDED RATIONALITY? : A
CRITICAL APPRAISAL OF HERBERT SIMON'S LEGACY /

Bruno Cunha Marchetti. -- Araraquara, 2025

156 p.

Dissertação (mestrado) - Universidade Estadual Paulista (UNESP),
Faculdade de Ciências e Letras, Araraquara

Orientador: Sebastião Neto Ribeiro Guedes

1. Racionalidade Limitada. 2. Herbert Simon. 3. Metodologia. 4.
História do Pensamento Econômico. 5. Economia Comportamental. I.
Título.

IMPACTO POTENCIAL DESSA PESQUISA

Essa pesquisa busca analisar aprofundadamente o conceito de “Racionalidade Limitada”, clarificando como seu criador, Herbert Simon, usou o conceito. A pesquisa desagrega as diferentes interpretações do conceito e apresenta dados bibliográficos sobre seu uso, clarificando como a pesquisa em Economia incorpora a Racionalidade Limitada.

POTENTIAL IMPACT OF THIS RESEARCH

This research seeks to conduct an in-depth analysis of the concept of “Bounded Rationality,” clarifying how its creator, Herbert Simon, employed the concept. The study unpacks the different interpretations of the term and presents bibliographic data on its use, shedding light on how research in Economics incorporates Bounded Rationality.

BRUNO CUNHA MARCHETTI

**AFTER ALL, WHAT IS BOUNDED RATIONALITY?
A CRITICAL APPRAISAL OF HERBERT SIMON'S LEGACY**

Tese apresentada à Universidade Estadual Paulista (UNESP), Faculdade de Ciências e Letras, Araraquara, para obtenção do título de Mestre em Economia.

Área de Concentração: Economia.

Data da defesa: 24/01/2024

Banca Examinadora:

Prof. Dr. Sebastião Neto Ribeiro Guedes
UNESP – Faculdade de Ciências e Letras – Campus de Araraquara

Prof. Dr. Ana Maria Afonso Ferreira Bianchi (titular)
USP – Universidade de São Paulo – Campus de São Paulo

Prof. Dr. Mário Augusto Bertella (titular)
UNESP – Faculdade de Ciências e Letras – Campus de Araraquara

Prof. Dr. Tatiana Massaroli de Mello
UNESP – Faculdade de Ciências e Letras – Campus de Araraquara

Prof. Dr. José Ricardo Fucidji (suplente)
UNICAMP – Universidade Estadual de Campinas – Campus de Campinas

To my professors.

ACKNOWLEDGMENTS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) - Código de Financiamento 001 (The present work was made with the support of the *Coordenação de Aperfeiçoamento de Pessoal de Nível Superior* – Brazil (CAPES)).

I would like to start by expressing my gratitude to the public Brazilian higher education system, specially São Paulo State University (*Universidade Paulista “Júlio de Mesquita Filho”* or Unesp, for the intimates), in which I graduated and developed this dissertation. I also have to thank CAPES (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*), which was responsible for the essential financial support that made this dissertation possible.

Even if I tried to be thorough, I must apologize for any important names that were left out of the acknowledgments.

I sincerely thank my supervisor, Sebastião Neto Ribeiro Guedes, for the continuing support since my graduation and his great patience. Thank you for always “buying” my ideas and helping me to make them work.

I thank Professor Tatiana Massaroli de Mello, for agreeing to be part of the qualification board of this dissertation. I would also like to thank Professor Tatiana for the constant support and for the many opportunities to assist her in class.

Likewise, I thank Professor José Ricardo Fucidji for also accepting to be in the qualification board. I would specially like to thank Professor Fucidji for the recommendations of the book *Machines Dreams*, by Phillip Mirowski, an essential author for this research and of whom I became a fan of.

I thank Professor Mário Augusto Bertella, for accepting to be a part of the defense board of this dissertation.

Furthermore, I thank professor Ana Maria Bianchi, who also accepted to be part of my dissertation’s defense. Thank you for the solicitude and availability.

I thank Professor Rogério Gomes for the continuing interest in my academic career. I am deeply grateful for the many orientations and for putting me “in the right path”.

I also have to thank Mariana for the great productivity boost.

I thank my friends for the continuing support. Running the risk of leaving important people out, I thank my colleagues from Unesp, Victor Hugo Castro, Caio Santarsiere, Alessandro Máximo, Luís Felipe Koga, Júlia Ruiz, Laura Paraíso and Felipe Martins; and my Institutional friends Rafael Duregger, Rafael Camatta, Luiza Fernanda, Eduardo Martins, Lucas Marçal, André de Jesus, and Fábio Fernandez.

I have to thank my dear friend, Theodoro Sposito for the great help with the bibliometric section of this dissertation. Thank you for showing me the ropes.

I also want to thank Dorricielk Gomes for the insightful comments. Thank you for encouraging me to be bold.

To my family, specially my mother, Claudia Marchetti, I thank you for the continuing support and faith in me. I also would like to thank my mother-in-law, Jane Neves, for the same reasons.

Finally, I thank my beloved companion, Alicia. This world would be gloomy without you in it.

“—You talk about redefining my identity.

I want a guarantee that I can still be myself.

—There isn't one. Why would you wish to?

All things change in a dynamic environment.

Your effort to remain what you are is what limits you.”

Ghost in The Shell (1996)

RESUMO

Racionalidade Limitada é uma das ideias mais celebradas na história da economia. Sua aparente simplicidade e ampla adaptabilidade, no entanto, ofuscam tanto um complexo desenvolvimento teórico quanto as circunstâncias que engendraram. Esta dissertação se insere na discussão sobre a Racionalidade Limitada em duas frentes: 1) identificar as origens conceituais da Racionalidade Limitada em Herbert Simon e 2) destacar os caminhos de desenvolvimento dessa teoria comportamental. Um argumento central permeia esta dissertação: o conceito de Racionalidade Limitada desenvolvido por Simon serviu como base para desenvolvimentos futuros; contudo, nenhum pesquisador seguiu sua metodologia específica. Assim, o impacto de Herbert Simon e do conceito de “Racionalidade Limitada” divergem, mesmo que esse fato seja pouco sublinhado. Dessa forma, o “legado de Herbert Simon” é, no melhor dos casos, ambíguo. Além de uma ampla revisão da literatura e da ciência de Herbert Simon (Capítulo I), esta pesquisa também investigou as interpretações tradicionais de Racionalidade Limitada nas várias vertentes da economia comportamental (Capítulo II). Ademais, esta dissertação também engendrou uma análise bibliométrica de dados bibliográficos (aproximadamente) da última década (Capítulo III). Os objetivos desta pesquisa são: 1) identificar como Simon impactou a economia, 2) como esse impacto evoluiu após ele, e 3) traçar um panorama da evolução da Racionalidade Limitada que está ocorrendo. Esperamos que este esforço seja frutífero para a discussão sobre a Racionalidade Limitada.

Palavras-Chave: Racionalidade Limitada; Herbert Simon; Metodologia Econômica; História do Pensamento Econômico; Comportamento Econômico

ABSTRACT

Bounded Rationality is one of the most celebrated ideas in economic history. Its simple demeanor and adaptability, however, obfuscate a complex development and the circumstances that generated it. This dissertation inserts itself in the discussion of Bounded Rationality, on two main fronts: 1) to identify the conceptual origins of Bounded Rationality in Herbert Simon, and 2) to highlight the paths of development of this behavioral theory. A main argument permeates this whole dissertation: that the concept of Bounded Rationality developed in Simon's work served as a basis for future developments; however, no researcher followed his specific methodological steps. Therefore, the impact of Herbert Simon and of "Bounded Rationality" differ, even if this fact is seldom noticed. As such, the "legacy of Herbert Simon" is ambiguous at best. Besides an ample review of the literature and science of Herbert Simon (Chapter I), this research also investigated the traditional interpretations of Simon's influence in the form of the many strands of behavioral economics (Chapter II). Furthermore, this dissertation also employs a bibliometric analysis of bibliographical data of (roughly) the last decade (Chapter III). The objectives of this research are: 1) to identify how Simon impacted economics, 2) how this impact evolved after him, and 3) to paint a broad picture of the evolution of Bounded Rationality that is happening right now. This effort, we hope, is fruitful to the discussion of Bounded Rationality.

Keywords: Bounded Rationality; Herbert Simon; Economic Methodology; History of Economic Thought; Economic Behavior

LIST OF FIGURES

Figure I – Thematic Map of Sample 1	106
Figure II – Co-citation Network of Authors of Sample 1	108
Figure III – Co-citation Network of References of Sample 1	110
Figure IV – Thematic Map of Sample 2	120
Figure V – Co-citation Network of Authors of Sample 2	121
Figure VI – Co-citation Network of References of Sample 2.....	123

LIST OF TABLES

Table I – Structure of Bounded Rationality	38
Table II – The Hardcore of Hebert Simon.....	61
Table III – Classification of Approaches of Bounded Rationality	81
Table IV – Classification of Behavioral Economics Approaches	92
Table V – Leading Publishers in Sample 1 (2013 – 2023).....	104
Table VI – Clusters in the Co-citation of References in Sample 1	111
Table VIII – Clusters in the Co-citation of References in Sample 1.....	124

LIST OF ABBREVIATIONS

AI	Artificial Intelligence
ABM	Agent-Based Model
BE	Behavioral Economics

SUMMARY

1. INTRODUCTION – The Mazes of Bounded Rationality	16
CHAPTER I	21
2. Herbert Simon on the creation of Bounded Rationality: Decision-making, natural and artificial. 21	
I.1 – Artificiality and Intelligence	22
I.2 – Behavior and Bounded Rationality.....	28
I.3 – The Methodology of Simon’s Artificial Science	37
I.4 – Economic Aspects of Herbert Simon and the Critique of Substantive Rationality	43
I.5 – Herbert Simon in the Orthodoxy-Heterodoxy debate.....	48
I.6 – The Hardcore of Simon’s Bounded Rationality.....	57
CONCLUSION OF CHAPTER I	61
CHAPTER II.....	62
3. Bounded Rationality After Herbert Simon: Three interpretations on the limits of cognition and what unifies them.....	62
II.1 – How to Reach “Bounded Rationality”?.....	64
II.2 – Two Periods of Behavioral Economics	68
II.3 – Approaches of Behavioral Economics.	75
II.4 – Types of Bounded Rationality	80
II.5 – Bounded Rationality as an Axiom.....	90
CONCLUSION OF CHAPTER II.....	93
CHAPTER III	95
4. The Legacy of Herbert Simon in Current-day Economics: Bibliographical evidences of the nature and uses of Bounded Rationality.....	95
III.1 – Data Collection and Bibliometric Methodology.....	96
III.2 – Data analysis of Sample 1	100
III.3 – Data visualization of Sample 1	104
III.4 – Summary of Sample 1	110
III.5 – Data analysis of Sample 2.....	112
III.6 – Data visualization of Sample 2	117
III.7 – Summary of Sample 2.....	122
III.8 – Interpretation of the results	123
CONCLUSION OF CHAPTER III.....	124
5. CONCLUSION – A Map to the Maze?	126
REFERENCES	132
APPENDIX I – Details of the Bibliographical Model.....	144

APPENDIX II – R script used to process and present the collected data.	147
APPENDIX III – Interpretation of the Thematic Map.....	148
APPENDIX IV – Complete list of Co-citations.....	149

1. INTRODUCTION – The Mazes of Bounded Rationality

I

Hebert Simon was an explorer; however, not a traditional one. He did not brave the frontiers yet to be conquered, instead, Simon dedicated his work and life in understanding another kind of frontier – the limits of the mind. An unfettered imagination, along with relentless curiosity were, the instruments used to boldly explore his chosen subject. His incursions and discoveries, however, were numerous, and it became impossible to write a single narrative that unified his trajectory. “I have been a scientist, but in many sciences. I have explored mazes, but they do not connect into a single maze” (Simon, 1996 [1991], p. XVII).

Simon also “mapped” his explorations. Knowing the secrets of the mind was not enough to him, he also sought to classify, organize, and systematize the acquired knowledge. The chosen tool was, naturally, the “behavioral model”, a familiar and favored intellectual tool of the author. His maps were as numerous as his many incursions, and also did not generate a single narrative, for he did not seek one – “My aspirations do not extend to achieving a single consistency in my life” (*ibid*).

Simon was, thus, a man of *models* – not a man of *a single model*. Models were the instruments that Simon used to make sense of the world, whether it was the mind itself¹, human society², artificiality³, scientific inquiry⁴, or his own life⁵. This unique type of methodology is the foundation on which his many contributions are built. What unifies his methodology and scientific style were not the models themselves (which vary greatly from one another), but instead his specific approaches to “doing science”. What ties these contributions to the man

¹ Here we refer to “*A Behavioral Model of Rational Choice*” (1957b), “*Theories of Bounded Rationality*” (1964), “*Models of Bounded Rationality*” (1997 [1982]), among others.

² “*Models of Man: social and rational*” (1957a) “*Invariants of human behavior*” (1990a), *Reason in human affairs*” (1990b), among others.

³ Mainly, “*The Sciences of the Artificial*” (1969).

⁴ Mainly, “*Models of Discovery*” (2012 [1977]).

⁵ Here we refer to his autobiography, “*Models of My Life*” (1996 [1991]).

who authored them is a (somewhat obsessive) interest in how behavior occurs. The lack of cohesion, consequently, is a result of the many themes that he tackled and the many models developed to understand them – not a lack of methodological consistency.

This inconsistency characterizes the entirety of Simon’s work, a fact he himself acknowledged: “Which of the wanderers through these different mazes will step forward at the call for the real Herbert Simon? All of them; for the “real” self is an illusion.” (Simon, 1996 [1991], p. XVI). Unlike other authors, who tried to ‘guide’ the interpretations of their work, Simon fully embraced “the death of the author”⁶ and rarely criticized others’ interpretations of his work.

As such, there is no single interpretation of Simon’s writings, a fact seldom contested by the author. Consequently, the collective interpretation of Herbert Simon, in a sense, *became* his contribution. His central idea of “Bounded Rationality” (of special interest to this research), therefore, is not and could not be one specified conception. Instead, this type of rationality is an amalgamation of different ideas, applications, mathematical formulas, behavioral presuppositions, and so forth.

II

The ideas of Herbert Simon were broad and proved useful to all persuasions. The insight of “bounding” rationality was immediately evocative for the entirety of economic science: economists with orthodox leanings saw in this insight an advancement of the increasingly inadequate model of ‘perfect rationality’, while heterodox economists saw in Bounded Rationality a powerful critical tool to escape the confines of optimizing behavior. As Kahneman notes (2017 [2011]): “Herbert Simon [is] perhaps the only scholar who is recognized and

⁶ “Death of the author” is a concept developed by Roland Barthes in literature studies, which can be defined as the recognition that “The author’s intention and biography are no longer relevant to the interpretation of the text”. The main consequence is “that once a text is published, it takes on a life of its own and becomes open to interpretation by readers”. Source: www.oxfordhomeschooling.co.uk/blog/the-death-of-the-author/.

admired as a hero and founding figure by all the competing clans and tribes in the study of decision-making” (p. 219, adapted).

The broad adoption of the ‘Simonian paradigm’ by economic science, much like the many “Simons” who explored the many mazes, is not, however, a unified whole. In the case of Bounded Rationality, for instance, certain elements were better suited to specific types of economic analysis, while others were not. This resulted in a generalized ‘partial absorption’ of Simon’s concepts, which evolved into various interpretations of Bounded Rationality that can be observed in economic literature. This unfolding of events, however, imposes a peculiar consequence on the interpretation of Simon’s work: no ‘true’ intellectual descendant of Simon can be conclusively identified. There are many that adopt or criticize concepts, ideas, and models of Simon, but it is incredibly difficult to identify an ‘heir’ to his specific style of scientific development.

As attempting to extract the ‘real’ Herbert Simon has proven to be a futile task, we are left with the impression that Simon saw *his* interpretation of *his* work as just one among many – not a greater truth. In this sense, Simon was a practicing ‘Lakatosian’⁷ who viewed his many models as part of research programs – not *the* research program. This fluid notion of “science” helps to explain his inclination in creating *theories* and *models* instead of striving toward a unified goal, shaping how these ideas were adopted by economic sciences.

III

To evaluate Herbert Simon’s legacy, therefore, presents a series of unique challenges. This research aims to address the idiosyncratic elements of Simon’s vast body of work in an attempt to classify it within the economic sciences of his time and to identify what were Simon’s original contribution. This forms the basis to identify *how* Simon’s contributions impacted economic sciences and how they evolved through the years. Much is attributed to Simon, but not everything that economists attributed to him can be found in his work. This dissertation,

⁷ Here we refer only to similarities between Lakatos’ theory of scientific paradigms and Simon’s general attitude towards science, as mentions of Lakatos are scarce in Simon’s work.

therefore, endeavors to “give Caesar's what is Caesar's”, pinpointing which ideas are Simon’s and which are only attributed to him.

This research begins as an attempt to navigate the cacophony surrounding the ideas of the “prophet of Bounded Rationality” (Simon, 1996 [1991], p. 250; Sent, 1997, p. 323 *apud* Barros, 2007). His ideas, however, have taken on a life of their own and, in a sense, transcended their author. Simon is still recognized as one of the founding fathers of Bounded Rationality and Behavioral Economics, however, recent evidence suggests that his work, while frequently cited, is not as deeply incorporated into economics as one might expect. The main sources of “Bounded Rationality” nowadays are other authors, who used the name, but changed the concept.

What we now call “Bounded Rationality”, therefore, is not one concept, but rather a collection of different formulations. It retains the name and acknowledges its creator, but radically alters Simon’s methodology. Over the years, key elements of Simon’s work changed, a silent evolution that only recently began to be understood. Far from advocating for a ‘purist’ view, or calling for a return to the ‘true’ Herbert Simon, this research seeks to identify three central elements in the evolution of Bounded Rationality: 1) how Simon originally conceptualized “Bounded Rationality”, 2) how this concept was adopted by economic science, and 3) how Bounded Rationality has evolved.

Therefore, this research does not present a generalized view of the impacts of Simon’s work as a whole, focusing “only” in Bounded Rationality. As such, the general understanding of Simon’s science is essential to understand how he sculpted the concept of Bounded Rationality, which justifies the lengthy review of Simon’s work. While extensive, however, this review is not exhaustive, and it is important to recognize that this dissertation cannot comment on the complete works of Simon. Other concepts of Simon’s work, thus, will be included, but only if they relate to Bounded Rationality in some relevant way.

IV

This research is organized as follows: The first chapter is a review of Simon's original texts, in an attempt to identify the "hardcore"⁸ of the "Simonian paradigm". The second Chapter proposes an interpretation of *what is* Bounded Rationality in economic sciences, focusing on what unifies the many interpretations of the concept – its axiomatic use. The third and final chapter will present bibliographical evidence of the reception of Simon's work into contemporary economic science, demonstrating the significant distance between researcher who *cite* the work of Herbert Simon and those who *use* the work of Herbert Simon. Specific methodological considerations will be addressed in each of the respective chapters.

In the spirit of Simon's attitude toward his ideas, this research seeks not to provide a definitive answer, but rather to explore the many answers to the question of "what is Bounded Rationality". No one answer suffices, but the analysis of the conceptual evolution illuminates a great deal. To keep the "many mazes" analogy, this research aims to be a "map for maps" of the expeditions of Herbert Simon into the maze of the mind, in a first stage. In a second moment, we also analyze how others read those "maps", and what they concluded. In a final moment, we venture to sketch a map of our own. This effort, we hope, may help to illuminate the complex geography of Bounded Rationality.

⁸ The term here is inspired by Lakatos (1968).

CHAPTER I

2. Herbert Simon on the creation of Bounded Rationality: Decision-making, natural and artificial.

“– Am I a real boy?
– No, Pinocchio [...]”
Pinocchio, 1940

Herbert Simon’s intellectual formation is a complex and interconnected web of influences, ideas, and idiosyncrasies. Even in his early years as a student at University of Chicago University, young Herbert never paid attention to the boundaries between fields of research, “never calling any one home for long” (Mirowski, 2002, p. 454). In Chicago, he studied economics and political science (Simon, 1996 [1991]) and graduated in 1936. In 1943, he earned his doctorate in political science in the same university⁹. During this time, however, Simon was also exploring the fields of psychology and administration¹⁰. Augier (2000) also reports Simon’s early interest in the development of “thinking machines” (computers), a field that would be central in his career. Other notable bibliographical observations include his time as a professor at Carnegie Institute of Technology and his early involvement with the RAND Corporation (Mirowski, 2002), a non-profit research organization closely tied to the U.S. military¹¹.

In this chapter, two central themes are presented: 1) a general interpretation of the way in which Herbert Simon developed his sciences, and 2) how this specific way of science-making

⁹ Source: <https://www.britannica.com/biography/Herbert-A-Simon>

¹⁰ Ibidem.

¹¹ A number of interesting, but ultimately pointless, bibliographical details could not be included in this brief overview. Some highlights are the case of Simon buying his house, which took one single day in total (Augier, 2000, p. 424) and the fact that Simon always ordered the same lunch (American cheese sandwich in white bread) (Modigliani, 2001, p. 87).

generated the concept of ‘Bounded Rationality’. Later, the effects of these elements on economics will be analyzed. The “leitmotifs” of Simon are of particular interest, such as his interests and methods in the study of the mind, his intellectual lineage and the *zeitgeist* he was part of. In a more philosophical note, we will explore Simon’s ideas on the limits of the “natural” and the “artificial”. Far from an idle curiosity, this neglected element of Simon’s scientific framework sheds light on several questions and connects many of the disparate elements of his work. The main argument is the following: the specific ways in which Simon interpreted the barrier between the artificial and the natural is the basis of his behavioral theory. This specific constellation of elements shaped his influence in economics and composes the ‘hardcore’ of Simon’s science.

This chapter is organized into six sections: Section I.1 deals with the role of the ‘artificial’ in Simon’s work. Section I.2 presents his ‘monomania’ about decision-making, identifying the type of behavior Simon focused on. The third section identifies the methodological elements of Simon’s science, highlighting his ‘special type of empiricism’. The fourth section presents the structure of Simon’s critique of Substantive Rationality¹² and its ripples in economics. The fifth is an attempt to locate Simon’s theories within the economic sciences, aiming to determine where Simon lands on the “orthodoxy-heterodoxy” spectrum. The final Section is both a summary and the presentation of the main argument for this first chapter – the identification of the *hardcore* of Simon’s methodology. This characterization forms the basis of the analysis in Chapter 2.

I.1 – Artificiality and Intelligence

The concept of ‘artificiality’ tends to take second stage in the discussion around Herbert Simon. Mostly remembered as the “prophet of Bounded Rationality” (Simon, 1976 [1947], p. 250; Sent, 1997, p. 323 *apud* Barros, 2007), his contributions to “artificial fields” (such as computer science, digital simulation, and artificial intelligence) are not forgotten, but their impacts pale in comparison with Simon’s models of rationality. Therefore, the relationship

¹² Also refereed as “perfect rationality” or “unbounded rationality”.

between the ‘artificial’ parts of Simon’s work and the other facets of his work is not well understood, despite the ample body of work¹³.

Moreover, the connection between ‘artificiality’ and Bounded Rationality is vague at best. In his life, Simon made little effort to systematize his specific methodology of research. As such, the scientific community had access to Simon’s final products (the models themselves), but little detail in the intermediaries that generate them. The consequences are summarized by Fiori (2011): “Simon’s original perspective, which connects [Bounded Rationality] and Artificial Intelligence (AI), has been essentially abandoned, and new theoretical tools have replaced it” (p. 587, adapted). Even so, this connection constitutes one of the central pillars in understanding the behavioral and cognitive theories of Herbert Simon. ‘The artificial’, therefore, is fundamental to understanding his science. This connection, however, is not trivial to parse out, as the pieces of this puzzle are scattered. The objective of this topic is to present why the concept of ‘artificiality’ is important to understand Herbert Simon, and on which basis we make this claim.

Let us start by defining what is artificial to Simon, as his use diverges from the common sense. For Simon, ‘natural’ and ‘artificial’ describe *only* the origin of something – without describing any other characteristic of it. The author took many pains to avoid the negative connotations that are usually associated with the concept of artificiality. He is emphatic: “[...] you will have to understand me as using “artificial” in as neutral a sense as possible, as meaning man-made as opposed to natural” (Simon, 1969, p. 4). With this definition, Simon avoids two usual interpretations. The first is ‘artificial’ as the opposite of “spontaneous” (e.g., “artificial” smile, as opposed to a “natural” one); the second is ‘artificial’ as the negation of nature (e.g., “artificial” food like processed meat, as opposed to “natural” food like fruits or vegetables). Both of these interpretations paint the artificiality in a negative light, as if it is something to be avoided in favor of the natural. Simon, however, completely rejects this notion: an organic vegetable is as artificial as processed foods, for it also depends on human labor; the vegetable was planted by human hands, harvested, and cooked by humanity – it is man-made. By all

¹³ The most notable example is the book *The Sciences of the Artificial*, which compiles a host of Simon’s work in sciences that he considered “artificial”.

means, the organic vegetable is artificial – for Simon, organic vegetables are as artificial as any kind of processed food.

While seemingly trivial, this definition carries with it a series of implications, mainly that the broad abstract notions of “nature” lose their meaning. For instance, “natural human behavior”, often an important element in the rhetoric of social sciences (especially economics), is deprived of any meaningful conceptual use – no product of human behavior *may be* natural (be it the price level, the unemployment rate, or the propensity for barter); it can only be artificial (generated by humanity). Human behavior that happens in society may *only* be artificial, for all forms of society are undoubtedly the result of human action and, therefore, artificial. As such, Simon does not believe in a “natural” form of human society.

The key point here is to understand that “artificial behavior” does not signify a shift away from “desired behavior” – it is not distorted. This is true when dealing with societies but also applies to the environment: “The world we live in today is much more a man-made, or artificial, world than it is a natural world. Almost every element in our environment shows evidence of human artifice” (Simon, 1969, p. 2). To Simon, this is not a judgment call but simply a fact of life.

Another consequence is directly linked to scientific investigation. As artificiality simply refers to the fact that the object is human-made, the division between ‘natural’ and ‘artificial’ sciences is simple: natural sciences study natural phenomena, and artificial sciences study man-made ones. The ‘artificial sciences’, therefore, encompass not only the study of machines (such as computer sciences) but also every field that focuses on human affairs. Consequently, areas such as economics, anthropology, sociology, language, arts, and theology are ‘artificial sciences’ (Simon, 1969). The same can be said about instrumental areas such as engineering (which studies the application of human knowledge) and computer science (which studies man-made machines). Artificial intelligence and digital simulations are merely two among many artificial sciences.

Simon also vehemently attacked the idea that sciences of the artificial consisted of the analysis of “made-up” facts. As the results obtained in artificial fields are not known *ex ante*, artificial sciences, by all means, are empirical sciences (Simon, 1995). The case of digital simulations, such as ABM models like that of Melo, Possas, and Dweck (2016) or the work of Alan Kirman (2010), is especially useful in illustrating this idea. In these simulations, the

interactions between variables are informed by the model, and the initial parameters, both defined by the researcher – but the outcome is not pre-determined. The model cannot predict the outcome – in simulations, properties “emerge” from the interaction of its elements. Here lies the empirical nature of even the most artificial of sciences – the behavior of what is created through human behavior (for instance, computers) is not known. Therefore, it must be tested with the same rigor as any other science. Artificial phenomena, therefore, are “empirical phenomena that can be studied by the methods of observation and experiment common to all science” (Simon, 1964, p. 99).

According to Simon (1995), “The case, in AI, for studying many different kinds of systems empirically is essentially identical to the case, in biology, for studying many species of organisms” (p. 99). Therefore, Simon does not identify any essential difference between the study of artificial and natural phenomenon – expect, of course, their origins. Natural and artificial sciences, therefore, are not in any way opposed; they may interface with one another – and often do. Simon uses the computer as an example. In this case, we observe the interaction between a human-made machine with the natural physical phenomena, such as the behavior of light, the metals, the electricity, and so forth.

Consequently, the line between natural and artificial is blurry. In fact, at the extreme, this line becomes unknowable. How can we put anthropology (which studies, among other things, human behavior mediated by the natural world), aeronautic engineering (which studies man-made machines interacting with the natural world), and theology (which studies human culture’s explanation of the natural world) in the same category and still have some cohesion among these sciences? The answer is simple: we can’t. Every ‘artificial science’ is so fundamentally different from the others that even the “origin” criteria proposed by Simon become insufficient.

The consequences of this blurred line are deep. For instance, how can we know if *we* are artificial or natural? How can we say that human thinking is somehow different from so-called “machine thinking”? The answer is the same as before: we can’t. Consequently, it is impossible to classify ‘human’ and ‘machine’ as either “natural” or “artificial”, for both are partially natural and part artificial beings. Mirowski (2002) synthesizes this consequence as such:

“This subversive project to irreversibly blur the boundaries between machine psychology and human psychology, between the biological and the conceptual, all through the instrumentality of computer simulation is the very essence of the cyborg imperative to erase the boundary between the Natural and the Artificial” (Mirowski, 2002, p. 467).

The use of the term ‘cyborg’ in the quote above is deliberate. According to Mirowski (2002)¹⁴, Simon is one among many ‘cyborg scientists’. ‘Cyborg’ means a “hybrid of machine and organism”¹⁵ (*ibid*, p. 5), and ‘cyborg science’ (originally proposed by Donna Haraway, 2013 [1991]) refers to a specific scientific trend during the Cold War, prompted by the advances in computers. Among the core beliefs of this group is the conviction that research in human affairs would benefit with the application of the methods and language that were being developed in computer sciences.

Unlike the “mechanistic” models inspired by Newtonian physics, such as those of Von Neumann and Morgestern (1944), ‘cyborg scientists’ do not tend to believe in determinism. The “von Neumann automata” (the term is Mirowski’s, 2002, p. 540), for instance, was the victim of a vicious and continuous attack by ‘the cyborgs’. According to them, the automata was not “sufficiently individualistic” (Mirowski, 2002, p. 438) – it has no freedom to choose, only to follow preferences in accord with the parameters defined *ex ante*. The “ceremonial exorcism” of Von Neumann (Mirowski, 2002, p. 438) consisted of the systemic refusal to accept a representative agent that lacked individual desires, ingenuity, or idiosyncrasies. Simon’s effort, along with other cyborg scientists, was to “humanize humanity” by making scientific models more individualized. Ironically, the cyborgs’ attempt to attain this goal by treating humanity as a more “individualized machine” – a computer. Mirowski (2002) summarizes the effects of the advances of “cyborguism” as such:

“With the advent of the cyborg sciences after World War II, something distinctly different begins to happen. Here and there, a cyborg intervention agglomerates a heterogeneous assemblage of humans and machines, the living and the dead, the

¹⁴ I must thank professor José Ricardo Fucidji for pointing out Philip Mirowski’s *Machine Dreams* as a reference for my research. This text proved an indispensable source and shaped many of the ideas presented in this dissertation.

¹⁵ The word ‘cyborg’ was conceived as the agglutination of two other words: “cybernetic” and “organism”. Source: www.etymonline.com/word/cyborg

active and the inert, meaning and symbol, intention and teleology, and before we know it, Nature has taken on board many of the attributes conventionally attributed to Society, just as Humanity has simultaneously been rendered more machine-like” (Mirowski, 2002, p. 13).

The blurred line between human and computer is central to Simon’s thought, which is why he integrates the ranks of the ‘cyborgs’¹⁶. Among the many individual characteristics shared by humans and machines are the so-called ‘cognitive limits’ – the inability to perform a task to its fullest. This observation, that people, as well as machines, attempt to do something but lack the means to do it fully, is the basis on which ‘Bounded Rationality’ will be developed. In earlier works, like *Administrative Behavior* (originally published in 1947), Simon conceived Bounded Rationality mainly as a critique to ‘rational behavior theories’ (such as rational expectations). The specific characteristics of Bounded Rationality, such as the use of heuristics or the step-by-step decision-making of procedural rationality, are only hinted at this point.

The developments of the specificity of Bounded Rationality accompanied Simon’s increasing interest in computer sciences. Therefore, Bounded Rationality was born as a critique of classical economic rationality but developed in the language and conceptual framework of computer sciences. ‘Rational limitations’ are analogous to the functioning of RAM memory; heuristics are task-specific, like computer programs; a (cyber)organism has so much processing power or energy that can be dispensed in a task at any single moment of time¹⁷ and so forth. Even the recurrent use of words like ‘computation’ or ‘memory’, as in “limit upon computation that human short-term memory can hold” (Simon, 1990a, p. 7) or the interpretation of agents as “information processor” (Simon, 1964, p. 2), gives hints on what Simon was thinking when developing his theories. This movement would deeply affect the evolution of the economic sciences, as it not only would criticize the neoclassical paradigm of the time but also lay down a foundation for alternative theories¹⁸.

¹⁶ Mirowski (2002) proposes a “fast and dirty” rule to exclude scientific strands as cyborg: “Bluntly: if it doesn't make fundamental reference to “the computer” (itself a historical chameleon), then it isn't a cyborg science” (Mirowski, 2002, p. 13). As we will see further, Simon makes ample mention of the computer.

¹⁷ Kahneman and Tversky famously would focus on this aspect of Bounded Rationality.

¹⁸ Mirowski (2002) identifies the fundamental critique of neoclassicism in the Cold War stemming from cyborg economists, and not an internal movement in neoclassicism away from deterministic mathematics: “In this

In the interpretation laid down in this section, Simon's concept of 'artificiality' serves to level human and machine cognition: "the family of intelligent systems had been joined by a new genus, intelligent computer programs" (Simon, 1995, p. 95). Simon highlights the "important advances in our understanding of procedural rationality, particularly as a result of research in artificial intelligence and cognitive psychology" (Simon, 1978, p. 1), believing that the paths that social science should take are intimately linked with computer sciences. Therefore, in Simon's mind, no ontological distance exists between artificial and natural beings – the studies of both follow the exact same procedures. This leveling between artificial and biological organisms is the essence of the "cyborg imperative" (the term is Mirowski's, 2002, p. 467), the central belief of cyborg scientists.

This interpretation of artificiality, along with the leveling between organic and artificial cognition, is the first of two elements that permeate the majority of Simon's work and which constitutes a fundamental pillar of his science. The next section presents the other pillar: Simon's interpretation of the process of cognition. These two pillars, we argue, enable the unique character of the science of Simon.

I.2 – Behavior and Bounded Rationality

Simon wasn't only a 'scientist of the artificial', but also a researcher with a specific theme of investigation. According to Augier (2000), Simon studied many different fields, but always with a specific focus: "I have been preoccupied with one topic, decision-making and rationality" (Simon, 1993, p. 395 *apud* Augier, 2000, p. 438). He adds: "All my life I have been studying one thing" (attributed to Simon, in Feigenbaum, 2013 [1989], p. 180) and concludes: "I am really a monomaniac" (*ibid*).

This affirmation may seem baffling due to the amplitude of Simon's work, but, in fact, his relative narrowness of theme helps to explain how the author managed to be "a scientist,

regard, it was the cyborg scientists, and not the neoclassicals, who sought to confront the disturbing implications of these mathematical paradoxes [...]" (Mirowski, 2002, p. 23). Therefore, the critique which would shape the style of revision of mainstream economics in the Cold War has its roots in the cyborg sciences, and not in traditional heterodox economics.

but in many sciences”. Despite the amplitude of his research, Simon “was really working on just one thing”¹⁹ (Augier, 2000, p. 412) – he always looked at behavior. This furnished him a common language and frame of reference, even if the subject matters were diverse. As Augier (2000) comments, “Simon was the perfect example of why disciplinary boundaries don’t always work” (p. 412) – his basic framework was preserved from one science to the other: “here lies the crux of understanding Simon’s approach to science” (Augier, 2000, p. 412). The question that arises is, “What is this framework, then?”. This section tries to answer this question.

Due to Simon’s focus in decision-making, the conception of intelligence becomes central. Intelligence is defined in Simon (1969) as “the work of symbol systems” (p. 23), which can be understood as any system that can process patterns – be the system a computer or a human brain:

A PSS [Physical Symbol System] is simply a system capable of storing symbols (patterns with denotations), and inputting, outputting, organizing and reorganizing such symbols and symbol structures comparing them for identity or difference, and acting conditionally on the outcomes of the tests of identity (Simon, 1995, p. 104, adapted).

According to Simon, “A physical symbol system (PSS) has the necessary and sufficient means for general intelligent action” (Simon, 1995, p. 104). This does not imply that every PSS is an intelligent system but does imply that every intelligent system is a PSS. This also extends to Artificial Intelligence (AI), which is the domain that “deals, in particular, with the phenomena that appear when computers perform tasks that, if performed by people, would be regarded as requiring intelligence – thinking” (Simon, 1995, p. 95). AI, therefore, is a special case of symbolic processing, which emulates biological thinking – what Alan Turing famously called the ‘imitation game’. This is why a calculator, is a PSS but not artificial intelligence, and AIs like Generative Pre-trained Transformer (or GPT) can be considered *intelligent* in Simon’s framework. The calculator can process pre-programmed mathematical operations, but cannot

¹⁹ According to Augier (2000), this was the response of Simon when asked about how he does so much and how does he work on so many areas: His response was that he didn’t; he only studied one subject but applied it in many different fields.

search for solutions. The human mathematician is intelligent, for he or she searches for answers, formulates them, and finds ways to express them in a machine – the calculator. In no moment does the calculator itself search for solutions, as its capability of action is limited *ex ante*.

This concept of intelligence reveals the deep influence of Turing, father of computer science, in Simon's work²⁰. Creator of the famous 'Turing test', this scientist proposed that a computer can be considered as capable of thought if the answers it gives are indistinguishable from those of a human being (undoubtedly, a thinking creature). As Mirowski (2002) observes: "Turing set the dominant tone for decades of conflating machine intelligence with simulation" (p. 462-463). This 'imitation game' therefore is the domain of Artificial Intelligence (Mirowski, 2002, p. 85).

As a result, the work of the AI scientists is mainly to develop systems that can imitate intelligent human behavior. A crucial part of this "human-like" acting is creating systems that adapt to different circumstances: "to begin building a theory of intelligent systems" (Simon, 1996 [1991], p. 96) and "to construct programs [...] that exhibited intelligence by using processes like those used by humans in the same tasks" (Simon, 1996 [1991], p. 96, adapted). He summarizes his view as follows:

"[...] in the domain of artificial intelligence, we need theories of the characteristics and underlying principles of systems capable of holding information in memory and retrieving it when needed, and theories of problem-solving systems, systems for inducing concepts, systems for learning, systems for navigating and operating in the external world (robots), systems for understanding human speech so on. The more powerful these theories, the more we can anticipate what properties systems in these domains must possess to exhibit intelligence (humanoid or other), and the better the systems we can design." (Simon, 1995, p. 115).

In the spirit of "a good theory of how humans create novels will create novels" (Newell and Simon, 1972, p. 11), Simon often stressed the fundamental need to develop theories that explain the task at hand (writing novels, playing chess, etc.) before any attempt at emulating these "intelligences". Only when a robust theory is available, can AI be made to do such tasks intelligently: "If we wish to know how an intelligent person will behave in the face of a

²⁰ Simon wasn't only "quite happy to claim Turing [...] as a precursor" (Mirowski, 2002, p. 463) but even suggested that the Turing test was "the method of science" (*ibid*).

particular problem, we can investigate the requirements of the problem” (Simon, 1990a, p. 6). Therefore, “intelligence consists precisely in responding to these requirements” (*ibid*). AI, therefore, depends on the current understanding of the structure of the problem that is trying to be solved: “The invariants must be sought in the mechanisms that allow them to solve problems and learn: the mechanisms of intelligence” (Simon, 1990a, p. 17). As such, a good theory of novel writing will write novels; a good theory of chess will play good chess; and a good theory of intelligence will create intelligence.

Consequently, AI for Simon is not an *imitation* of intelligence but *a form* of intelligence that *imitates* another form of intelligence. Computers are fully capable PSSs and have the ability to engage in non-mechanic activities (the best example is computer simulations). In Simon’s criteria, this *is* intelligence – not only an imitation of it (remember, there is no ontological difference between artificial symbol processing and organic symbol processing). The question, therefore, is not *if* a specific system is intelligent or not, but *in which way* this system is intelligent. Human beings are highly intelligent creatures, in part for their capability to adapt to a vast number of circumstances. Computers are less adaptable but far more potent in areas such as basic mathematics or logic. In Simon’s conception, it is wholly illogical to classify computers or human beings as one being “more intelligent” than the other – both types of intelligence are different. To him, “Homo sapiens shares some important psychological invariants with certain nonbiological systems – the computers”²¹ (Simon, 1990a, p. 2).

As a result, the goal of AI is not to imitate intelligence – there is no need to do it; it already is an intelligent system – but to imitate the results achieved by other types of intelligence: biological systems. The main trait shared among machines and human ‘minds’ is limitations – the incapability of resolving problems in a complete manner. Both need “shortcuts”, heuristics, to deal with the infinitely complex world. The process of “learning” consists, basically, of a process of uncovering new heuristics or improving existing ones.

Due to the limitations of intelligent systems, optimal outcomes are effectively impossible. Maximizing solutions (end-states where the desired goal is identical to the end result) can only happen if an agent has an “unbounded minds”, free of any restriction and

²¹ His infatuation with “computer minds” merited the mocking observation by Walter Weimer, who described Simon as “psychologist of machine” (Baars, 1986, p. 307-308 *apud* Mirowski, 2002, p. 456).

capable of complete cognition. As all intelligent systems are subject to limitations, Simon considers optimization impossible in practice. As Loasby (1989) comments, for Simon, “Life is a succession of problems, to be tackled successively. It appears to embody no grand theories, no opportunities for global optimization.” (p. 141). In place of optimization, decision-making in Simon is guided by his famous concept of ‘satisficing’, a world-state where the desired goal is “close enough” to the actual results; it is satisfactory.

One must note the utilitarian nature that ‘intelligence’ takes on Simon’s work – the mind serves to solve problems and to approximate results from the desired ones. Simon gives no consideration to the emotional aspects of intelligence, or non-instrumental knowledge, such as art appreciation or idle abstraction. Therefore, Simon closely relates ‘intelligence’ – “the abilities to learn from experience, adapt to new situations, understand and handle abstract concepts, and use knowledge to manipulate one’s environment”²² – and ‘rationality’ – “the use of knowledge to attain goals”²³. Effectively, to Simon, intelligence is the path to rationality, or as Brette, Lazaric, and Vieira (2017) summarize: “[...] Simon defines rationality in terms of results. In this perspective, human behavior is rational to the extent that it produces appropriate results, as could be assessed by an external observer — whatever the nature of the process that has led to this result” (p. 576).

In Simon’s work, ‘rationality’ can be understood as the relation between a set of cognitive mechanisms (symbol processing, learning, computation, etc.) and the actions that they enable. An action is more ‘rational’, therefore, if the desired goals and actual results align. It is more ‘rational’ if the distance between the two is shorter. Intelligence, on the other hand, is the means in which a system searches and applies different rationales (Simon, 1969). As such, rationality is context specific – but not intelligence.

In a manner of speaking, rationality is the result of interaction between intelligence and the environment. This link is stylized as a “problem space” (Simon, 1964). Different from the “task environment”, which is independent of the agent, a ‘problem space’ relates directly to the possible solutions for the task and, therefore, also depends on the agent’s ability to compute and apply these solutions. As Fiori (2011) highlights, the “internal representations” – how an agent

²² Source: <https://www.britannica.com/science/human-intelligence-psychology>

²³ Source: <https://www.britannica.com/topic/rationality>

interprets the outside world – also are important, for they, along with the task environment and the problem space, “determine the relation between the state of the world and individual subjectivity” (p. 591). Brette Lazaric and Vieira (2017) explained it in another manner: “Because organizations, as well as men, are mostly imperfectly adapted organisms, understanding and predicting their behaviors require more than simply examining the surrounding environment in the light of the aims at hand; it requires opening the black boxes that lie within” (p. 7).

The more structured a problem space is, the closer the approximated results may be to the desired outcome. As a result, ‘computing power’ (the ability to process information) is secondary – the focus is the relation between agents’ “rationalities” and the problem at hand:

“[...] humans, whose computational abilities are puny compared with those of modern super-computers or even PCs. are sometimes able to solve, with little computation, problems that are very difficult even by computer standards-problems having ill-defined goals, poorly characterized and bounded problem spaces, lack of strong and regular mathematical structure” (Simon, 1995, p. 98).

This is strongly reminiscent of the way AI works: “[...] a running AI program contains a definition of the goal and knowledge about the task domain, as well as problem-solving processes. It also contains strategies, some of which may be task-specific.” (Simon, 1995, p. 107). Consequently, the characterization of ‘Bounded Rationality’ as mainly a collection of critiques to Substantive Rationality is misleading. As argues Mirowski (2002):

“[...] it is an egregious error to write off Simon's Bounded Rationality as merely the shopworn complaint of the naive empiricist that “people really don't think like that,” perhaps pleading for a more nuanced input from psychology into economics (although he sometimes does personally come across this way)” (p. 463).

Modeling behavior, consequently, is an attempt to explain the interaction between an agent’s limitations, their goals, and the task environment. This type of modeling is the “small-

m methodology”²⁴ of Simon – its research method. Bounded Rationality, a fundamental touchstone of this methodology, can be understood as the systemic nature of problem spaces that arise from the interaction of complex systems and agents with cognitive limitations. These problem spaces are doomed to be only partially solved – every time, for optimization is impossible.

Procedural rationality, for instance, is simply the chaining of solutions to problem spaces, one after the other, as a response to a complex task, effectively separating a larger problem into small pieces:

“The study of human behavior in the face of difficult tasks shows why we need a theory of processes (procedural rationality) as well as a theory of the requirements of the task (Substantive Rationality). A theory based only on task requirements could not tell us how behavior depends on knowledge of relevant cues or strategies.” (p. 5).

The limitations of Bounded Rationality, therefore, are not the problem spaces themselves, but the structure that they need to have to fit into this arrangement. The reader should have noted that Bounded Rationality can only meaningfully guide the interpretation of an agent when a task has a solution or discernible end. Other behaviors, such as habits, cannot be described in the framework presented here. This reveals the instrumental nature of intelligence implied in Bounded Rationality – again, “the use of knowledge to attain goals”. Therefore, Simon’s rationality is result-oriented. Loasby (1989) illustrates this point by explaining the role of “learning” in Simon’s rationality: “Simon looks back only to draw the lessons of immediate experience and then quickly forward to the next problem” (p. 141).

As a consequence, Simon does not develop a theory of mind that encompasses erratic mental processes, such as sentiments or intuition. The mental life of Simon’s agent is reduced to symbols and symbol manipulation (Fiori, 2011). As such, emotions or intuitions are only incorporated into Simon’s theories as additional parameters of choice²⁵, stylized as variables or

²⁴ According to Dow (2021), after the famous work of Deirdre McCloskey (1985), it became customary to differentiate between “big-M Methodology” and “small-m methodology”. The first refers to the philosophy of science; the second to non-philosophical study of method.

²⁵ I attribute this formulation to Ronaldo Buzzo.

reduced to “noise”. Consequently, Simon does not provide a theory which incorporates emotions or irrationality in economics, even if some researchers suggest it²⁶. As Fiori (2011) summarized:

“This implies that emotional or automatic responses to complex phenomena are not considered in light of their specific neurobiological or psychological dimension, in that Simon’s AI examines only general and symbolic traits of intelligent behavior, irrespective of the physical features of subjects (men, computers etc.).” (2011, p. 594).

Simon’s framework is the byproduct of a number of critiques of Substantive Rationality, assembled in the ‘cyborg’ manner described above and culminating in an agent which acts in an analogous manner to a computer. Simon implemented a number of general steps in his research. Even so, the most common use of Bounded Rationality is “the negative conception of global rationality” (Barros, 2010, p. 459, own translation), not defined positively²⁷.

That being said, Bounded Rationality presents an identifiable structure which is different from a reformed version of traditional unbounded rationality. For instance, the stopping rule of traditional microeconomics is the identification of the optimal solution; the stopping rule of Bounded Rationality is ‘satisficing’, which consists of a random selection between satisfactory outcomes. This is the method chosen by Simon to deal with situations in which an agent has to choose between incomparable categories and when optimization is not possible: “[...] a satisficing choice can still be made as soon as an alternative is found that (a) is satisfactory along all dimensions of value, (b) has satisfactory outcomes for all resolutions of the uncertainty, or (c) is satisfactory for all parties concerned, respectively” (Simon, 1990c, p. 10). As such, “more than minor tampering with existing optimization theory is called for” (quote attributed

²⁶ The origin of this notion probably is in the works of Daniel Kahneman and Amos Tversky, which do incorporate emotions in economic decision-making (affecting utility) and are often associated with Simon’s work. According to Fiori (2011), “emotions play a wider role [in Kahneman and Tversky] than in Simon’s theory, where they are described in symbolic terms as mechanisms able to help agents select and focus conscious attention on specific objects, and consequently define priorities in the search associated with problem-solving activity.” (p. 603).

²⁷ What something *is*, instead of what it *is not*.

to Simon in Gigerenzer, 2004, p.1). His framework of Bounded Rationality, then, can be summarized as such:

Table I – Structure of Bounded Rationality

Step	Description
1. Satisficing Criteria	A desired outcome is defined <i>ex ante</i> , which defines the agent's criteria for 'satisfaction'.
2. Problem-space Apprehension	The task environment and the limitations of the agents define the problem space, and the agent starts to consider plans of action.
3. Plan of Action	A plan of action is selected by the 'satisficing' criteria, and the agent starts to act, periodically revising this plan if the current course of action is perceived as "insatisfactory".
4. Satisficing	When the first satisfactory is achieved, the decision-making process ends.

Source: own elaboration.

The interactions between environment and agents create different “rationales”, which are more or less appropriate to one task or another, resulting in a “solution”. If the solution is approximately the desired outcome, it is considered “satisfactory”; otherwise it is not.

In this case of procedural rationality, a fifth step is added. Based on the solution of step 4, another task environment is generated, and a new problem space arises (step 2), initiating another plan of action (step 3) and resulting in an end-state result (step 4). Step 5 is repeated until the process is perceived as “complete” (the solution to the problem is considered “satisfactory”). Note that decision completion and procedural completion are different. Decisions refer to specific tasks, while processes refer to a chain of different tasks. Therefore, (traditional) Bounded Rationality applies to decisions, while procedural rationality applies to decision processes (Simon, 1990a)²⁸.

This general process of decision-making is the identifiable characteristic of Simon's research process. It is his “small-m methodology”, the processes which constitute his science.

²⁸ A note should be made on Gohmann's (2019) “Common Perspective of Bounded Rationality” (p. 311, own translation), which is reminiscent of the model present here and describes a “common model” of decision-making in Behavioral Economics (see *Guia de Economia Comportamental*, p. 311). We opted to not adopt this explanation here, for it is not specific to Herbert Simon, as Gohmann used the satisficing criteria of “utility”, which, we argue, is not the only one adopted by Simon.

In the next section, we tackle the “big-M Methodology” of Simon – the philosophy of science which structures Simon’s thought process.

I.3 – The Methodology of Simon’s Artificial Science

This section will present the methodological basis of the development of Simon’s specific way of “science-making” and how it relates to the two first topics of discussion (artificiality and behavior). He was influenced by many sources but incorporated them in unique ways, rarely seen in other scientists. Here, we attempt to show how this conceptual absorption occurred.

Simon graduated in a time and place effervescent with economic ideas. He lived through the 1929 Great Crash and the end of traditional *laissez-faire* capitalism. Among the most famous reactions to these developments are the rise of the Vienna Circle and the Keynesian Revolution. Economic science in the US was especially plural in the 1930s (Hodgson, 2007; Rutherford, 2011). Neoclassical economists coexisted with the Institutional Economics of the Veblen-Commons tradition (now deemed Original Institutional Economics or OIE) and the solidifying Keynesian approach. Simon, thus, was influenced by a plethora of ideas and traditions, an inescapable fact to academics at the time.

In methodological terms, the main discernible influence of Simon was the Vienna Circle²⁹, often associated with the Austrian School of Economics. The Vienna Circle was a group of intellectuals engaged in the study of language and scientific philosophy³⁰. The influence of the Circle in Simon can be traced to one of his most notorious professors, Rudolf Carnap. Carnap was a notable member of the Circle and espoused a radical interpretation of empiricism, which would be known in posterity as ‘logical positivism’. According to it, “scientific knowledge is the only kind of factual knowledge and [...] all traditional metaphysical doctrines are to be rejected as meaningless”³¹. Scientific knowledge, in this

²⁹ The Vienna Circle was also known as the *Verein Ernst Mach* (Ernst Mach Association). Among the figures who had ties with the Vienna Circle is notable science philosopher Karl Popper.

³⁰ Source: <https://www.britannica.com/topic/Vienna-Circle>

³¹ Source: <https://www.britannica.com/topic/logical-positivism>

interpretation, refers only to testable knowledge, as in the criterion of ‘falsificationism’ of Karl Popper (the term is Blaug’s, 1992). As a result, this empiricism is not a reproduction of Hume’s method of observation³², for logical positivists reject the existence of any kind of metaphysical, extra-mundane knowledge.

Carnap, and other intellectuals of the Vienna Circle, believed that not only an objective underlying reality exists (i.e., the “real” is not defined by subjectivity, nor is it metaphysical), but also that it was the role of the scientists to identify this reality. This position of logical positivism can be summarized in Viskovatoff (2001): the conviction that “knowledge lies in sense-impressions, which are ‘selfjustifying’ since they are ‘immediate’, that is, bits of knowledge not derived from other bits of knowledge” (p. 324-325)³³. Induction, the linking of knowledge, may result in mistakes due to false correlations made in the abstraction process – this is the famous ‘problem of induction’, proposed by Hume.

In logical positivism, language plays a key role, for clear communication (precise words and precise meanings) is essential to communicate this underlying reality. In this interpretation, loosely defined words are not sufficiently rigorous to guide scientific investigation. Carnap’s scientific efforts presented two main elements: 1) a profound preoccupation with the formalization of knowledge and 2) the recognition that this scientific program required a shared, unambiguous language between scientists³⁴. Hence, the role of the logical positivist to Carnap is not only to discover reality, but also to present it in a rigorous manner. For logical positivism, this is the *only* form of scientific knowledge, epitomized in Bridgman’s Operationalism: “the attempt to define all scientific concepts in terms of specifically described operations of measurement and observation”³⁵. Rigorously defined and reflecting an underlying truth, knowledge in logical positivism attempts to unshackle itself from all political implications, becoming a “neutral fact”, effectively aiming to be apolitical in its implications³⁶.

³² Ibidem.

³³ This passage was a comment on the work of Wilfrid Sellars, *Empiricism and the Philosophy of Mind* (1956).

³⁴ Source: <https://www.britannica.com/topic/logical-positivism>

³⁵ Source: <https://www.britannica.com/topic/operationalism>

³⁶ It is interesting to note that this position is almost the exact opposite of pragmatic philosophy, which also influenced Simon, although in a much more timid way.

These ideas shaped Simon's views on what was science and the scientific method. As Crowther-Heyck (2005) points out, "[Carnap's] intensive study of formal logic and positivist philosophy had a profound, lasting effect on [Simon's] thinking" (p. 73 *apud* Fiori, 2011, p. 593, adapted). This is important, for Simon graduated in a time when economics was beginning its long road toward increasingly formalized mathematics. For logical positivists, however curiously, this formalization did not increase the rigorousness of the knowledge portrayed in economic models but, instead, further distances it from the underlying truth – economic formalization heavily relies on induction. This occurred at a point of time when the validity of economics as a science was at stake, for the discipline did not adopt the empiric methods that are required to attain rigorous, scientific knowledge³⁷.

Some of these elements can be seen in Simon's condemnation of "casual empiricism and [...] armchair introspection" (Simon, 1996 [1991], p. 59). According to the author (Simon, 1969), modern economics makes use of "heroic abstraction and idealization of human rationality" (p. 23), which deviate from the phenomena at hand. Simon believed that the study of empirical phenomena was the essence of the scientific method (Simon, 1969). Observation, therefore, is the fundamental unit of investigation, betraying his amazement at the advances of natural sciences. In fact, one of his ambitions was to emulate these advances in the humanities: "The social sciences, I thought, needed the same kind of rigor and the same mathematical underpinnings that had made the "hard" sciences so brilliantly successful"³⁸. Therefore, by attacking the increasingly abstract mathematical efforts of economists, we must agree with Gigerenzer (2004) when saying that "Bringing sanity into theories of rationality was a major guideline in Herbert Simon's scientific life" (p. 389).

However, we must keep in mind that this so-called "sanity" has specific baggage that links it inextricably with logical positivism. Due to the influence of Carnap, Simon's conception of "truth" is close to that of logical positivism – testable knowledge, formalized models, rigorous observations, etc. In Simon's (1996 [1991]) own words: "I [...] embraced a logical positivism that I have never relinquished" (p. 44). He adds in parentheses "(I would prefer to

³⁷ Logical positivism was just one of many attempts to remedy this situation. Among other attempts to introduce "empiricism" to economics is the famous work of Institutionalist Wesley C. Mitchell that, to the contrary of logical positivism, did not attempt to strictly formalize economics.

³⁸ Source: <https://www.nobelprize.org/prizes/economic-sciences/1978/simon/biographical/>

call it empiricism now)” (*ibid*). As such, we can “identify our author in logical positivism” (Cunha, 2022, p. 637, own translation), but we also should keep in mind that Simon was not a traditional logical positivist. As Fiori (2011) points out, Simon “accepted some conclusions of logical positivism” (p. 593)³⁹ – but not every conclusion. Instead of the more strict interpretation of logical positivism, Simon espoused a “special type of realism” (the term is Cunha’s, 2022, p. 635) that resembles logical positivism but differs from it in some key ways. The similarities were already explained (conception of truth, favored empiricism, importance of rigor, etc.), but what of the differences?

Firstly, Simon never adopted the strict views on language that defined much of Carnap’s work. Simon did not focus his career on figuring out the “operational definition of terms [...] crucial to making statements clear and testable” (Crowther-Heyck, 2005, p. 101 *apud* Fiori, 2011, p. 593). Instead, Simon preferred “abstract and non-specific definitions” (Cunha, 2022, p. 650, own translation), which are common in social sciences, but they go against the tenets of logical positivism. An interesting example is Simon’s definition of ‘organizations’. As points out Cunha (2022), Simon never presented a complete definition of the term, even if his Nobel was partially granted due to “his pioneering research into the decision-making process within economic organizations”⁴⁰.

Incredibly, the same applies to “Bounded Rationality”: the term was never defined (Barros, 2010). This vagueness of definition was recognized by Simon himself: “I guess a major reason for my using somewhat vague terms — like Bounded Rationality — is that I did not want to give the impression that I thought I had “solved” the problem of creating an empirically grounded theory of economic phenomena.” (Herbert Simon in an e-mail to Gerd Gigerenzer, available in Gigerenzer, 2004, p. 10). This shift away from logical positivism was gradual, as Simon’s early works were more closely tied to logical-positivists conception of rigor if compared to his later works. Therefore, these conceptions of “rigor” shaped Simon’s taste for formalization of knowledge, but he did not adopt an analogous position to that of Carnap.

Secondly, Simon had an ambiguous relationship with the “apolitical” aspects of logical positivism, seeing in them an important guideline to make science rigorous but also recognizing

³⁹ Fiori (2011) was referring specifically to Simon (1976 [1947], p. 45) in here.

⁴⁰ Source: www.nobelprize.org/prizes/economic-sciences/1978/simon/facts/

the importance of social aspects in defining human life. The tendency to ‘apoliticism’ was specially apparent in his youth, but gradually faded. On any occasion, he never focused on social matters.

Simon’s ‘special type of realism’, therefore, can be interpreted as a general admiration of the testable and rigorous methods of natural sciences and an attempt to “translate” these principles to social sciences and economics. However, social sciences are fundamentally “untestable”, for every social phenomenon is unique – it happens in an evolutionary historical system, subjected to non-ergodic entropy⁴¹. Therefore, social conditions that caused one phenomenon or another are unrepeatable. Simon’s focus on simulations and AI acquires another implication here: this would be a tool to help translate the repeatability of natural sciences to the humanities.

What becomes clear at this point is that Simon’s methodology had many phases. Brette, Lazaric, and Vieira (2017) identify an important methodological break after 1958, the year of publication of *Organizations*, along with James G. March. Before this moment, Simon focused in the research of organizations, in line with his degree in political sciences. In this early phase, Simon’s conception of science was more in line with logical positivism, even if not overtly so (Simon, 1976 [1947]). After the book’s publication, however, his interests shifted towards a more individualistic methodology, focusing on individual agents instead of collective assemblies. This was relatively early in his long career⁴², and defined much of the bulk of his “mid-life” works. According to Brette, Lazaric, and Vieira: “From that point on, Simon’s interest in psychology as well as his skepticism toward economics increased” (2017, p. 8).

Besides this relatively early breaking point, another shift can be gleaned between his ‘mid-life’ and his ‘late-life’ works. In short, the logical-positivist roots of his early and mid-career were effectively abandoned in the final stages of his life. His late works were not shy of

⁴¹ Recent developments in Experimental Economics, as those indicated by Bianchi and Silva Filho (2001), put this concept in check. However, Simon made few works in line with Experimental Economics.

⁴² *Organizations* was published in 1958, eleven years after *Administrative Behavior* (1947). The latter was the most successful early work of Simon.

including many references to other approaches to alternative strands of economics⁴³ and other sciences⁴⁴, and thus tended to be much more speculative than his early ones. The more mature Hebert Simon, therefore, drifted ever further away from logical positivism. Unlike in the case highlighted by Brette, Lazaric, and Vieira (2017), this research could not find a single, distinctive breaking point. This transformation was gradual.

The main consequence of this “special-type” of logical positivism is a broad definition of empiricism. For Simon, non-deterministic outcomes are empiric phenomena, even if the results come from replicable man-made systems. AI and simulations, for Simon, are empirical phenomena and generate empirical data (Simon, 1995). Even if the system is made by human hands, the results are not. They are not curated, nor arbitrarily defined. As artificial and natural phenomena are on the same level (as argued in Section I.1), there is no ontological difference between data generated by simulations and generated in the natural world. Therefore, Simon considers the ‘artificial sciences’ as empirical, in the exact same terms as the natural sciences.

As such, the influence of logical positivism in Simon is ambiguous. On one hand, Simon’s methods of research can be directly linked to logical positivism. On the other hand, he was never a true believer in the type of reality espoused by logical positivism and its “excessive formalization” (the term is Fiori’s, 2011). As Cunha (2022) notes, Simon was an example of “this type of naturalist, one observer of itself and others, that did not perform any laboratory experiments” (phrase attributed to William James, available in Cunha, 2022, p. 641, own translation). This is not a demerit of the author, but it is important to understand that ‘empiricism’ for Simon did not mean, necessarily, fieldwork, laboratory experiments, or even the observation of nature. ‘Empirical’ to him meant, foremost, the unsupervised results of elements of a system (artificial or natural).

His style of knowledge formalization, consequently, never adhered to the more strict forms of logical positivism, and (ironically) left a great many gaps in the intended knowledge that was to be communicated. For Simon, ‘rigor’ meant a strict adherence to *what he considered* empiricism – not a strict adherence to formal definitions written in stone. Conceptual definitions

⁴³ In Simon (1978), the author makes reference to Institutional Economics of Commons and Ely, also implying those of Veblen. In Simon (1979), the author dedicated his career “in the service of Political Economy” (p. 493), even if his contributions to the area *per se* were marginal at best.

⁴⁴ In Simon (1978), the author also speculated on the Freudian methodology of psychoanalysis.

in Simon, like his behavioral theories, are context specific – in practice, they don't reflect an “underlying truth”.

It is impossible to say why Simon gradually shifted away from traditional logical positivism, or even if he considered himself a logical positivist. In any occasion, however, the influences of logical positivism are discernible in his science, even if he also deviated considerably from it. One of the more obvious pieces of evidence of his deviation is captured in his own words: “I have never thought of either Bounded Rationality or satisficing as precisely defined technical terms, but rather as signals to economists that they needed to pay attention to reality, and a suggestion of some ways in which they might” (phrase attributed to Herbert Simon in an e-mail, available in Gigerenzer, 2004, p. 10). As we will see in Chapter II, this loose definition of Bounded Rationality is essential to understand the developments of this concept.

In this section, we saw how Simon's ‘special type of empiricism’ came to be and how it changed through the different stages of his life and career. In the next section, we attempt to understand the nature and structure of Simon's critique of Substantive Rationality.

I.4 – Economic Aspects of Herbert Simon and the Critique of Substantive Rationality

As we argued throughout this dissertation, Simon's interpretation of Bounded Rationality can be defined positively – it is a concept with theoretical essence, not mainly a negation of unbounded rationality, even if this was the most common use, as indicated by Barros (2007). That being said, it cannot be denied that Bounded Rationality functions as a criticism of traditional Substantive Rationality. In fact, it can be argued that the critical function of this concept is its most impactful aspect. As such, the explanation of the structure of this criticism (why Bounded Rationality is a critique of Substantive Rationality) is important, even if not the focus of this dissertation. Interestingly, the bulk of the criticism of Substantive Rationality did not come from Simon, but from Simon's commentators – economists that adopted Bounded Rationality and applied it in their specific research objects.

As such, the goal in this section is not to paint Simon as mainly a “critic of neoclassical economics” (Sent, 2005, p. 230), as some authors did (besides Sent, 2005, other examples are Angner, 2019; Gigerenzer, 2004 and 2020). Simon rarely rejected ideas wholesale, even if he

disagreed with them, and his criticism rarely involved the complete refutation of one concept or another. In other words, Simon had many critics of neoclassical economics, but none of them were “dealbreakers”. In fact, Simon used orthodox neoclassical economics in some works (the most notable examples are Holt, Modigliani and Simon, 1955 and Holt et al., 1960), a fact commonly lost on the more critical commentators.

So what was the structure of the critique of neoclassical economics by Simon? His problems with traditional economics start with the ideal type of *homo economicus*⁴⁵, in two main aspects: 1) *homo economicus* does not have to deal with any significant limitations, and 2) it can only apply *one* solution for every problem – optimization. The first point refers to the unreality of *homo economicus* as a description of real-world behavior. As Simon often notes, even machines have perceivable and significant non-negligible, limitations. This makes *homo economicus* not analogous to any machine (even supercomputers), but more in line with supernatural beings – a ‘Laplacean demon’, as Gigerenzer and Goldstein (1996) deem it. As Conlisk (1996) summarizes, Simon turns rationality into a scarce resource, a type of scarcity unknown to *homo economicus*.

The second point refers to the adaptability of *homo economicus*. As it can only perform one type of problem-solving technique – optimization – every task environment needs to align with this process of problem-solving. This follows the logic of “if you have a hammer, all problems become nails”; if you can only optimize, all decisions are problems of optimization. Neoclassical economics indeed does this by constructing its theories in a way to circumvent or avoid situations in which optimizing behavior is impossible. Traditional neoclassicism, for instance, does not deal with strong uncertainties, preferring probabilistic, weak uncertainties (Dequech, 2011); it does not allow its agents to behave irrationally (transitivity, monotonicity, stability of preferences, etc.) and presupposes ergodic worlds where the distributions of probability may be known (the future is conceivable).

Consequently, while *homo economicus* acts perfectly, it may only do so in a limited set of circumstances. The ‘stopping rule’ of neoclassical economics is a good example of these limits. As *homo economicus* will *only* stop the decision-making process when optimization is

⁴⁵ Used broadly in neoclassical economics, *homo economicus* refers to a rational agent with no intrinsic limit to information processing. This is a wholly rational being, capable of “optimization behavior” – actions that generate the most utility given income restrictions.

achieved, it will never make any decision if optimization is not possible. *Homo economicus* is supernaturally potent in the maximization process, but it is useless in every non-maximizable situation. In practice, *homo economicus* requires that every problem space to be similar, but Bounded Rationality does not have this limitation. Ana Maria Bianchi (2019) summarized this dynamic as such: Bounded Rationality admits the existence of the “multiple self” – the idea that “a person can be seen as (or even be) a set of sub-selves” (p. 222, own translation), while *homo economicus* can’t.

The center of Simon’s critique of neoclassical rational behavior rests in two pillars – unrealism and rigidity. Therefore, it is a critique of behavior in both intensity (capacity for acting and thinking; unlimited in *homo economicus* and bounded in Simon’s) and structure (explanations do not conform to a single, unifying logic; maximization in *homo economicus* and satisficing in Simon are not analogous).

This last point, that Bounded Rationality is a critique of structure, may seem controversial. Now, isn’t ‘satisficing’ just ‘maximization’ with less intensity – the second or third best instead of *the* best? Simon argues this is not the case – the gap between rational behavior and *almost* rational behavior is wider than it may seem at first. Stiglitz (2010 *apud* Gigerenzer and Gaissmaier, 2011) resumes this idea as such: “It simply wasn’t true that a world with almost perfect information was very similar to one in which there was perfect information” (p. 243) This is seen in Simon’s views on the work of Milton Friedman:

“Many economists, Milton Friedman being perhaps the most vocal, have argued that the gap between satisfactory and best is of no great importance, hence the unrealism of the assumption that the actors optimize does not matter; others, including myself, believe that it does matter, and matters a great deal” (Simon, 1969, p. 29).

It is under this logic that we argue that Bounded Rationality is not a reformed version of Substantive Rationality, but another explanation of behavior – the structure of explanation of behavior are different. This may suggest Simon was, indeed, this “staunch critic” of neoclassical economics⁴⁶, but we argue that this is not the case. Due to the adaptive nature of rationality (refer to Section I.2), bounded and Substantive Rationality explain *different* behaviors. Every set of solutions “must be, in a certain fundamental sense, incomplete. It is a

⁴⁶ Sent, 2005; Earl, 2010; Steingraber and Fernandez, 2013; and Gigerenzer, 2020, are some examples of authors that propose this.

truth of mathematics and logic that any program (including those stored in human heads) must be unable to solve certain problems” (Simon, 1990c, p. 5).

Consider the following quote: “In contrast to a situation where the adaptation process is itself problematic, we can predict the system's behavior without knowing how it actually computes the optimal output. We need to consider only Substantive Rationality.” (Simon, 1969, p. 26). In this case, Simon proposed that Substantive Rationality is sufficient to describe the behavior of an agent in a well-structured problem space, in which the outcomes are known and computation was possible. In this case, a solution that took Substantive Rationality as a basis is appropriate – it reaches a reasonable solution that, in this case, closely resembles (or, in the limit, is) the optimization case. To think of a solution for an abstract problem only becomes unfit when someone insists that, actually, this abstract solution also applies to non-abstract circumstances.

Therefore, Simon’s critique of unbounded rationality is constructed around its unfitness to describe real decisions – but not as a theoretical tool. In fact, Simon made recurrent references to the usefulness of Substantive Rationality as a normative artifice, something to define what a “good” or “satisficing” solution could be – “the requirements of the task” (Simon, 1990c, p. 5), in his words. Unlike in neoclassical economics, however, Simon’s Bounded Rationality does not present a “normative normativity” – a required set of normative solutions defined *ex ante*. Instead, “normativities” are defined as part of the problem space, derived from what is satisfactory to an agent in each case and taking account of the desired outcomes. As a consequence, if the desired outcome is maximization and the agent can maximize the problem space, internal representations, and task environment are aligned. In this case, optimization is abstractly feasible.

This is what is behind the many recommendations that Simon made to orthodox economists. For instance, Simon (1978) comments that economics will have “to devote major energy to building a theory of procedural rationality to complement existing theories of Substantive Rationality” (p. 14-15). In this case, Simon recognized the need to “complement”, but not to reject, traditional Substantive Rationality. In fact, in commenting on Holt et al. (1960) (a research where Simon applied optimizing microeconomics), Simon was clear about his thoughts on rational optimizing theories: “I have no urge to apologize for our decision procedure as a useful management science tool” (Simon, 1990a, p. 15-16), defending that the version of maximizing behavior functioned as an “approximation” of real behavior.

Another blow to the idea that Simon was mainly a critic of neoclassical economics is the vagueness of his critique to it. For instance, many classical preoccupations of economics were only marginally dealt with by Simon, such as equilibrium, economic policy, or the effects of uncertainty on economic activities⁴⁷. Loasby (1989) synthesized this situation as: “Simon’s central assumption implies an interest in processes rather than equilibrium, but, since all models are incomplete representations of phenomena to which they are applied, equilibrium concepts may be used on occasion” (p. 142). The findings of this dissertation are in line with Loasby’s observation. A similar vagueness can be gleamed in matters of economic policy, uncertainties, and other classic economic matters – in other words, the critique of mainstream economics, even the critique of rational microeconomic behavior, was not Simon’s focus.

Consequently, Simon’s critique of rationality is both of intensity and structure. However, this is not a radical critique. The structure of Simon’s Bounded Rationality does not contradict neoclassical economics in any fundamental way, even if it implies the necessity of some revision.

In the end of the day, did Simon believe that people were, indeed, bounded rational agents? Or this is simply one convenient tool to analyze behavior? It is hard to say, but violations of Substantive Rationality are often observed. As Conlisk (1996) succinctly summarized:

“Hundreds of studies of this type have been done, mostly by psychologists but more recently by experimental economists also. There is a mountain of experiments in which people: display intransitivity; misunderstand statistical independence; mistake random data for patterned data and vice versa; fail to appreciate law of large number effects; fail to recognize statistical dominance; make errors in updating probabilities on the basis of new information; understate the significance of given Sample sizes; fail to understand covariation for even the simplest 2x2 contingency tables; make false inferences about causality; ignore relevant information; use irrelevant information (as in sunk cost fallacies); exaggerate the importance of vivid over pallid evidence; exaggerate the importance of fallible predictors; exaggerate the ex ante probability of a random event which has already occurred; display overconfidence in judgment relative to evidence; exaggerate confirming over disconfirming evidence relative to initial beliefs; give answers that are highly sensitive to logically irrelevant changes in questions; do redundant and ambiguous tests to confirm an hypothesis at the expense of decisive tests to disconfirm; make frequent errors in deductive reasoning tasks such as syllogisms; place higher value on an opportunity if an experimenter rigs it to be the

⁴⁷ Simon made offhand comments on equilibrium. In Simon, 1969 he spoke of an “Order without a Planner”; and in Simon, 1978 he suggested that only relatively weak conditions need to be in place for equilibrium to be possible), but did not systematize his views on the matter.

"status quo" opportunity; fail to discount the future consistently; fail to adjust repeated choices to accommodate intertemporal connections; and more." (p. 670).

These contradictions are sufficient, we think, to make a compelling case that people are not substantive rational agents. In the next section, we explore some more of the curious place that Herbert Simon occupies in the economic sciences.

I.5 – Herbert Simon in the Orthodoxy-Heterodoxy debate

Herbert Simon defies the usual classifications of orthodoxy and heterodoxy in economics. Possibly for this reason, Simon found broad acceptance from all points through the spectrum. Bounded Rationality is used by neoclassical economists such as Kenneth Arrow⁴⁸ (1986), mathematical modeling-oriented economists such as Ariel Rubinstein⁴⁹ (1998), the strands of economics that combine elements from the traditional orthodoxy and heterodoxy, such as Richard Nelson and Sydney Winter⁵⁰ (1985) or Oliver Williamson⁵¹ (1989), and more traditional heterodox economists, such as Paul Davidson⁵² (2010) or Marc Lavoie⁵³ (2022).

As we argued in Section I.4, the specific strand of “economics” that Simon is part of is hard to classify. Simon escapes the usual trappings of orthodoxy; however, he also does not approach the established traditions of heterodoxy. His methodology exists somewhere between

⁴⁸ “Obviously, I am accepting the insight of Herbert Simon [...] on the importance of recognizing that rationality is bounded” (Arrow, 1986, p. 397).

⁴⁹ “[The] motivation for studying models of Bounded Rationality springs from our dissatisfaction with the models that adhere to the “perfect rational man” paradigm” (Rubinstein, 1998, p. 2, adapted).

⁵⁰ “Man's rationality is "bounded": real-life decision problems are too complex” (Nelson and Winter, 1985, p. 35).

⁵¹ “I have previously argued that contracting man is distinguished from the orthodox conception of maximizing man in two respects. The first of these is the condition of Bounded Rationality.” (Williamson, 1989, p. 138).

⁵² “Keynes was the first behavioral scientist!” (Davidson, 2010, p. 251).

⁵³ “We shall also see that the framework developed by all these economists is quite coherent with the notions of uncertainty and Bounded Rationality” (Lavoie, 2022, p. 124, adapted).

a reformed version of logical positivism (Section I.3), an alternative to maximizing behavior of orthodox economics, and a specific framework for decision-making (Section I.2). This section aims to locate this unique style of Simon's in the Orthodoxy-Heterodoxy debate. Was Simon an orthodox economist? A heterodox one? Do these classifications even apply? In this section, we propose an interpretation. Due to the complexity of the theme at hand, this section is lengthier than the other sections in this chapter and, unlike them, is subdivided.

Orthodoxy and Mainstream

The concepts of 'orthodoxy' and 'mainstream' are often confused, for, as argued by Dequech (2007), some economists treat them interchangeably. Dequech differentiates both by proposing that the latter refers to an intellectual heritage and the former refers to sociological criteria. Let us start defining 'mainstream'. For Dequech (2007), "The concept of mainstream economics is based on prestige and influence and includes ideas taught in prestigious schools." (p. 279). Therefore, orthodoxy (which will be defined latter) is part of the mainstream, but the mainstream is not only orthodoxy. In fact, he argues that the period between the 1990s and the 2000s was one of a pluralistic mainstream. Economics that were not in line with traditional aspects of orthodoxy were being taught in the most prestigious schools, printed in the most prestigious books, and were in the minds of the most prestigious economists. Some examples are Behavioral Economics, New Institutional Economics and Evolutionary Game Theory, all of which are mainstream but deviate from (traditional) orthodoxy in one way or another.

Therefore, it is important to stress that both 'mainstream' and 'orthodoxy' are not static categories. This is harder to note in the case of orthodoxy, which can be directly traced back to Adam Smith and continually occupied one of the most prestigious positions in the economic sciences, even the most prestigious position in some occasions. If focused on the USA and the UK, it is generally true to say that the 'orthodoxy' is neoclassical⁵⁴, but the same cannot be said

⁵⁴ The geographical identification is important. Germany, for instance, followed the tradition of the German Historical School and did not have a prestigious strand of neoclassical economics before the second half of

about the ‘mainstream’. What this implies for our analysis is that the classification of Herbert Simon as ‘mainstream’ does not mean that he was an ‘orthodox’ economist necessarily.

Little doubt exists on the mainstream status of Herbert Simon. Besides the ample impact of Bounded Rationality and his high prestige, one can also cite the fact that he won the Nobel Prize in economics in 1978⁵⁵, the Turing Award in 1972, the National Medal of Science in 1986 (given to scientists in the service of the US government, bestowed by the president himself; Ronald Reagan in the case), the Lifetime Achievement Award in 1993, the John von Neumann Theory Prize in 1988, and the Dwight Waldo Award in 1995⁵⁶. If mainstream status indeed depends on prestige, Simon is comfortably placed in its ranks.

What is to note is that Simon became part of the mainstream in an era different from the current one. Hodgson (2007) and Rutherford (2011), for instance, highlight the fact that the mainstream in the US in the early 20th century included both Original Institutional Economics and Keynesian Economics, as alluded to in Section I.3. It is almost impossible to say *when* Simon became “mainstream”, but a convincing argument can be made that it was somewhere in the 1970s, for his Nobel was in 1978. This is a notable time frame, for the 1970s was a notable period of methodological monism and of a profound dominance of orthodoxy over other approaches⁵⁷. As such, it is significant to note that Simon became part of the mainstream in a moment of reconstruction of neoclassical economics.

the 20th century. Other countries without a strong neoclassical tradition could be cited, like Russia (ex-USSR), the ex-Yugoslavian countries, or China. Brazil is another example of a place where ‘orthodoxy’ does not describe the ‘mainstream’.

⁵⁵ It is widely known that Alfred Nobel did not include ‘economics’ as part of the Nobel Prize. The “Nobel Prize of economics”, therefore, is not handled by the Nobel Foundation but the Sweden Central Bank. Therefore, the proper name to this Prize is *Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel*. For simplicity’s sake, we followed the convention to just refer to it as “Nobel Prize in economics”.

⁵⁶ Excluding the Nobel, these prizes were given, respectively, by the following organizations: The Association for Computing Machinery, The USA Government, The American Psychological Association, Institute for Operations Research and Management Science and American Political Science Association.

Source: amturing.acm.org/award_winners/simon_1031467.cfm

⁵⁷ An illustration of the neoclassical methodology can be gleaned in Gary Becker’s *The Economic Approach to Human Behavior*, in 1976. In this book, Becker (1976) proposes that “the economic approach provides a

This was not the case at the end of Simon's career, as the 1990s and 2000s observed a pluralist mainstream (Dequech, 2007). If this assessment holds up is a matter of debate. Recent events, such as the 2008 financial crisis, did little to change the dominant status of orthodoxy, and, in fact, a resurgence of neoclassical economics seemed to take place after the crash (Mäki, 2021; Dow, 2021). This gives credence to the interpretation presented by Bianchi (2010), who proposed that the "crisis could delay the return of orthodoxy and extend the duration of this moment of pluralism" (p. 12, own translation) but could not change its inertia away from pluralism. This is the "message of John Davis"⁵⁸ (Bianchi, 2010, p. 9, own translation): "To us, pluralists, [this message] can be resumed as such: THINGS WILL GET WORSE!" (*ibid.*, adapted).

Orthodoxy, neoclassicism and the negative definition of heterodoxy

In this dissertation, the term 'orthodoxy' is closely aligned with the formulation proposed by Dequech (2007). Adopting the conception of Colander, Holt, and Rosser (2004), which defines orthodoxy as "generally [...] what historians of economic thought have classified as the most recent dominant 'school of thought'" (2004, p. 490). The most recent dominant school of thought was neoclassical economics. According to Colander, Holt, and Rosser (2004) and Dequech (2007), neoclassical economics shares three main characteristics: 1) focus on maximizing rationality; 2) focus on equilibrium analysis; and 3) negligence of "strong" or "fundamental" uncertainties (as defined in Dequech, 2011)⁵⁹.

Besides these three established conceptions, this research also identifies a fourth central element in neoclassicism: methodological individualism. This method hearkens back to the

valuable unified framework for understanding all human behavior" (p. 14), based on traditional microeconomics and rational behavior.

⁵⁸ The text in question is Davis (2008).

⁵⁹ Weak uncertainties are those in which the distribution of probability is known (Dequech, 2011). Some examples are the Knightian Risk of Savage's uncertainty. Strong uncertainties are those in which the distribution of probability is not, or cannot be known. Consequently, the possible "world states" are not perceivable *ex ante*. The main example is the Keynesian uncertainty.

work of Max Weber and can be defined as the idea that “social phenomena must be explained by showing how they result from individual actions, which in turn must be explained through reference to the intentional states that motivate the individual actors”⁶⁰. Consequently, “neoclassicism” in this dissertation will be defined as economic approaches with these four analytical characteristics: 1) maximizing rationality; 2) equilibrium; 3) weak uncertainties; and 4) methodological individualism.

The definition of ‘heterodoxy’ is trickier, however. Dequech (2007) conceptualizes two definitions for it, one negative and the other positive. The negative definition is simply the opposite of either ‘orthodoxy’ (what is not orthodoxy) or ‘mainstream economics’ (what is not mainstream). Keeping in mind that ‘orthodoxy’ and ‘mainstream’ are closely linked in the current period, every economic approach that does not adhere to the four principles listed above is heterodox in this classification – heterodoxy is the complementary set $(1 - p)$ of either orthodoxy or mainstream (p) in this interpretation.

As suggested above, there were periods in which approaches traditionally considered heterodox were part of the mainstream. This weakens significantly the definition of ‘heterodoxy’ as the opposite of ‘mainstream’, for it implies that one cannot be both at the same time. Consequently, we will not consider the definition of ‘heterodoxy’ as the negative of mainstream. The concept of ‘negative heterodoxy’ in this dissertation, therefore, signifies only the negative of orthodoxy – economic approaches that do not share the four points. This classification, however, should be seen as it is: a simplified and imperfect grouping, which is useful only to identify broad trends and with limited application in some cases. As such, this classification serves more as a tour guide than an anatomical one.

The positive definition of heterodoxy

This dissertation will also propose another definition of heterodoxy. The reason for this decision is simple: there are common characteristics between heterodox approaches that do not solely rely on their contrast to orthodox economics. The main limitation of the negative

⁶⁰ Source: plato.stanford.edu/entries/methodological-individualism/

definition is that these characteristics are systematically disregarded. For instance, the negative definition identifies both Austrian School Economics and Marxist Economics as heterodox approaches, with little differentiation. The tricky part is to identify the characteristics which unify these very different strands of economics – not a trivial task!

To positively define something means to identify some set of characteristics present in every element. As Dequech (2007) notes, this might not be possible in the case of heterodoxy: “in some cases, the result of the application is an empty set” (p. 296). We encountered something similar in Section I.1, when defining ‘artificial’ and ‘natural’. In that case, ‘artificial’ was defined as things that originate from human action – i.e. things that do not occur in nature. This is a negative definition, for the main characteristic of ‘artificial’ is not being natural. This negative definition led to the conclusion that no true limit existed between artificiality and nature, a step in the line-blurring process pointed out by Mirowski (2002). Consequently, the result was an “empty set”.

However, we argue that the positive definition of heterodoxy does not necessarily imply an empty set. As argues Sposito (2024), “Positive sociological identification has made its most important advances in the work of Frederic S. Lee” (p. 26) by highlighting criteria of association and community history⁶¹. According to Sposito (2024), the work of Lee (2009) identified three fundamental elements in the identification of a positive heterodoxy: 1) “common values and objects of study” (Sposito, 2024, p. 32); 2) “a significant presence of heterodox economists in more than one association” (*ibid*) and the transit of researchers between these associations; and 3) “the presence of strong convergences between different approaches” (*ibid*). Therefore, while the negative approach signals a “common enemy” (the term is Sposito’s, 2024, p. 25), the positive definition “highlights institutionally shared aspects” (*ibid*).

This definition highlights why the “translations” of heterodox concepts into orthodox ones are often rejected by heterodox economists. Some examples are the adaptation of the Keynesian general theory to neoclassical language in the IS-LM model of Hicks (1937) or the appropriation of Commons’ (1990 [1934]) concept of ‘transaction’ by New Institutional

⁶¹ One of the main arguments for this approach is the associations formed by self-identified heterodox economists as offshoot associations of the American Economic Association. Some examples are the Association for Evolutionary Economics (AFEE, mainly linked to Original Institutional Economics) and the Union for Radical Political Economics (URPE, mainly linked to Marxism).

Economics (NIE) such as in Williamson (1989). These “translations” prompted reactions from traditional Keynesians or Institutionalists due to the fact that they change the ‘common values’ and ‘objects of study’ of those two approaches⁶². This rejection of the “translation” also helps to explain why heterodox Keynesians organized in the ‘Post-Keynesian’ movement and heterodox Institutionalists organized in the ‘Original Institutionalism’ movement (of which the foundation of the Association of Evolutionary Economics was part of) – this was a defense strategy.

In short, heterodoxy can be defined as a loose group of research programs that share 1) common values and themes; 2) associations; and 3) concepts. While loosely defined, this positive definition highlights correctly the fact that traditional (positive) heterodox approaches (Marxist, Keynesian, Institutionalist) share traits that are not found on negative approaches (New Institutionalism, Austrian, New Keynesian, etc.), such as a distrust of the self-regulating nature of markets (lack of equilibrium), a “lack of faith” in human rationality alone, a focus on factors outside economics, and the central discussion of ‘social provisioning’ instead of ‘market efficiency’ as the core of economic investigation.

Applying the definitions to Herbert Simon

Now, let us apply the negative and positive concepts of ‘heterodoxy’ in the work of Herbert Simon. Applying the negative definition, we conclude that Simon is, indeed, a heterodox economist, in virtue of his not conforming to the four elements presented as characteristic of neoclassicism (focus on maximizing rationality, equilibrium, neglect of uncertainty, and methodological individualism). The only characteristic which we can safely say is present in his work is methodological individualism. Much like traditional microeconomics, individual agents as well as organizations are individualized, and outcomes

⁶² Some authors, like Langlois (1986), even argue that Original Institutionalism (deemed by him as “old institutional economics”) is not even the basis in which the New Institutional Economics developed. Instead, he argues that the basis of NIE is the work of Austrian-school Karl Menger. However, lack of conclusive historical evidence weakens this argument.

are defined by the aggregation of individual decisions⁶³. Therefore, the characteristics of individual choice translate to aggregate choices – no new behavior properties are generated when agents are congregated. Therefore, even when Simon talked about organizations, he did not focus on the specific characteristics that behavior assumes when performed in groups, with no important fundamental changes⁶⁴.

The other points either describe partially or do not describe the work of Simon. He focused his work in “rationality”, but not the maximizing kind, even if he made concessions to it (sections I.2 and I.3 dealt with this matter in more detail). In this matter, it is sufficient to say that Simon’s Bounded Rationality relates to the “practical impossibility of the exercise of full rationality” (Barros, 2007, p. 6, own translation).

This also relates to Simon’s interpretation of uncertainty, which he dealt with marginally. As argues Dequech (2001): “The theory of Bounded Rationality overemphasizes rule following” (p. 912) and serves, mostly, as a theory of how agents deal with complexity. While Dequech (2001) is correct in pointing out that “Simon’s emphasis on complexity and neglect of novelty and surprise imply that his concept of Bounded Rationality has limited applicability in situations of fundamental uncertainty” (p. 921), he does not bring attention to the fact that “rule following” itself is interpreted by Simon as a source of novelty.

The key to Bounded Rationality in fundamentally uncertain situations is the ever-central concept of ‘satisficing’ (refer to I.2). As Simon (1990a) observes, satisficing is sufficient for a theoretical agent to make decisions when “an enormous, or even potentially infinite, number of alternatives are to be compared” (p. 9) or when “the problem has so little known structure that all alternatives would have to be examined in order to determine which is optimal” (p. 9-10). In fact, the ‘search method’ described in I.2 is partially able to exist due to some level of uncertainty, which, granted, does not need to be fundamental or strong but can be. Furthermore,

⁶³ This will be important latter, as Tomer (2007) argues that Simon does not start from methodological individualism, an assessment with which we disagree.

⁶⁴ This is in opposition to what we can loosely call ‘methodological collectivism’, such as Marxism, in which a class acts as a social agent. Marx (2015 [1867]), as other heterodox economists such as Keynes (2017 [1936]), Veblen (2017 [1899]) or Commons (1990 [1934]) recognize the ‘composition fallacy’ which presupposes that “that parts or members of a whole will have the same properties as the whole”. Source: www.scribbr.com/fallacies/fallacy-of-composition

the centrality of adaptability can be interpreted as an important decision mechanism for conditions of uncertainty.

In any case, Simon did not focus his studies on matters of radical or fundamental uncertainty, but it is also not correct to say that he “neglected” fundamental uncertainties. Simon never denied the theoretical importance of uncertainty and the need to make intelligent systems that can react accordingly to it – even if, again, he himself did little to tackle these questions directly. The concession that we should make on this regard is that rule-following in line with Bounded Rationality can be implemented to deal with behavior under fundamental uncertainty⁶⁵.

Finally, equilibrium exists in a similar place to the fundamental uncertainties – it is not the focus of Simon. His focus on the ‘behavior’ part of economic behavior relegated much of traditional preoccupations of economics as secondary, as argued in I.4. In summation, by the negative definitions, Simon is not a neoclassical economist and, therefore, is not orthodox.

However, the positive definition of heterodoxy that we proposed gives us another answer. Let us apply the three categories (common values, heterodox association, and convergence with heterodoxy) to Simon. Due to the secondary nature that many economic questions take in Simon’s work, it is almost impossible to say if Simon himself shared the “heterodox values” of which Lee (2009) talked about. It is safe to say, however, that Simon did not focus on them. His preoccupations were not with social matters, power dynamics, or institutional evolution. We must remember Augier (2000) – he was a monomaniac of behavior. In any case, it is safe to say that the science of Herbert Simon was not explicitly in line with heterodox values or object matters.

Pertaining to association, this research could not find any evidence of Simon participating with heterodox associations. While his time at orthodoxy-oriented associations are well documented, such as Carnegie-Mellon and the RAND Institute, the same cannot be said for any heterodox association – let alone about Simon’s transit between heterodox associations. Due to the limited scope of this research, it is not possible to rule out Simon’s association with

⁶⁵ See Nelson and Winter (1985) for an example of Bounded Rationality in conditions of uncertainty.

heterodox organizations; however, we can safely say that it was not a central part of Simon's intellectual or professional development.

The last point, convergences, presents the most promising case to present Simon as positively heterodox: Simon converges with heterodox traditions by taking (some) distance from Substantive Rationality and, consequently, approaches heterodoxy in the quality of a critic. However, he does not converge with heterodox theories of behavior, such as the habits of Original Institutionalism, the 'animal spirits' of Keynes, or the Marxian dialectics of social class. Therefore, the approach between Simon and traditional heterodoxy occurs due to the rejection of Substantive Rationality – but not any specific theory of behavior. Simon also was very fond of “translating” his basic framework of Bounded Rationality to other domains, which latter developed into the many conceptions of Bounded Rationality that now exist (this will be presented in Chapter II). As such, we argue that Simon is not a heterodox economist by the positive definition either.

Consequently, Simon exists in a curious place – he is negatively heterodox but not positively so. This helps to explain the many divergences between the place of Herbert Simon in economics. Economists like Sent (2005), Tomer (2007), and Earl (2010) paint Simon mainly as a critic of neoclassicism and Bounded Rationality as mainly a heterodox conception, while Conlisk (1996) and Dequech (2011) characterize him closely to orthodoxy, somewhere along the *protective belt* of neoclassicism.

Unable to settle in only one characterization, we adopt this double diagnosis: Simon is negatively heterodox, but not positively so. This “superposition”, we believe, is more useful to understand Simon's place in economics than any closed placement as either ‘orthodox’ or ‘heterodox’. This double diagnosis, in fact, helps to explain the acceptance of Bounded Rationality into both orthodoxy and heterodoxy.

I.6 – The Hardcore of Simon's Bounded Rationality

Throughout this chapter, we identified a series of characteristics of Herbert Simon's science, often overlooked in specialized literature. This final section attempts to draw the

characteristics of the scientific “hardcore”⁶⁶ of Herbert Simon, based on the findings laid out in the other sections. As we understand, Simon’s interpretation of Bounded Rationality is consistent with the elements presented throughout this dissertation, even if it was not an exhaustive list. Therefore, this characterization of Simon’s hardcore is functional – it serves as a basis of comparison between other strands of economics that adopted the concept of Bounded Rationality, which is relevant in the analysis of Chapters II and III.

As argued in Sections I.1 to I.3, Simon’s worldview was heavily influenced by computer sciences, especially the conception of the “imitation game” of Alan Turing. To him, the simulation of intelligence consisted in creating artificial systems that make decisions comparable to those made by biological systems. The most important characteristic of any intelligent system is its symbol-manipulation potential. If symbols are manipulated in a manner consistent with what we consider “intelligent”, then we can say that a system is intelligent.

Rationality is a subsection of intelligence, defined as the interaction between agents and goals. A decision is called ‘rational’ if the distance between the desired goal and the actual result can be considered ‘satisfactory’. A decision becomes more rational the closer the desired outcome is to the actual outcome. Simon, therefore, comes closer to the definition of ‘rationality’ formalized by Viskovatoff (2001) – “valid inference”. Instead of defining rationality as ‘optimization’, Viskovatoff proposes that rationality consists of the correct correlation between causes and effects, mediated by the values of the decision maker. Simon’s definition of rationality is much closer to this than any form of optimization – which he considered impossible. As such, decision mechanisms in Simon are many and context-sensitive.

Decisions in Herbert Simon are flexible and adaptable, as are the definitions used by the author, but the decision structure based on computer sciences is consistent. This basic mechanism of decision-making is the beating heart of Herbert Simon’s theory and its core as a scientific research program. However, this ‘monomania’ about decision-making turns Simon’s

⁶⁶ Note that this use is not analogous to that of Lakatos (1968), as only one scientist cannot compose a research program by himself. We chose to keep the term “hardcore” for it communicates the objectives of this research – to identify the main points of a scientific approach. The use of other terms (such as “fundamentals”) wasn’t as successful in communicating this objective.

position opaque to many traditional economic matters, including the author's location in the economic traditions. Table I present a summary of the findings:

Table II – The Hardcore of Hebert Simon

Characteristic	Description
The “Natural” and the “Artificial” are on the same ontological level.	As intelligent systems, organisms, “cyber-organisms” and (thinking) machines work in analogous ways (sections I.1 and I.2; main references: Simon, 1969; Simon, 1990a; Simon, 1983; Simon, 1995).
Focus on limited “problem-solving” and “decision-making”.	Decisions, the result of goal-oriented behavior, is the main theme of Herbert Simon’s work. Intelligent systems share traits which make them ‘rational’ (acting towards a goal), but limited so (bounded). Decision-making, consequently, is context specific (Section I.2; main references: Simon, 1956; Simon, 1957b; Simon, 1990b).
Decision-making “relativism”	As decision behavior is goal-oriented, the degree of ‘rationality’ of each decision is defined contextually. A problem space is formed with the agent’s limitation, the task environment, and a set of possible solutions (Section I.2; main references: Simon, 1976).
“Empirical” research methods, including simulations.	Simon considered non-deterministic phenomena, such as simulations, as “empirical”, as processes independent of human agency generate these results (Section I.3; main references: Simon, 1969; Simon, 1995; Simon, 1997 [1982]]; Simon (2012 [1977])).
Relates with other sciences only in matters of decision-making.	Simon’s hardcore is strictly linked with <i>decision-making</i> . Comments on other matters, such as equilibrium economics or policy, tend to be tangential. Economic matters tend to be stylized as specific ‘problem spaces’ (Section I.4; main references: Simon, 1957a; Simon, 1978; Simon, 1990b).
Adaptable conceptual definitions.	Besides the general outlines of Bounded Rationality, Simon did little to impose <i>his</i> view on what Bounded Rationality should be. This made Bounded Rationality a highly adaptable concept, which was applied in a wide spectrum of economics but made it impossible to classify Simon in the wider economic traditions (Section I.5; main references ⁶⁷ : Simon, 1978; Simon, 1991; Simon, 1996 [1991]).

Source: Own elaboration.

Therefore, Simon proposes a theory of adaptive decision-making inspired by computer sciences that does not differentiate artificial intelligence beings from organic ones. Besides the specific methodological elements (as in research methods), Simon does not compose a general theory of decision-making. In fact, Simon did little to synthesize the elements of his specific

⁶⁷ The quote from Simon in Gigerenzer (2004a) – “I have never thought of either Bounded Rationality or satisficing as precisely defined technical terms” (p. 10) – is also a strong evidence of this item.

methodological approach, leaving it to posterity. Instead, Simon proposes many applications of his specific style of decision-making. As Loasby (1989) summarizes:

“It is perhaps the most remarkable feature of Simon’s work that it appears to be the product of a man who is prepared to face the unknown without a general theory, but with faith that, despite the bounds of their rationality, human beings can display enough flexibility and imagination to recognize and to solve their problems fast enough to maintain, and even improve their situation” (p. 154).

Therefore, more than a *theory* of Bounded Rationality – or even *theories* of Bounded Rationality – Simon’s legacy for economics is mostly a robust framework of decision-making. By all means, this is a *Tabula Rasa*, in which many different theories can be analyzed. In this interpretation, Simon suffered a deeper process of ‘death of author’ than most economists did, as his concepts were reinterpreted in many different ways. The author interfered little with the many applications of Bounded Rationality, even if it is suggested that he disagreed with some of them. Consequently, the interpretations and applications of Bounded Rationality multiplied and gained “life”.

This process culminated in the many current interpretations of Bounded Rationality and a gradual detachment between Herbert Simon and Bounded Rationality. This is not to suggest that Simon was forgotten – but these interpretations often resemble Simon’s work only marginally. Therefore, even if many economists adopted his idea, no “true” heir of Simon – one who continues his methods and research focuses – could be identified in this research, as it will be shown in chapters II and III. This is baffling considering the influence of Herbert Simon in economics, but it starts to make sense when considering his style of science.

This is not also to suggest that the many interpretations of Bounded Rationality are a “corrupted inheritance⁶⁸” of Simon’s original work. In fact, Simon often saw the many

⁶⁸ Pessali (1998) uses the term “corrupted inheritance” (own translation) as an indication of a concept that loses its original meaning and implication when “translated” to other circumstances. The author uses it to refer to the use of the term “transactions” in the work of John R. Commons and his adaptation by Oliver Williamson of the same concept. According to Pessali (1998), Williamson “corrupted” Commons interpretation by stripping the concept of “transactions” of its implications. Here, we use the term in the same manner, but, we argue, this is not the case with Simon. Many implications remain the same, even if specifics are different.

applications of his concept in a positive manner (Fiori, 2011), in line with the idea that many ‘Simons’ exist. In a manner of speaking, the collective interpretation of Herbert Simon *became* the contribution of Herbert Simon, even when the specific Bounded Rationality used by one author or another does not adhere to any elements of his “hardcore”.

CONCLUSION OF CHAPTER I

What is the legacy of this multi-faceted scientist of economic? The answer to this question is simple and obvious: Bounded Rationality! But what *version* of Bounded Rationality Simon proposed? This first chapter aimed to understand this question. We argued that even if Simon did not unify his many theories of Bounded Rationality, it is not strictly correct to call Bounded Rationality only the inverse of optimizing behavior. Bounded rationality is this indeed, but not *only* it. A positive characterization can be made, even if a broad theory of Bounded Rationality still eludes the economic sciences.

Moreover, this chapter intended to show the complex interrelation between the intellectual development of Simon and his method of science. Inspired by an idiosyncratic interpretation of logical positivism, computer sciences, and a wide group of economic approaches of the 1930s, Simon evolved to be the only member of his species. Inspired by the rigor of natural sciences, Simon was formed as a political scientist, made his post-graduation studies in economics, was deeply interested in computer technology (incorporating it into his specific science), which was latter applied to military research in the RAND institute – and all this while feeding a permanent monomania about decision-making.

In this light, it is not surprising that Simon escapes almost every traditional label. In what category did the economics of Herbert Simon fit? Was he even an economist? If so, he was an orthodox or heterodox one? This difficulty stems from a very personal mind and very particular methods. Simon could be unique – but he was not isolated. In fact, it was the contrary, Simon found a way to establish a deep and meaningful dialogue with the economic sciences of his time, which resonates to this day – despite his very uncommon methods. The remaining of this dissertation aims to collect evidences from this resonance.

CHAPTER II

3. Bounded Rationality After Herbert Simon:

Three interpretations on the limits of cognition and what unifies them.

“More is lost by indecision than wrong decision”

The Sopranos, 2007.

Simon left a powerful and evocative concept for the scientific community – Bounded Rationality. This is the crowing jewel of his legacy. In the previous chapter, we proposed an interpretation of Simon’s version of this concept and its place in the economics of his time. It is only natural that such an idea would be adapted, tweaked, and transformed by the many scientists who applied it in their research – and this is exactly what happened. In this chapter, we examine the new meanings that Bounded Rationality acquired after Simon.

The reader, however, may already have in mind the fact that Simon did not make many strides to clearly define or present his specific interpretation of ‘Bounded Rationality’. In fact, he proposed many different formulations of it – but this does not mean that he lacked a cohesive idea that guided them. We argued that the reason for his many formulations lies in the fact that his specific interpretation was “adaptive”, guided by general rules but not strictly determined by them. This relativistic approach to behavior led Bounded Rationality to be interpreted mainly as a “negative concept”, as proposed by Barros (2007). However, Bounded Rationality is not an “empty box”, nor it is a strict formula for behavior. This latitude of interpretation left many gaps in the theories: What are the implications of this type of rationality for equilibrium? Is this a heterodox behavioral theory? Does this invalidate the mainstream approach of Substantive Rationality?

With the wide adoption of Bounded Rationality, the gaps gradually began to be filled. As such, it becomes relevant to identify how Bounded Rationality changed in its many adoptions in economics – and this is exactly what we attempt to do here. The chosen methodology was a review of the different definitions and classifications of both Bounded

Rationality and Behavioral Economics (or BE, henceforth). As the development of Bounded Rationality and BE is intimately tied, it is not possible to analyze them separately.

Three interpretations are selected as a sample: the classification of eight strands of Behavioral Economics proposed by Tomer (2007) (later expanded to ten by de Castro, 2014); the differentiation between ‘Old’ or ‘Original’ and ‘New’ Behavioral Economics, as proposed by Sent (2004), Nagatsu (2015) and Angner (2019); and the three uses of Bounded Rationality identified by Gigerenzer (2020). These three contributions allow us to elucidate, respectively, the specialized strands of BE that formed throughout the years (providing insight into their use of Bounded Rationality), its chronological evolution and important milestones, and the concepts used in the current-day discussion on the matter.

The objective of this analysis, similar to that of Chapter I, is to assess the attempts to understand the “hardcore” of the economic research programs that use Bounded Rationality. Simon’s hardcore (or at least the approximation of it, identified in the last chapter) functions as a basis of analysis, in which the other strands will be compared. Does Simon’s research preoccupations coincide with the current research programs that adopted his ideas? The answer, as the reader may already infer, is no. Bounded Rationality acquire many different meanings through the years. But exactly how Simon’s Bounded Rationality differ from those others? This is what this chapter aims to explore.

This chapter is organized as follows: Section II.1 presents a summary of Simon’s Bounded Rationality, focusing on the theoretical construction of Bounded Rationality, an important matter in the attempted comparison. Sections II.2 to II.4 present the three classifications of “Behavioral Economics” or “Bounded Rationality”⁶⁹. Section II.2 presents Sent’s (2004), Nagatsu’s (2015) and Angner’s (2019) approach, a chronological differentiation

⁶⁹ For simplicity’s sake, we adopt the definition of BE as “economics with Bounded Rationality”. This use of the term is explicitly consistent with Tomer (2007) and Angner (2019) and does not contradict the others definitions, facilitating the arguments presented. This definition, however, has some notable limitations. Firstly, it includes approaches that only marginally deal with Bounded Rationality. Secondly, it excludes the application of other strands of psychology other than the Cognitive one, which is the foundation of Bounded Rationality. In any case, this characterization is useful in this research, for it directly correlates “Bounded Rationality” and Behavioral Economics, avoiding the more nebulous characterization of BE as the “reunification of psychology and economics” proposed by Camerer (1999).

between ‘Old BE’ of Herbert Simon, George Katona, and Tibor Scitovsky and the ‘New BE’, after the works of Kahneman and Tversky. Section II.3 discusses Tomer’s (2007) classification, based on six criteria, focusing on specific methodological differences between economic strands that use Bounded Rationality. Section II.4 presents the classification of Gerd Gigerenzer, which differentiates three broad methodological categories of Bounded Rationality, distinct by their methodological applications.

Finally, Section II.5 proposes a “summary” of this review, and an interpretation of what unifies the different formulations of Bounded Rationality. Similar to the condensation of Simon’s characteristics in his ‘hardcore’ in Chapter I, the classification proposed in this section is a condensation of the many characteristics of Bounded Rationality, shared by the different definitions presented here. The main crux of this argument is the “axiomatic” quality of the Bounded Rationality, its ‘self-justifying nature’. As such, the term ‘Bounded Rationality’ congregates a number of implications, even if the concepts themselves do not.

II.1 – How to Reach “Bounded Rationality”?

The basic idea of Bounded Rationality can be explained in one single sentence: the imposition of limits on the capacities of people to get what they want. Section I.2 presented the basic structure of Bounded Rationality, while Section I.4 presented its critique of Substantive Rationality. However, we made an omission in these two instances. How exactly does one “reach” Bounded Rationality? In other words, what methodological steps need to be taken for a theory of Bounded Rationality to be created? Notice that in Section I.2 the focus was not on the “imposition of limits”, but the functioning of this type of rationality once they were already in place. As such, we assumed that rationality was already bounded. So exactly what needs to happen to an agent’s cognition for it to be “*boundedly rational*”?

This may seem a secondary question, but it is not. Satisficing presupposes a comparison between one actual state of the world and a desired state – but how can we know this desired state? We can infer that it depends on the agent’s proclivities or the task at hand, but this does little to explain behavior. To end the explanation of the ‘desired state’ on “it depends” creates a perfectly fitting criteria that explains all possible situations – even non-nonsensical or bizarre behavior. Let us take the examples of Stigler (1966 *apud* de Castro, 2014): “We see a person

drinking the crankcase oil of his car and, in sequence, this person contorts in agony and dies” (de Castro, 2014, p. 99, own translation). Adopting a microeconomic framework, can we say that this person revealed a preference for motor oil and, for this reason, chose to drink it? What was the ‘desired state’ in this case? His optimal solution was to drink even more oil on other occasions, but his Bounded Rationality only let him take the “satisficing” option of drinking it only once? Can we find logic in such illogical behavior?

Satisficing starts to show its cracks as a behavioral theory in cases like this. In this preposterous situation, what was the goal that guided the decision process? We probably could say that someone who drinks motor oil is “limitedly” rational – for the simple reason that a rational person (in the definition of Viskovatoff, 2001) would have the insight to understand the consequences of this action and would probably not do it. Then, how can Bounded Rationality help us to explain this case? If we adopt a perfectly adaptive criteria of ‘desired state’, we can start by defining the satisficing criteria as such:

Satisficing criteria: The agent wants to drink motor oil. Satisficing options are a set of outcomes related to this criteria.

This situation also begs another question: the agent *wants* to drink motor oil, but *should* he? This is the Achilles-heel of the perfectly adaptive criteria of the ‘desired state’ – as it’s only related to decisions, no consideration is given to the cultural norms of behavior, the specific values of an agent, biological or technical limitations, or even (on some degree) the consequences of specific actions. In the motor oil case, one can make the convincing argument that there is a ‘rational’ or even ‘optimizing’ solution – not drinking the oil and continuing to live. However, the “value agnosticism” of the decision structure of Bounded Rationality does not take this into consideration. As such, if we take a step back, however, we can see this “satisficing” criteria as what it is – an *ex post* explanation for a behavior that defies any *ex ante* attempt to rationalize it that has no explanatory value at all. In short, a theory of *something* must not explain *everything* – it needs constraints for its explanations to mean anything.

In this case, as in many others, there is an untold expected goal in the mind of the agent. For instance, when searching for a job, we can infer that this agent seeks higher wages; when buying a house, the agent probably will prefer a safer neighborhood; when investing, we can

expect that the agent will seek higher returns. It is very difficult to make the case that a person can be rational when deliberately losing money by investing in Ponzi-scheme ventures.

This reveals an important quirk of Bounded Rationality *à la* Simon. Decision-making is relative to the context, as is the satisficing criteria – but not the desired outcome. In the investment case, the investor may have some extravagant goal for a return-for-profit (possibly tending to infinity), but given the investment option that he has, a satisficing criterion emerges. Again we should stress the importance of the problem space in this case: a goal (income $\rightarrow \infty$, in this case) encounters a limit (maximum income is less than ∞), and a number of solutions present themselves. This is the basis of the “decision-making relativism” presented in Section I.6 – not necessarily a relativity of goals, but a relativity of problem spaces that take shape when goals are confronted by the real world. In this case, the same exact goal (profit maximization) is applied to a number of different circumstances and creates an exactly equal number of satisficing outcomes.

As such, how can we know the agent’s desires that are directly linked to the formation of satisficing criteria? Simon is not clear on this matter – remember, he had no general theory, as suggested by Loasby (1989). For the lack of an original theory and his many applications of Bounded Rationality in economics, an argument can be made for the case that Simon’s main criteria for “desires” tended to be the neoclassical Substantive Rationality or something in this line.

Let us illustrate this point: the basic method of Bounded Rationality is to take theories of decision and impose limits on them – hence the name. This is exactly the path taken in Simon (1990c): “Theories of Bounded Rationality can be generated by relaxing one or more of the assumptions of [Subjective Expected Utility] theory⁷⁰” (p. 15, adapted). In this case, Simon takes the traditional neoclassical microeconomic rationality as a basis of analysis, adding to it limitations of cognition. The next step, taking heed of Stiglitz’s (2010) warning of the fact that perfect and *almost* perfect behavior differs greatly, is to identify how this behavioral theory changes when limitations are applied to it. Neoclassical microeconomics is changed in two ways on this occasion: 1) the optimization solution cannot be attained, and 2) the “stopping rule” must be changed.

⁷⁰ “SEU theory” in the original.

However, as we saw in Chapter I, Simon did not think of Bounded Rationality only in this context – Bounded Rationality can be applied in an infinite number of cases. But when dealing with economics, how can we know what people want? Simon’s silence on the matter presents us with two main alternatives: 1) the decision goal is defined exogenously (Simon, 1978, 1990a, 1990b, 1995, among others) and 2) the decision is defined by neoclassical standards (Simon, 1955, 1976, 1990c, among others). In some instances, like Simon (1969), the author seems to adopt both alternatives, depending on the occasion⁷¹. As such, Bounded Rationality can be used to study decisions made with a “personalized desired state”, or one prescribed by neoclassical theory, one specific type of desired state. Curiously, therefore, Neoclassical Substantive Rationality acquires the quality of a special case of Bounded Rationality.

With this in mind, can we identify the other cases of Bounded Rationality? In other words, what other theories of behavior would benefit from the application of cognitive limits? The answer is not many. Note that the process of “bounding” rationality only makes sense if it is unbounded at first. For instance, there is no need to “bound” the rationality of the investor described in Keynes’ *General Theory* (2017 [1936]) – this Keynesian agent is fully capable of making mistakes. The same applies to almost every traditional heterodox approach, such as Marx’s (2015 [1867]), Veblen’s (1898), Commons’ (1990 [1934]) or Schumpeter’s (2021 [1911]). None of these authors, or the traditions that they inaugurate, presuppose an unbounded agent needing to know their limits – only neoclassicism does this. In this case, the “special case” dynamic is inverted – Bounded Rationality is a special case of Substantive Rationality.

In any occasion, we can infer that Bounded Rationality and Substantive Rationality are intimately related. In fact, without heavy contextualization, it is difficult to think of a Bounded Rationality untied from Substantive Rationality, at least in the sense of conceptual development. Theories of Bounded Rationality do not require that the theory in question accept Substantive Behavior⁷², but the concept of Bounded Rationality does.

This again is not to say that Simon was a neoclassical economist – as argued in Section I.5, this classification is not trivial. The fact that we should note is that Simon never made a systematic use of an original normative theory of “desired states”. While he also did not

⁷¹ There are no “phases” in which Simon favored one explanation or another. Both were concomitant.

⁷² Arguably, Nelson and Winter (1985) and Davidson (2010) are examples of this case.

systematize all the uses that he made of the mechanisms of Bounded Rationality, an independent “hardcore” can be constructed analyzing how Simon applied this theory and how he explained behavior. However, no analogous effort can be made to define the normative standards of behavior for Simon. He tends to be vague about it or relegates it to either the model-maker (in the case of AI) or the economics discipline (in the case of economics). A good theory of novel writing will write novels, but Simon’s monomania is on the writer, not the writing.

As such, Bounded Rationality in Simon can be understood in two ways: 1) the desired state is determined outside the model (context-specific, defined by the researcher, stated by the agent, etc.) or 2) dependent on standards defined by Substantive Rationality. This research could not identify a solid case to determine if Simon favored one or the other. Literature that argues for the first tends to paint Simon as a subversive (Sent, 2005; Tomer, 2007; Angner, 2019; etc.), a reading that is contradicted by Simon’s insertion in the mainstream. Literature that argues for a close adherence to Simon into the mainstream (Conlisk, 1996; Loasby, 1996; Dequech, 2001; etc.) tends to paint Bounded Rationality as a stricter concept than it is in Simon’s work. Both lines tend to ignore the broad work of Simon in other fields, such as AI. In any occasion, theories that adopt Bounded Rationality will either presuppose that Simon followed the normative theory of orthodox economics or will propose a critique of it and an alternative behavioral theory.

This completes the explanation of Simon’s Bounded Rationality: this is a theory of behavior based on a specific and independent structure (Section I.2) that implies a number of implications to mainstream economics (I.4) but that does not present an independent normative theory (this section). This last point needs either to be implemented exogenously or captured from other theories. In summary, Simon’s Bounded Rationality proposes an independent decision structure, but only when a “desired state” was already defined. Therefore, there are two main ways to “reach” Bounded Rationality: constructing independent behavioral explanations or adapting existing theories. The following classifications fall in one category or another.

II.2 – Two Periods of Behavioral Economics

One of the first and most successful attempts to classify Behavioral Economics is what we refer to in this dissertation as the ‘two periods interpretation’, proposed by Sent (2004), and

later expanded upon by Nagatsu (2015) and Angner (2019). The central thesis is this: BE, while a consolidating field in the 1950s and 1960s, presented concerns and worldviews that were abandoned in favor of a more “modern” economic approach. This shift occurred throughout the 1970s and mainly consisted in an approximation towards the dominant neoclassical mainstream of the time. As such, BE is classified into two broad periods: ‘Old’ or ‘Original’⁷³ and ‘New’ Behavioral Economics. This section is subdivided to facilitate the explanation.

Original Behavioral Economics

The first moment – Original Behavioral Economics – consisted of three main bodies of work: George Katona (Raposo, 1984), Tibor Scitovsky (Bianchi, 2018)⁷⁴ and, of course, Herbert Simon (Sent, 2004; de Castro, 2014). These were very disparate researchers, unified in their dissatisfaction with the behavioral theories of the economic mainstream between the 1950s and 1960s. All of them demonstrated an interest in interdisciplinary methodology, especially psychology: Katona, a psychologist by training, shared something similar to Simon’s ‘monomania’ and focused his interests on behavior itself, with little mention of more common elements of economic analysis⁷⁵. Scitovsky, on the other hand, had more “traditional” economic interests and were dissatisfied with economics’ “lack of humanity” – its joylessness⁷⁶ (Bianchi, 2018). Simon’s requires no further elaboration.

⁷³ ‘Original’ is not a classification adopted by Sent (2004), Nagatsu (2015), or Angner (2019). However, due to the connotations of the word ‘old’ (outdated, aged, etc.), we prefer to use the term ‘original’. This is analogous to the case of ‘Original’ and ‘New Institutional Economics’. Moreover, authors like Gerd Gigerenzer, an active scholar, are closer to ‘Old BE’ than ‘New BE’. For these reasons, we favor the term Original BE.

⁷⁴ This author is Marina Bianchi, not to be confused with Ana Maria Bianchi, mentioned earlier. Ana Maria also explicitly recognized Tibor Scitovsky as a founder of BE.

Source: https://www.youtube.com/watch?v=0i8EWqi9LBA&ab_channel=CanalUSP

⁷⁵ Among Katona’s works are *Organizing and Memorizing* (1940); *Psychological Analysis of Economic Behavior* (1951); *Psychological Economics* (1975), among others.

⁷⁶ Scitovsky was a law graduate. His most impactful book, *The Joyless Economy* (1976). Other works include *Welfare and Competition* (1951); *Economic Theory and Western European Integration* (1958) and *Money and the Balance of Payments* (1969).

The most relevant element that unified these authors was the search for alternative behavioral theories, but not necessarily psychological ones. Authors that interpret BE as an effort to incorporate psychological findings into economics often overlook the other sources of knowledge that BE drew from. Simon is a good example: one of his main inspirations was computer science. In short, BE wasn't *only* looking at psychology when looking for alternatives. Consequently, this search is broader than the common narratives of “applying psychology to economics” or “reunifying psychology and economics” (proposed respectively by Camerer, 1999, and Nagatsu, 2015).

It is also worth noting that Simon's relationship with psychology is not as clear-cut as some might think. Simon had no formal training in psychology and primarily derived his interests in behavioral theories from economics and AI, as argued earlier. It's no wonder, then, that Simon's approach with psychology occurred within Cognitive Psychology⁷⁷, a psychological strand closer to the development of computer models of the mind (Schultz and Schultz, 2011). Aside from occasional nods to other psychological research programs, Simon's effort consists in an application of *Cognitive Psychology* to economics – not a broad attempt to adapt psychological approaches to economics, as some authors seem to suggest. Simon illustrates this by saying, “Psychologists think that I am an economist, but economists think I am a psychologist. In fact, I feel allegiance to none of these academic tribes, but regard myself as a citizen of the world – a behavioral scientist”⁷⁸ (attributed to Simon, in Augier and March, 2003, p. 136). Consequently, the “psychology” of Simon (and other Behavioral Economists) is a preoccupation with the determinants of behavior (cognition, intelligence, rationality) – but not necessarily with the research field of psychology proper.

Moreover, even among the first precursor generations, Simon's interest in computer science is unique. This research's scope did not assess in depth the approaches of Katona or Scitovsky, and thus few comments on both are included here. However, what becomes clear even in this brief review is that both Katona and Scitovsky shared Simon's broad interest in alternative behavioral theories, but not his interest in behavioral simulations or AI. Consequently, Simon's specific methodology of combining economics, cognitive psychology,

⁷⁷ Cognitive Psychology is part of the field known as ‘Cognitive Science’, a loose research orientation based on the study of behavior and the application of computer models. See Miller (2003) for a brief overview.

⁷⁸ This is an excerpt of a letter from Simon to Sigmund Koch, 06/02/1961 (Augier and March, 2003).

and AI remained constrained to himself. Curiously, therefore, Original BE cannot be classified as part of the ‘cyborg sciences’, even if Simon himself, its most relevant author, can.

As such, Original BE can be identified as a discernible, but not cohesive, intellectual movement in economics initiated in the 1950s. The main goal of this movement was to find alternative behavioral theories that could either complement or replace traditional microeconomic utility-maximization theories. This search’s center of inertia is the conviction that neoclassical microeconomics is not an appropriate behavioral theory – it is neither realistic nor particularly explanatory.

New Behavioral Economics

The second moment – New Behavioral Economics – has its beginnings marked by the publication of Kahneman and Tversky’s “blockbusters”⁷⁹: *Judgment under Uncertainty: Heuristics and biases* (Kahneman and Tversky, 1974) and *Prospect Theory: An analysis of decision under risk* (Kahneman and Tversky, 2013 [1979]). These two articles initiated the recognition of BE as a sub-field of economics. Before this, BE was not seen as a larger research program – at this time, what would become BE were simply tweaks or critiques of economic behavior, orbiting other, more holistic, scientific research programs.

Paradoxically, however, New BE remained much closer to neoclassical economics, even if a discernible research program was beginning to take shape. The main goal of this new research effort is to increase the “explanatory power” (Thaler, 1980; 2016) of traditional economic theory. Thaler (1980) illustrates this point. According to him: “[...] in certain well-defined situations many consumers act in a manner that is inconsistent with economic theory. In these situations economic theory will make systematic errors in predicting behavior” (p. 39). New BE, therefore, exists in these crevices where the light of neoclassicism does not (yet) shine as bright.

In general terms, the methodology of New BE does not reflect Simon’s approach described in Chapter I. Instead of proposing a different behavioral structure, New BE imports

⁷⁹ The term is Thaler’s, 2016.

much of the traditional microeconomic framework. Note, for instance, that “systematic errors” only make sense if a normative theory exists and defines “systematic successes”. Errors are determined by the distance between actual and desired outcomes – a level of satisfaction, in other words. Therefore, New BE presupposes a normative theory that Simon never developed (refer to the previous section). New BE authors like Thaler (1980) are explicit on this matter, whereas Kahneman and Tversky (1974; 2013 [1979]) leave it implied.

As such, some authors (Earl, 2010; Gigerenzer and Berg, 2010; Dequech, 2011) put forth the idea that New BE exists within the protective belt of neoclassical economics. As such, this effort is more an attempt to disguise the same old theory: “neoclassicism with psychological names”, as Gigerenzer and Berg (2010, p. 140) suggest. While there is little doubt that a significant shift occurred, what remains in dispute is its meaning. Thaler (2016) interprets New BE as a bold new frontier for economics, whereas Gigerenzer and Berg (2010) suggest that it is a distortion of the original aims of Bounded Rationality.

Angner (2019), who leans closer to Thaler’s interpretation, goes even further and proposes that a “consummate conversion of neoclassical economists into behavioral ones” (p. 205) followed the 1970s shift. According to him, mainstream economics no longer reflects traditional rational neoclassical theories (maximizing behavior, perfect information processing, local satiation, etc.). The foundations of the new mainstream, therefore, are derived from BE – Bounded Rationality, for short. In parallel to Friedman’s famous comment – “We’re all Keynesians now” (attributed to Milton Friedman in Angner, 2019, p. 205) – Angner proposes that “We’re all behavioral economists now” (Angner, 2019, p. 205).

We should note, however, that Angner (2019) is quite specific with what he means by ‘behavioral economics’. According to him, mainstream economics follows not the whole of BE – Original and New – but specifically New BE. Referring to Original BE, he argues that “it is a mystery how neoclassical and behavioral economists could come to disagree so vigorously for so long” (p. 197) and emphasizes that New BE is much less strict on this matter: “neoclassical and behavioral economists nowadays often go out of their way to downplay the differences (p. 195).”

Note, however, that New BE is much closer to neoclassicism than its counterpart. As such, this ‘consummate conversion’ may not be as dramatic as Angner (2019) seems to suggest – a neoclassical hardcore still remains at its center. Moreover, we should note that Angner’s

classification implies that Original BE was substituted in favor of the new approach – Original BE was a critique of neoclassicism, which is no longer the case with New BE⁸⁰. Further, we will explore the current state of both Original and New BE; for the time being, it suffices to say that no substitution took place: both approaches co-existed in the 1970s and still do today.

Furthermore, the linear evolution implied by Angner (2019) also presupposes a connection that did not necessarily exist during this shift. Plainly, the connection between Herbert Simon and New BE is not clear. Incredibly, the evidence for this common narrative is not conclusive. Gigerenzer (2004; 2020), for instance, brings attention to the fact that Kahneman and Tversky did not cite Simon in his early influential works – nor did they mention the concept of Bounded Rationality there. Lopes (1992 *apud* Gigerenzer, 2004) interprets the few mentions of Simon in the Tversky, Kahneman, and Slovic (1982) anthology as more a recognition of an important figure in the same field rather than evidence of a tributary link to Simon. In short, Kahneman and Tversky made only a ‘nod’ to Simon at first.

As such, Simon did not focus his works on identifying ‘systematic errors’, like Kahneman and Tversky did. Nonetheless, the popular history of BE often identifies him as one of the precursors of the notion that Bounded Rationality implies errors (biases) – but this is not the case. In any occasion, however, this confusion was to last: “Nevertheless, the notion that Bounded Rationality is the study of cognitive illusions has since become widespread” (Gigerenzer, 2004, p. 5). The reason for this misunderstanding is simple: while Simon’s Bounded Rationality made space for systematically erroneous behavior, this was not the norm. Remember, Simon did not lay down a cohesive theory of ‘satisfaction’ that allowed behavior to be inherently “biased”. Kahneman and Tversky, on the other hand, implicitly adopt neoclassical normativity, which allows behavior to be systematically biased (Gigerenzer, 2020). In Simon, as situations change, “erroneous” behavior also change; it becomes more or less appropriate to each problem space; in Kahneman and Tversky, the actions themselves are wrong, independently of context – the error is in the mind.

Moreover, in Simon, errors are an important part of the learning processes that increase the discretionary power of agents. For him, decision errors are not to be completely hammered out of the decision-making process – they serve an adaptive function. Even the concept of

⁸⁰ This, again, is one of the reasons that made us favor the term ‘Original’ instead of ‘Old’.

‘satisficing’ becomes murky if we consider that that error is the norm – how exactly can we be satisfied if we are so often mistaken? As such, Simon does not seem to focus on possible errors of decision-makers, trying instead to contextualize the analyzed decisions. He offers an illustration: If a dog dances, “The marvel is not that it dances well; the marvel is that it dances at all” (Simon, 1995, p. 114, attributed to Samuel Johnson) – the wonder is not that decision-makers choose perfectly – the wonder is that they choose at all⁸¹.

Lopes (1992) and Gigerenzer (2020) help us illuminate why Bounded Rationality, a concept of Simon, became so ingrained with Kahneman and Tversky’s idea of ‘systematic errors’: this association originates from the 1970s shift between New and Original BE. The grouping with Simon, Kahneman and Tversky under the umbrella of ‘Behavioral Economics’, therefore, does not reflect specific methodological convergences regarding Bounded Rationality, but rather a broad convergence.

The key distinction between Original and New BE lies in two important impacts on economics: firstly, it “translated” behavioral cognitive theory into “neoclassical language”, making it possible to apply BE into this framework. The protagonists of New BE also would inaugurate the important Heuristics and Biases (H&B) research program⁸² – the formalization of the prescriptions to look for errors in behavior. This research program would prove to be one of the most successful approaches to BE – if not *the* most successful one.

Therefore, New BE, unlike Original BE, did not seek alternative behavioral theories but aimed to reform neoclassicism. The underlying assumption remained the inadequacies of neoclassical rational theory, but the prescription of ‘what to do’ changed considerably – the broader interest in alternative behavioral theories of Original BE did not translate to New BE. Psychology remained a source of inspiration, but mainly the Cognitive strand, and little else.

⁸¹ We offer another illustration: Anthony Stafford Beer, a researcher in cybernetics, proposed the principle of POSIWID: “The purpose of a system is what it does”. As such, the purpose of a system cannot be its intended function. In a manner of speaking, Simon adopts a version of this principle in his behavioral theory: the purpose of behavior is what it does – not what it intends to do. This is one possible justification for the principle of satisficing. Source: <https://www.forbes.com/sites/benjaminkomlos/2021/09/13/the-purpose-of-a-system-is-what-it-does-not-what-it-claims-to-do/>

⁸² The term is Fiori’s, 2011

No consideration was given to the field of psychology, much less computer science. Slowly, the interdisciplinary nature of BE faltered, and a more endogenous BE started to take shape.

Summary of the two periods interpretation.

The main contribution with the distinction of the ‘two periods’ approach is the correct identification of a significant shift in BE. It, however, presupposes a process of substitution that did not occur. Original BE was not ‘outdated’ and later replaced by New BE. Instead, both approaches coexisted – Simon was still an active and relevant scientist in the 1970s. He, however, seems to be “the last of the Mohicans” of the first generation of Original BE in the mainstream. Katona had some success in establishing its own research program, but this is seldom noticed outside specialized literature. Scitovsky, on the other hand, is a rare sight, even by specialized authors (Sent, 2004 does not cite Scitovsky as one of the founding fathers of BE, for instance).

In Bounded Rationality, the groundwork for a transformation was set. The changing face of BE, therefore, would deeply affect the evolution of this concept. Even if Simon proposed it originally, Kahneman and Tversky (possibly unbeknownst to them at the time) would revolutionize what ‘Bounded Rationality’ meant in economics, as many research programs would spring from this evolving core. The next section briefly deals with some of them.

II.3 – Approaches of Behavioral Economics.

While the ‘two periods’ classification provides a general framework of the evolution of Bounded Rationality, little methodological detail is given. With this in mind, Tomer (2007) probably proposed the more in-depth classification of Behavioral Economics, helping us to identify how Bounded Rationality evolved. To him, BE signifies, simply, the use of theories of Bounded Rationality as an explanation of behavior. Consequently, Tomer includes approaches

not often interpreted as ‘Behavioral’ *per se*, such as Nelson and Winter’s (1985) evolutionary theory and George Akerlof’s (2002) macroeconomics.

Tomer’s (2007) preoccupation is to identify what differentiates mainstream economics from BE. To this end, he adopts six comparative categories: 1) Narrowness, deliberate restrictions from the specific approaches (i.e., limits of the approach’s research scope); 2) Rigidity, the adherence of an approach to his own methods of research; 3) Intolerance, the general *a priori* propensity to consider irrelevant the advances of other approaches; 4) Mechanicalness, the adherence of an approach to the idea that people act in a deterministic manner⁸³; 5) Separateness, the distance from one approach to another; and 6) Individualism, the focus on individual behavior in place of aggregate behavior.

According to him, mainstream economics scores high in every category: it is strait, rigid, intolerant, mechanistic, separated, and individualistic. We wholly agree with this characterization, with one caveat: Tomer (2007) equates ‘mainstream’ with orthodoxy. As such, the pluralistic mainstream of the 2000s does not share these traits, while that of the 1970s does. With this in mind, it is more rigorous to say that orthodox economics presents these six traits.

After presenting his classification of orthodox economics, Tomer (2007) presents his classification of BE. He identifies eight discernible approaches: 1) Herbert Simon’s and the Carnegie-Mellon approach; 2) George Katona and the Michigan School approach; 3) Psychological Economics from Kahneman and Tversky; 4) Harvery Leibenstein’s X-Efficiency theory; 5) George Akerlof’s Behavioral Macroeconomics; 6) Richard Nelson and Sidney Winter’s evolutionary theory; 7) Behavioral Finance; and finally, 8) Vernon Smith’s Experimental Economics. De Castro (2014) develops further Tomer’s classification, adding to it the following approaches: 9) Peter Earl’s Economical Psychology; and 10) Gerd Gigerenzer and the Berlin Group approach to Ecological Rationality. Let us start with Tomer’s classification of Herbert Simon.

⁸³ Tomer (2007) describes mechanicalness as such: “actors are viewed by the discipline as behaving in machine-like ways” (p. 466). This refers to deterministic behavior. Due to the implications of ‘machine behavior’ in Simon’s work (presented in Section I.1) and to avoid confusion, we opted to omit the comparison to machines.

Tomer (2007) classifies Simon on the exact opposite of mainstream economics: Simon is wide, flexible, tolerant, un-mechanical⁸⁴, interconnected to other approaches, and not individualistic. We mostly agree with this classification, but another caveat must be made: this research does consider Simon's approach non-individualistic. Let us examine this closer.

According to Tomer, "Individualism implies that the ultimate constituents of the social world are individuals" (2007, p. 466). While subtle, Tomer does not imply that "Individualism" is an approach's adherence to methodological individualism, but only an implication of it ("Individualism implies...", *ibid*). As such, Tomer's Individualism is closer to the focus of research in individual behavior itself – not methodological individualism proper. Therefore, it is plausible that Tomer's methodology identified Simon as individualistic due to Simon's early works in organizational (aggregate) behavior. As indicated by Brette, Lazaric, and Vieira (2017), however, Simon's focus on organizations would suffer a great shift after 1958, when Simon decidedly focused his research on individual behavior (refer to Section I.3). It is also worthy of note that Tomer (2007) does not cite any first-hand reference of Simon in his classification. The main references to Simon in his paper are Augier (2003) and Augier and March (2003)⁸⁵. As such, another plausible explanation is that Tomer (2007) interpreted Simon's early focus on organizational behavior as part of a greater conviction of "methodological collectivism" – which Simon did not present.

In any occasion, we argue that Tomer (2007) misplaced the role of individualism in Herbert Simon. Simon always adopted a form of methodological individualism, and, therefore, his science is also individualistic, even when dealing with organizations⁸⁶. As a consequence,

⁸⁴ While ironic, Simon's Bounded Rationality is not mechanical, even with its closeness to computer sciences. It can be said that Simon "un-mechanizes" machines.

⁸⁵ This dissertation could not find this article. As such, it was not possible to review Augier's (2003) characterization of Simon. In Augier and March (2003), little mention is made of the individualistic quality of Simon's science, except for general statements such as "Most of his major work was linked to an enduring emphasis on the cognitive processes of intentional action decision-making and problem-solving by individuals and organizations" (p. 137).

⁸⁶ Again, this is analogous to microeconomics – firms are individualized not because they are individuals, but because the firm is interpreted as one single cohesive unit. This form of methodological individualism, Dequech (2007, p. 291-292 *apud* de Castro, 2014, p. 106) argues, was one of the key ingredients for BE to be inserted into mainstream economics.

we agree with Simon's classification in the first five categories, but not the sixth. Table III presents a figure with Tomer's (2007) and de Castro's (2014) classification⁸⁷, along with suggested alterations to the sixth category (Individualism).

⁸⁷ While suggesting the inclusion of two approaches, de Castro (2014) does not classify them in Tomer's (2007) categories. The references used to categorize each approach are specified below Table III.

Table III – Classification of Approaches of Bounded Rationality

Approach	Narrowness	Rigidity	Intolerance	Mechanicalness	Separateness	Individualism
Herbert Simon and the Carnegie-Mellon School	Wide	Flexible	Tolerant	Non-mechanical	Integrated	Individualistic*
George Katona and the Michigan School	Wide	Flexible	Not informed	Non-mechanical	Integrated	Not individualistic
Psychological Economics (Kahneman, Tversky)	Narrow	Rigid	Intolerant	Non-mechanical	Integrated	Individualistic*
Harvery Leibeinstein and the X-efficiency theory.	Narrow	Flexible	Tolerant	Mechanical	Separated	Individualistic*
George Akerloff's Behavioral Macroeconomics	Wide	Flexible	Tolerant	Non-mechanical	Integrated	Not individualistic
Nelson and Winter Evolutionary Theory	Wide	Flexible	Tolerant	Non-mechanical	Integrated	Not individualistic
Behavioral Finance (Shiller, Thaler)	Narrow****	Rigid	Intolerant	Non-mechanical	Integrated	Individualistic*
Vernon Smith and Experimental Economics	Middle point between narrow and wide.	Flexible	Tolerant	Non-mechanical	Middle point between integrated and separated.	Not individualistic
Peter Earl's Economical Psychology**	Wide	Flexible	Tolerant	Non-mechanical	Integrated	Not individualistic
Gerg Gigerenzer and the Berlin Group***	Wide	Flexible	Tolerant	Non-mechanical	Integrated	Not individualistic

Source: own elaboration, based in Tomer (2007) and de Castro (2014).

*Own classification. In Tomer (2007), only mainstream economics (omitted in this table) is considered individualistic. This research interprets these approaches as “individualistic”, instead of the “not-individualistic” characterization of Tomer (2007).

** Based in Earl (2010) and de Castro (2014)

*** Based in Gigerenzer (2004; 2020); Gigerenzer and Gaissmaier (2011); and de Castro (2014).

**** Behavioral Finance is classified as a middle point between narrow and wide in Tomer (2007). We argue, however, that due to the dependency of neoclassicism and the “special case” nature proposed by Thaler (1980), this approach is narrow.

Tomer's (2007) classification reveals some broad tendencies in Behavioral Economics. First, the approaches associated with Original BE, Simon and Katona, share some common traits. Both are wide, flexible, non-mechanical, and integrated. The key difference between the two is their relation to Individualism – we consider Simon individualistic, but not Katona. This is due to a difference of focus. Simon incorporates social behavior only as a stylized form of environment, while Katona recognizes that behavior fundamentally changes when it is collective. In other words, aggregates are not analogous to individuals – they function in another way⁸⁸.

Second, we can identify that both approaches related to the H&B paradigms – Psychological Economics (Daniel Kahneman, Amos Tversky, David Laibson, George Lowenstein, Mathew Rabin [Tomer, 2007]) and Behavioral Finance (Richard Thaler⁸⁹, Robert Shiller, Andrei Shleifer, and Hersch Shefrin [*ibid*]) – share four characteristics with orthodox economics: narrowness, rigidity, intolerance, and individualism. These points of convergence help to explain why these approaches were so successful in the mainstream⁹⁰. The other characteristics (non-mechanicalness and integration with interdisciplinary fields), therefore, can be understood as the research gap that New BE fulfills in the mainstream.

The evolution of Bounded Rationality, therefore, fostered research programs with little or no commonalities among them. Bounded Rationality can be narrow or wide; rigid or flexible; intolerant or not; mechanical or not; separated or integrated; individualistic or not. Moreover, some versions of Bounded Rationality are not reminiscent of Simon's concept. We already presented the case of Kahneman and Tversky, wrongly identified as heirs of Simon's rationality, but other examples can be highlighted. Harvey Leibenstein's, for instance, is much closer to

⁸⁸ A common illustration of this is the 'fallacy of composition': "Inferring that something is true of the whole from the fact that it is true of some part of the whole". Methodological Individualism falls into this fallacy, for it presupposes that the characteristics of individual behavior are analogous to those of collective action.

Source: <https://www.logicallyfallacious.com/logicalfallacies/Fallacy-of-Composition>

⁸⁹ Richard Thaler also is part of the Psychological Economics approach, according to Tomer (2007). For simplicity's sake, we opted to include him only in one category.

⁹⁰ Kahneman won the "Economics Nobel Prize" in 2002, a Prize he would probably share with Amos Tversky, hadn't Tversky not already passed away in 2002 (the Nobel does not grant posthumous prizes). Thaler won the Prize in 2017.

neoclassicism itself, proposing a form of Bounded Rationality in line with orthodox microeconomics. This also is not a form of Bounded Rationality directly tied with Simon.

Therefore, what becomes clear at this point is that Bounded Rationality now means many different concepts. Some of them are reminiscent of or evolved from Simon's original concept – but some did not. At this moment, the identification of a methodological unity of BE becomes effectively impossible: this loose congregation of ideas would coalesce into the current interpretations of 'Bounded Rationality'. Gigerenzer (2020) identifies three types: 'ecological', 'irrational' and 'optimizing under constraints'. The three types are presented in the following section.

II.4 – Types of Bounded Rationality

Tomer's (2007) classification is a significant complement to the 'two periods' interpretation and advances towards a more holistic interpretation of the approaches to Bounded Rationality. However, Tomer still does not identify how Bounded Rationality evolved – only its aftermath, the research program itself. This is deliberate, for his efforts were to interpret BE in comparison with the mainstream, but a question remains: what theory of Bounded Rationality are used in each approach? What makes it narrow or wide; flexible or rigid; tolerant or not; mechanical; separated; or individualistic? In other words, what concept of Bounded Rationality gives an approach to BE its characteristics? This section, constructing on the previous two, seeks to provide a general answer to this question.

According to Gigerenzer (2020), there were two methodological breaking points in the conceptual development of Bounded Rationality – not one, as the 'two periods' interpretation suggests. The first break – the transition from 'Original' to 'New' BE, explained in Section II.2 – focus on how people adapt decisions under different circumstances (Simon's approach) to the identification of systematic errors from the agent's limited cognitive apparatus. Hence, Bounded Rationality changed from its "original" form to another – what we called "Biased" Bounded Rationality in this dissertation.

The second shift occurred sometime after the 1970s and consisted of an even more neoclassical take on Bounded Rationality, already alluded to in the previous section. The main

characteristic of this second transformation was the use of a third type of Bounded Rationality, one that, like Biased Bounded Rationality, does not conform to the principles used by Simon. Unlike Biased Bounded Rationality, however, this third concept was not focused on identifying or correcting systematic cognitive errors. Instead, it maintained the structure of neoclassical maximizing behavior but added to it cognitive constraints. As such, maximization was still presumed possible, but a number of constraints impede it from happening systematically.

This concept – which Gigerenzer (2020) deems the ‘optimizing under constraints’ Bounded Rationality, or Optimizing Bounded Rationality as it will be referred to from now on – does not presuppose that perfect and almost perfect behavior differ significantly. As such, this version does not presuppose that Bounded Rationality is a structural critique of Substantive Rationality. Consequently, cognition limits operate in a way similar to income restrictions – agents have a given limit to cognition. What decisions can an agent make to maximize their utility in this case? In other words, the neoclassical logic is preserved, and, unlike in Simon, optimizing outcomes are possible. This “double take-over” of the original concept of Bounded Rationality, Gigerenzer (2020) suggests, was “so subtle that few people notice” (p. 55).

The historical literature is not clear on specific events or dates of this second shift. Arguably, however, it precedes Simon. Thaler (1997), for instance, argues that Irving Fisher’s research forwards many points of BE by incorporating psychology into economics. It’s also worthy of note the fact that the 1950s and 1960s saw two other significant attempts to reform the neoclassical behavioral model: Adaptive Expectations of Phillip Cagan, Milton Friedman, and Marc Nerlove (Evans and Honkapohja, 2001a) and Rational Expectations, originally proposed by John Muth and popularized by Nobel laureate Robert Lucas (Tobón, 2014). The evolution of Optimizing Bounded Rationality, therefore, cannot be said to conclusively originate from Simon himself – it is also possible (and in fact more probable) that this general trend of behavior reform to be part of this greater movement of reforming rational behavior theory, only later “capturing” Bounded Rationality as part of it.

Gigerenzer (2020) presupposes, therefore, that the “neoclassification” of BE, and the drift away from Simon’s concept, did not stop at Kahneman or Tversky but kept going. The basic premise of Gigerenzer’s interpretation is this: three versions of Bounded Rationality took shape after the 1970s – one closer to Simon (‘Ecological Rationality’, according to him, referred to in this dissertation as ‘Optimizing Bounded Rationality’), another in line with Kahneman and

Tversky ('Bounded Rationality as irrationality' to him, referred as 'Biased Bounded Rationality' here), and another closer to traditional neoclassicism ('optimization under constraints' to him, 'Optimizing Bounded Rationality' in this dissertation). As such, "Bounded Rationality has, surprisingly, three faces" (Gigerenzer, 2020, p. 55). These three faces are wildly different from one another, both in research scope and in methodological consequences. Every type of Bounded Rationality presupposes an underlying normative theory – or lack thereof. In sequence, we present the three interpretations and their characteristics.

Simon's Original concept – Ecological (or Adaptive) Bounded Rationality

According to Gigerenzer (2020), the main trait of Bounded Rationality originally was the adaptive nature of the agent to the environment. His version of Bounded Rationality, deemed by Gigerenzer as 'Ecological Rationality'⁹¹ consists on the formalization of the agent-environment adaptive dynamic. In fact, Gigerenzer goes beyond, claiming that Ecological Rationality in the form of *homo heuristicus* is, indeed, the version of Bounded Rationality of Simon: "Homo heuristicus personifies Simon's original program to study decision making under uncertainty" (Gigerenzer, 2020, p. 61). Like *homo economicus*, *homo heuristicus* is an ideal type that serves as a representation of the fundamental elements of a decision agent. Unlike *homo economicus*, however, *homo heuristicus* is not capable of optimizing behavior; the focus of this ideal type is the adaptation of agents to an environment by the use of cognitive heuristics.

As the reader may already infer from Chapter I, we do not wholly concur with Gigerenzer when he proposes that this *was* the behavioral ideal type of Simon, even if the underlying 'spirit' of the argument is correct. Let us start with the caveats. For starters, in the material used in this research, Simon never once used the term 'Ecological Rationality'. This term more aptly describes the works of Gigerenzer and Vernon Smith, even if both authors use the term in different ways⁹² (Dekker and Remic, 2019). While it is true that Simon often stressed

⁹¹ 'Ecological' here refers to the environment of decision and has no implication on environmental concerns like the preservation of nature.

⁹² Gigerenzer's (2002) Ecological Rationality is based on the work of psychologist Egon Brunswik, which emphasizes adaptation of an organism to his environment. Smith (2003), on the other hand, interprets

the role of the environment, it is misleading to suggest that this was his focus. Simon incorporated the environment in his basic framework of Bounded Rationality (refer to Section I.2), but his focus was on decision-making itself – not the specifics of the environment of decision-making. In other words, the context of a decision only defined the problem space to the agent; besides this fact, Simon did not strive to understand the implications of the environment to decision-making like Gigerenzer does.

Moreover, Ecological Rationality puts AI and behavioral simulation in a secondary role. In fact, Gigerenzer is critical of the use of AI as a tool for behavioral research: “AI, Gigerenzer [...] maintains, is concerned with ‘disembodied intelligence’” (Fiori, 2011, p. 607) – not real organisms in real-life environments. In short, Gigerenzer does not recognize the equivalence between organisms and machines as does Simon; to him, “the way in which an evolved human brain solves problems is essentially different from that of a computer” (*ibid*). In other words, Gigerenzer gives little attention to the ‘cyborg’ nature of Simon’s science, a fact that, we argue, is central.

Finally, a profound methodological difference exists in the evolutionary aspects of Gigerenzer and Simon. Gigerenzer is preoccupied with the development of decision-making through evolutionary processes (Fiori, 2011) – a long learning process that equips humans with heuristics for decisions, what Gigerenzer calls the ‘Adaptive Toolbox’ of heuristic decision-making. In Gigerenzer, society is essential for this reason – it is a reservoir of past knowledge and wisdom that makes many heuristics available to individuals and guides their uses as tools for decision-making. Simon, on the other hand, does not present an evolutionary theory of any kind. As argues Loasby (1989), the learning processes in Simon’s theories only recall immediate experiences – no long process of learning through cumulative causation is part of Simon’s framework. This long learning, however, is central to Gigerenzer. Therefore, it would be more correct to suggest that Simon gave importance to the decision environment, never neglecting it, but the suggestion that Simon *used* Ecological Rationality is too bold a statement and does not appropriately reflect Simon’s paradigm.

rationality as an emerging property of the system, in line with theories of distributed cognition. Refer to Dekker and Remic (2019) for further detail.

Gigerenzer's Bounded Rationality, therefore, is heavily based around the original version of the concept – but does not use it exactly in the way that Simon used it. The similarities of methodological procedures and theoretical basis are sufficient for us to classify both as part of a research orientation, even if not *the* same orientation. Consequently, it is useful to classify both Simon and Gigerenzer as part of the Original BE: This approach keeps the search for alternative behavioral theories, taking the role of the environments into account, and not presupposing a complete normative theory as a guideline. As such, Gigerenzer did not practice the same science as that of Simon, but one with close methodological parallels and a discernible intellectual heritage to the 'prophet of Bounded Rationality'.

The main characteristic of Ecological Rationality, and Adaptive Bounded Rationality by extension, is the rejection of the idea that biased behavior is necessarily undesirable. Behavior is not presupposed to be biased *a priori*, and, even if it is, this type of behavior is not necessarily undesirable. Ecological Rationality sees heuristics as decision tools, appropriate to each decision – therefore the need for an 'Adaptive Toolbox'. In this analogy, heuristics are like screwdrivers – each one is appropriate for a specific type of screw; each heuristic is appropriate for a specific type of decision. Therefore, heuristics are not faulty mechanisms by themselves, as suggests by Kahneman and Tversky (1973; 2011 [1979]) – heuristics may be either advantageous or not; it depends on the circumstances of a decision.

'Adaptive Bounded Rationality', as we chose to call it, closely aligned with the version originally used by Simon but never formalized by him. This concept congregates the researcher who kept alive the search for new behavioral theories, never settling on a normative behavioral theory. Consequently, the 'satisficing' criteria are fluid. Incidentally, due to the non-adhesion of this type of Bounded Rationality to the neoclassical framework, this is the more 'heterodox' version of Bounded Rationality.

Kahneman and Tversky's concept – Biased Bounded Rationality

Biased Bounded Rationality originated in the works of Kahneman and Tversky (1973; 2011 [1979]) and is probably the most common interpretation of the concept. In practice, this version of Bounded Rationality is closely tied with Thaler's (1980) suggestion to identify the

“special circumstances” where neoclassical rationality is insufficient, complementing it with behavioral theories. This intuition, therefore, is direct: people make mistakes. The role of Biased Bounded Rationality and H&B more generally is to identify the circumstances of these deviations, their reasons, and what to do about it.

Deemed by Gigerenzer (2020) as ‘irrationality’, this form of cognition presupposes a distinction between rational and irrational behavior, with the latter being desirable, but the former predominant in everyday decisions. Kahneman (2017 [2011]) himself, however, denies this description of ‘irrationality’: “I often cringe when my work with Amos is credited with demonstrating that human choices are irrational, when in fact our research only showed that Humans are not well described by the rational-agent model” (p. 402). In any case, Biased Bounded Rationality does, indeed, focus on systematic errors, which can be interpreted as ‘irrationality’ in Viskovatoff’s (2001) sense – invalid inference. Therefore, Gigerenzer’s critique is correct if it does not presume that people act mainly manically or in a random manner but rather presumes that people stray from rationality and, consequently, are irrational.

Kahneman (2003; 2017 [2011]) proposes the ‘dual systems’ explanations to differentiate the two “modes” of behavior – deemed arbitrarily as ‘system 1’ and ‘system 2’⁹³. System 1 is the irrational part of human cognition; it is fast, instinctive, and relies heavily on heuristics as tools of decision. System 2, the rational one, is slow but capable of complex and deep thinking – maximization, in economics’ terms. Playfully deemed by Thaler and Sunstein (2021 [2008]) as, respectively, the “Mr. Spock lurking within you” (p. 44) and “everyone’s Homer Simpson” (*ibid.*), the dynamics of these cognitive systems make us prone to irrational, fast, and possibly erroneous behavior, reserving cognition for decisions where we truly need to be rational. Everyday behavior is inescapably biased because it relies so heavily on heuristics – and not on maximizing behavior. Unlike in Adaptive Rationality, heuristics are interpreted as deleterious to the decision-making process. As such, Biased Bounded Rationality tends to understand that maximization is possible if the underlying biases are lifted.

⁹³ Kahneman and Tversky (1974 and 2013 [1979]) did not include any mention of the terms ‘systems 1’ or ‘system 2’. One of the first mentions of this classification is Kahneman (2003), after Tversky’s death. Due to their alignment of ideas, Tversky probably would favor this classification of Kahneman if he had contact with it; however, this research could not find any mention that Tversky was aware of this scheme of system differentiation.

Therefore, Biased Bounded Rationality, unlike the Adaptive concept, presupposes that heuristics are naturally unfitted to generate ‘rational’ outcomes. As such, Gigerenzer (2020) argues that this type of Bounded Rationality depends on the existence of an underlying neoclassical normative theory: “Unlike Simon, the program maintained rational choice theory as normatively correct and blamed the discrepancies on people’s lack of rationality” (Gigerenzer, 2020, p. 60). Researchers that adopt Biased Bounded Rationality tend to deny this allegation of conceptual dependency to neoclassicism, but fail to formalize an independent normative theory. Kahneman and Tversky, for instance, defend the conceptual independence of BE to neoclassicism, proposing that the former is a critique of the latter (Kahneman and Tversky, 1973; 2011 [1979]; Kahneman, 2017 [2011]), but also condemns the ‘normative agnosticism’ of Gigerenzer’s Ecological Rationality, explicitly adopting the notion that behavior can be erroneous: “We think that such judgments [those not based on statistical frequentism and formal logic] should be treated as mistaken; they violate accepted constraints on the use of probability statements in everyday discourse” (Kahneman and Tversky, 1996, p. 586, adapted). In neither case, however, is an independent normative theory of ‘rationality’ proposed. Satisficing in this case depends on the ‘rational outcome’, defined by this murky normative theory.

As such, Biased Bounded Rationality, we argue, is closer to neoclassicism’s normative theory, even if the explanations themselves are not. This concept presupposes a ‘rational’ outcome, which is systematically out of reach due to the use of heuristics. This concept functions mainly as a complement to neoclassical behavior theory – not an independent explanation of it. Therefore, researchers who adopt this concept do not exactly look for alternative explanations for behavior, but rather to identify faults in the neoclassical rational behavior model.

The “neoclassical” concept – Optimizing Bounded Rationality

As already mentioned, a third branch of Bounded Rationality can be identified in current literature. Gigerenzer (2020) deems it ‘Bounded Rationality as optimization under constraints’. This is in line with neoclassical rationality, preserving its choice structure and normative theory. In essence, ‘Optimizing Bounded Rationality’ (as we chose to call it) preserves the structure of utility maximization, implying that, if cognitive costs are taken into consideration, Bounded Rationality is perfectly able to produce optimal results.

Similar to Biased Bounded Rationality, this concept is not directly linked with Simon’s framework. According to Gigerenzer (2020), this concept was popularized by George Stigler (1961), who changed “satisficing search to optimal search” (Gigerenzer, 2020, p. 59). Unlike the satisficing described by Simon (a random choice among satisfactory options), Stigler proposes that the ‘stopping rule’ of a decision is the moment when the costs associated with one option are exceeded by the benefits of it. In the words of Gigerenzer: “Stigler’s approach poured search theory back into the old bottle of expected utility maximization, introducing perfect foresight once again, and reducing uncertainty to risk” (*ibid*) – “Bounded Rationality is thereby seen as nothing but constrained optimization in disguise” (*ibid*).

This approach is much more open to its neoclassical influences than the Adaptive kind. Fundamentally, there is no attempt to distance this type of Bounded Rationality from orthodox economics. This group is part of the “repair program” of neoclassical economics of Bernoulli, identified by Gigerenzer and Berg (2010). Methodologically, this approach echoes Conlisk’s (1996) proposition. Taking the definition of economics by Lionel Robbins (2007 [1932]), Conlisk (1996) goes so far as to propose that “one could as well argue that economists are by definition required to study Bounded Rationality” (p. 686), for it also is a case of resource allocations – “Human cognition, as a scarce resource, should be treated as such” (p. 693).

In essence, this concept unties Bounded Rationality from its fundamental insight: if people are limitedly rational, optimization is impossible. Optimizing Bounded Rationality, by incorporating cognition as another form of restriction, enables a version of Bounded Rationality that is perfectly fitted to the orthodox microeconomic framework, enabling optimization once again.

Due to its closer adherence to neoclassicism, this concept of Bounded Rationality is the one more separated from other disciplines. Biased Bounded Rationality, even if it conforms to many of neoclassicism’s main tenets, still relies on (cognitive) psychology for inspiration in his

behavioral models. This is not the case in Optimizing Bounded Rationality. In fact, little first-hand references outside economics are used as a theoretical basis for this concept. Simon, for instance, “criticized Rubinstein⁹⁴ for showing no awareness of psychological research apart from that of Kahneman and Tversky” (Gigerenzer, 2020, p. 59). In other words, the main ‘outside’ reference to Optimizing Bounded Rationality is Behavioral Economics – not psychology.

Therefore, Optimizing Bounded Rationality is deeply tied with neoclassical economics. This concept, unlike the others, presupposes a cohesive and clear theoretical normativity: optimal outcomes, as defined by traditional microeconomics. Unlike the Biased concept, however, optimization is not an unattainable goal but a possible reality if taken into consideration that cognitive limits can be incorporated in economic models in an analogous manner to other restrictions.

Summary of the three Bounded Rationalities

Articulating the three interpretations of BE presented here (‘two periods’, Tomer’s, and the ‘three Bounded Rationalities’), we can observe three discernible, if not cohesive, uses for Bounded Rationality: Adaptive, Biased and Optimizing. Table IV presents a visualization of the concepts developed so far in this chapter. The following section proposes an interpretation of these results.

⁹⁴ One of the top researchers that adopt Optimizing Bounded Rationality.

Table IV – Classification of Behavioral Economics Approaches

Classification	Approaches	Authors	Type of Bounded Rationality	Traits
Original Behavioral Economics	Carnegie-Mellon; Michigan; Experimental Economics; Ecological Rationality (Berlin Group); Economical Psychology	Herbert Simon, George Katona, Tibor Scitovsky, Gerd Gigerenzer, Vernon Smith, Peter Earl	Adaptive Bounded Rationality	Focus on the relation between environment and cognition; Closer to heterodoxy. Adaptive normative decision criteria.
H&B Behavioral Economics	Psychological Economics, Behavioral Finance	Daniel Kahneman, Amos Tversky, Richard Thaler, David Laibson, George Lowenstein	Biased Bounded Rationality	Focus on systematic cognitive errors (biases). Presupposes “rational behavior” as normative theory.
Neoclassical Behavioral Economics	X-Efficiency, BE aligned with Neoclassicism	Harvey Leibeinstein, Kenneth Arrow, Ariel Rubinstein, John Conlisk, Raj Chetty, Matthew Rabin, Robert Shiller, Andrei Shleifer, Hersch Shefrin, Cars Hommes*	Optimizing Bounded Rationality	Focus on behavioral modeling, in line with neoclassical economics. Presupposes neoclassical normative criteria.
Instrumental Bounded Rationality	Nelson and Winter Evolutionary Theory; Behavioral Macroeconomics, New Institutional Economics	Richard Nelson, Sydney Winter, George Akerlof, Oliver Williamson	Varied; often adopts Simon’s or Kahneman’s version.	Varied.

Source: own elaboration, based on Sent (2004), Tomer (2007), de Castro (2014), Nagatsu (2015), Angner (2019), Gigerenzer (2020).

*This researcher was not already mentioned, but are presented in Chapter III, along with arguments to why they belong in each category.

II.5 – Bounded Rationality as an Axiom.

So far in this chapter, we have argued that Bounded Rationality has changed over the years. While hardly a divisive idea, the way that it changed seems to not be completely captured by the literature. Simon's lack of a 'general theory' made Bounded Rationality diffuse, and branching paths formed. Unlike the two-period distinction suggests, one strand of Bounded Rationality never replaced the others. Instead, we observe a simultaneous evolution of all three versions.

Concerning Simon, we can say that his mind was unique, even among his equals. The precursors to BE, George Katona and Tibor Scitovsky, did not adhere to Simon's hardcore, and, consequently, the first generation of Original Behavioral Economics did not share a cohesive nucleus of research. Katona and Scitovsky were not especially interested in machine learning or AI and cannot be considered 'cyborg scientists', even if Simon can. Their interests were more in line with (biological) psychology.

The evolution of Bounded Rationality eventually transcended Simon himself, acquiring new meanings in the economic sciences. In this chapter, we highlight three main interpretations of Bounded Rationality: Adaptive, Biased, and Optimizing. A fourth auxiliary concept is also proposed as a catch-all category that includes economic approaches that use Bounded Rationality but do not focus on behavior: Instrumental Bounded Rationality. When using the term 'Bounded Rationality', economists tend to refer to one of these concepts.

As such, a profound change can be observed – no longer is the broad term 'Bounded Rationality' is reminiscent of Simon's original approach presented in Chapter I. In its place, three main different concepts have taken shape. These are not "interpretations" of Simon – but rather different theoretical structures that are not necessarily linked to the author, even if the name remains the same. In Lakatos' (1958) terms, 'Bounded Rationality' is adapted to the hard-cores of other research programs⁹⁵. What, then, unifies the many uses of Bounded Rationality? The answer: not much. Let us now identify these few elements.

⁹⁵ I thank Dorricielk Gomes for pointing this out in these terms.

A similar appeal is common to all forms of Bounded Rationality: the claim that, “obviously”, people do not act in accordance with neoclassical rational choice theory. This re-examination, however, is not exclusive to BE – even among neoclassical economists, this perspective is widely acknowledged. Milton Friedman is a prominent example. Rather than proposing that people act rationally, Friedman (1953) argues that the specifics of behavior are unimportant, as long as predictions are accurate. Even Gary Becker, one of the crusaders of rational behavior, concedes the limitations of the “believability” of rational behavior. Rather than arguing that people always act rationally he proposes that social mechanisms, beyond rational agents, also enforce rational outcomes⁹⁶. In this view, agents do not necessarily need to be rational for rationality to emerge. Adaptive Expectations offers another illustrative example: in this theory, equilibrium (the optimal outcome) does not require agents to be perfectly rational. Through a dynamic process, ‘rational economic outcomes’ are achieved, even if agents never come closer to the exigencies of perfect rationality. A final example is Kenneth Arrow, mentioned earlier. Thus, many neoclassical economists, implicitly or explicitly, recognized the fact that rational economic outcomes do not depend solely on perfect rationality of agents.

So, what is unique in the recognition of this insufficiency in Bounded Rationality? The answer lies in its focus. Bounded Rationality recognizes the ‘emperor’s new clothes’ for what they truly are – everyone sees the naked emperor, but no one dares to shatter the illusion; people do not act rationally, but nonetheless we persist in acting like they do. The current iterations of Bounded Rationality, in this sense, are unified by a simple recognition of the obvious – people are not rational! Another question, however, remains: if people are not rational, what are they? What do we do now? The three main versions of Bounded Rationality presented in the previous section can be understood as an attempt to formulate an answer to these questions. Notice, however, that each answer is quite different from the others.

Consequently, few methodological traits are shared among Adaptive, Biased, and Optimizing Bounded Rationality. In other words, few commonalities are shared between the hard cores of research programs that adopt each interpretation of Bounded Rationality. There is no shared methodology, no common implications regarding the effects of Bounded Rationality,

⁹⁶ See for example, Kirzner (1962): “The essence of Becker's demonstration is to show how individuals, however irrational, are in the aggregate forced, in their passive roles of price-takers, to adjust systematically to changes in price” (p. 382).

no ‘general theory’, no underlying psychology, and not even a clear theoretical heritage. Each form of Bounded Rationality is effectively unique and dissociated from the others – except for the starting observation: people are not rational!

In this light, we propose that what currently unifies Bounded Rationality is an axiomatic attitude towards economic behavior. An axiom is “an indemonstrable first principle, rule, or maxim that has found general acceptance or is thought worthy of common acceptance whether by virtue of a claim to intrinsic merit or on the basis of an appeal to self-evidence”⁹⁷. In other words, this “common attitude” is the “first principle” of general acceptance; the “intrinsic merit” lies in the self-justifying notion that theory should closely describe actual human behavior; the “obvious” fact that “people are not rational!” is part of this self-evident conceptual construction that Bounded Rationality acquired in recent literature. Once again, no clear list of methodological elements can be identified within Bounded Rationality. This, we argue, is the primary distinction between Simon’s original concept and Bounded Rationality in economics. Bounded Rationality has thus evolved into a general attitude towards research, requiring few theoretical commitments for its application.

Gradually, Bounded Rationality became a *Tabula Rasa*. Little of the positive theory proposed by Katona and Scitovsky survived. Even Simon, its most prominent researcher, had limited influence on the specifics of this behavioral theory. Some authors, such as Gigerenzer and Smith, remained closer to Simon, but they represent a minority. Most economists adapted the “spirit” of Bounded Rationality to their theories – but not the structures of behavior originally presupposed in it. ‘Cognitive limits’ can signify small tweaks to neoclassical rationality or imply a whole different behavioral structure. ‘Satisficing’ may mean the original mechanism described by Simon or something else entirely. The environment can either be the most important factor of a decision or a secondary preoccupation. The evolution of Bounded Rationality only made these divergences more apparent.

Moreover, has economics truly relied on only *one* behavioral theory all this time, requiring such an extensive effort to reform it? The answer is no – many other explanations for behavior exist in economics, such as those of Keynesian, Marxist, Institutional, and even Austrian economics. Why did BE have to look outside economics, a field rich with behavioral

⁹⁷ Source: <https://www.britannica.com/topic/axiom>

theories, for its source of inspiration? Moreover, why did BE limit his interests to Cognitive Psychology? Why was no wider effort made to incorporate other forms of psychology if this was his goal? The answer lies in the characteristics required for this new behavioral theory: it needed to be different, but not too different from established economics. Heterodox economics, as well as other forms of psychology, were likely too disruptive from the desired theory.

This impulse to adopt psychology as a form of increasing the descriptive power of a theory is not immediately clear as it may first appear. Economics had other incursions with psychology, most notably the pragmatic philosophy that provided the basis for Original Institutional Economics in the USA. In fact, the dissatisfaction with the behavioral theories of neoclassical economics far outdates Simon's work. Veblen (1898) famously critiqued the mechanistic aspects of neoclassicism, while Mitchell (1910) explicitly calls out the inadequacies of the psychological basis that structure this economic approach:

“If economists cannot hand over to others the study of how men have acquired their present measure of rationality without falling into misconceptions which vitiate the work which they themselves attempt to perform, they clearly do need the psychological training” (Mitchell, 1910, p. 216).

Mitchell's words are of special note, for they closely echo Simon's position.

As such, the appeal to consider other types of behavior is nothing new; what is unique in both BE and Bounded Rationality is to make it the focus of a research program. Each strand of Bounded Rationality looks at a specific “sub-focus”: Adaptive Bounded Rationality observes the relation between agent and environment; Biased Bounded Rationality focuses on ‘systematic errors’; and Optimizing Bounded Rationality looks at maximizing behavior. No common theory, critique, or goal unifies them – only the axiomatic conviction that “people are not rational!”.

CONCLUSION OF CHAPTER II

In this chapter, we analyzed the conceptual evolution of Bounded Rationality. Three main contributions were articulated: the “two periods” interpretation (Sent, 2004; Nagatsu,

2015; Angner, 2019), the research programs list (Tomer, 2007; de Castro, 2014) and the “three Bounded Rationalities” proposition (Gigerenzer, 2020). This articulation, synthesized in Table IV, serves as a general “map” of the uses of Bounded Rationality after Herbert Simon. Every style of Bounded Rationality fits a “niche” in BE.

What we can observe, however, is that the current approaches to Bounded Rationality share little in common with one another. Incredibly, however, the different strands of Bounded Rationality do not share a common ancestry. Original BE is the one closer to Simon and the one which more closely follows his legacy. H&B is deeply tied with Kahneman and Tversky’s work, focusing instead on the deviation of ‘rational’ behavior. This approach only incorporated Bounded Rationality after its methodology was established. Therefore, the effects of Simon in H&B were limited. ‘Optimization under constraints’ is an approach that only Gigerenzer (2020) brought significant attention. Closer to neoclassicism, the methodology of this strand can be more reliably traced to other forms of neoclassical reforms of Substantive Rationality (such as Adaptive or Rational Expectations) than any form of Bounded Rationality *per se*.

Finally, we propose that few commonalities unify the current uses of Bounded Rationality. The most notable shared trait is an axiomatic use of the concept: agents do not adhere to the tenements of maximizing neoclassical rationality – they have limited rationality. What these “limits” imply, however, depends on each approach – no common answer is possible.

Consequently, Simon’s concept changed considerably since its inception. From a loose, yet cohesive methodology, Bounded Rationality evolved to a flexible concept, able to be applied in a variety of situations. This, off course, came with a price. Bounded Rationality lost its “core”. Gigerenzer and Selten (2001) may had put it best: “bounded rationality has become a fashionable label for every model of human behavior” (p. 4). The next chapter discusses the recent literature on Bounded Rationality.

CHAPTER III

4. The Legacy of Herbert Simon in Current-day Economics: Bibliographical evidences of the nature and uses of Bounded Rationality.

“If you don't know where you are going, any road will
get you there”

- Lewis Carrol, Alice in Wonderland, 1864

So far, we have developed a simple argument: the version (or versions) of Bounded Rationality that were adopted in economics, while loosely related to Herbert Simon, are not closely linked with his specific methodology. Instead, Bounded Rationality's main contribution to economics is a ‘general methodological spirit’, versatile in its application. This versatility is the main trait that has allowed this type of rationality to be so widely adopted – but it came at a cost. A cacophony of meanings has formed around this concept, unified only in name – not in substance. Consequently, the meanings of the term “Bounded Rationality” are numerous, even though the boundaries between one concept or another are not clear.

The evidence for this narrative is mainly bibliographical, but other tools can be applied to this end. This final Chapter consists of a bibliometric analysis of the recent developments in Bounded Rationality. The goal here is to differentiate the impact of “Bounded Rationality” from the impacts of Herbert Simon in recent literature. Therefore, we strive to distinguish the contributions *of* Herbert Simon from those *based* on his work. This, we hope, will help elucidate which types of Bounded Rationality (Adaptive, Biased or Optimizing) have impacted the development of economics, and in what ways.

Effectively, what this chapter attempts is a hypothesis test. Based on the evidence presented in Chapter II, the hypothesis to be tested is as follows: the absorption of Bounded Rationality into economics is linked to the work of Herbert Simon (H0) or not (H1). Accepting the null hypothesis means that the popular narrative, which closely links Simon and Bounded Rationality, is correct. On the other hand, the alternative hypothesis signifies a shift away from

Simon, and a transformation of the concept – a ‘transcendence’ of its original creator, as suggested in Chapter II.

This chapter is divided into eight parts and is richly subdivided: Section III.1 presents the method of data gathering and details the methodology of the bibliometric analysis in question. This analysis is divided into two samples: the first is composed of articles that discuss Bounded Rationality; the second includes papers that discuss both Bounded Rationality *and* Simon. For simplicity’s sake, the samples are named Sample 1 and Sample 2, respectively. Section III.2 presents the descriptive information and analysis of Sample 1, while Section III.3 provides data visualization. Section III.4 is a summary of the results of Sample 1. Section III.5 is analogous to III.2, presenting the descriptive analysis of Sample 2, a subgroup of Sample 1. Naturally, Section III.5 presents the descriptive information of Sample 2, and Section III.6 presents the visualization of the data from Sample 2. Section III.7 summarizes Sample 2. Section III.8 presents a summary of the findings and correlates them with the earlier chapters.

III.1 – Data Collection and Bibliometric Methodology

Bibliometrics is a field of study that quantifies, compares, and analyzes data, much like traditional econometrics. The main difference, as the name suggests, is the analyzed elements: econometrics deals with datasets expressed in numbers, bibliometrics compares datasets composed of words. In simple terms, bibliometrics analyzes the correlation between sets of words, comparing them for similarity or dissimilarity and identifying distances between groups of words.

Pertaining to the bibliometrics of Bounded Rationality, some works should be highlighted. Truc (2021) presents a comprehensive analysis of the subfields of Behavioral Economics⁹⁸, focusing on cross-disciplinary citations with psychology. Jain et al. (2021) map the field of behavioral biases, closely linked to Kahneman and Tversky’s research program. Costa, Carvalho, and Moreira (2019) provide a general overview of the fields of Behavioral

⁹⁸ Truc (2021) divides Behavioral Economics into seven fields: Social Preferences; Biases, Heuristics and Nudges; Behavioral Finance; Risk and Uncertainty; Intertemporal Choice; Neuroeconomics and Behavioral Game Theory.

Economics and Behavioral Finance, aiming to differentiate the two. Finally, Aoujil et al. (2023) analyzed the interface between BE and Artificial Intelligence.

While interesting, none of these articles directly related to the question at hand. Firstly, ‘Bounded Rationality’ is mostly implied, as the term is seldom cited. In Truc (2021), ‘Bounded Rationality’ is only mentioned twice, and the same is true for Aoujil et al. (2023). Both Costa, Carvalho, and Moreira (2019) and Jain et al. (2021) do not mention Bounded Rationality at all. Secondly, the discussion around Herbert Simon is incidental in these papers. Only Costa, Carvalho, and Moreira (2019) and Aoujil et al. (2023) include Herbert Simon in their references (Truc, 2021, only mentions Simon in passing, and Jain et al., 2021, does not mention the author). Third, none of these bibliographical researches examined the sociological aspects of BE, such as their place in the heterodoxy-orthodoxy debate or its relation to current mainstream economics.

Therefore, these references only tangentially address either Herbert Simon or Bounded Rationality, even if their main theme is BE. The inspiration for the bibliometric analysis, therefore, comes from other articles such as Almeida and de Paula (2019), Maltaca and Almeida (2022), Brites and Almeida (2023), Almeida, Cruz-e-Silva, and Brites (2023), and Sposito (2024). These articles use bibliometric tools to analyze the intellectual structure of subfields of economics, relating them to specific traditions of thought, association, and journals. The work of Almeida and co-authors focuses on criteria of economics associations, an important element for identifying the inner workings of heterodoxy. Sposito (2024) engages a similar approach, adopting the classification of heterodox and orthodox journals proposed by Lee et al. (2010). As far as this research could ascertain, this effort was not yet made in the case of BE, and it could prove a valuable step in guiding the discussion of the place of BE into the heterodoxy-orthodoxy debate. Therefore, the main inspiration of this bibliometric section is the work of Almeida and co-authors, even if the influence is only methodological.

The mathematical method of bibliometrics can be synthesized as follows: documents (book chapters, files, news, or, in this case, scientific articles) are converted into numeric coordinates and compared. Similar documents (with similar coordinates) are clustered into groups, allowing for the comparison of documents (Huang, 2008). The distance between the “centers”⁹⁹ of one cluster and other clusters’ “centers” defines their conceptual structure.

⁹⁹ The rigorous term is “centroid”. See Appendix I for an explanation.

A number of attributes of these documents may be analyzed. In this dissertation, we chose a main form of bibliometric coupling known as ‘co-citation’¹⁰⁰, specifically the co-citation of authors and references. The reason for this option is simple: these two forms of co-citation reveal the sources of the analyzed articles. For example, this dissertation cites many works by Herbert Simon; they will be co-cited when other documents (papers, articles, theses, etc.) also cite the same references. As such, a “common ground” between two articles forms, and their common references can be identified. In the case of a co-citation of references, for instance, this common ground is the references cited in both, giving an idea of the themes discussed in both articles.

The attributes identified in the co-citation method are then grouped into clusters. These clusters permit the analysis of datasets (groups of articles), making it possible to identify the characteristics of a research subject¹⁰¹. In simpler terms, co-citation gives us a “map” of the subject. Following the workflow proposed by Aria and Cuccurullo (2017), this research was conducted in three steps: 1) data collection (Section III.1); 2) data analysis (III.2 and III.4); and 3) data visualization (III.3 and III.5).

The data used in this bibliometric analysis were obtained in the databases *Scopus*¹⁰². The database *Web of Science* was also considered; however, as the number of indexed journals in this database is significantly smaller than in *Scopus*, the results were less representative of the field. For this reason, this dissertation analyzed only *Scopus*.

Journals were classified as either “heterodox” or “orthodox”, in line with the definition proposed by Lee et al. (2010)¹⁰³, in an attempt to identify which economic tradition absorbed

¹⁰⁰ Co-citation is a subcase of bibliometric coupling. Coupling is defined as an attribute present in two different documents (Aria and Cuccurullo, 2017). Co-citation is the citation of a same source in two different articles (*ibid*).

¹⁰¹ Due to the technical nature of the bibliometric method and their tangential relevance to the main discussion of this dissertation, the explanations for the workings of bibliometrics were omitted from the main body of the text. Refer to Appendix I for a brief explanation.

¹⁰² Available at scopus.com.

¹⁰³ Lee et al. (2010) define a ranking of the quality of the research from heterodox journals based in “a comprehensive or total citation-based ranking of the journals that combines a bibliometric ranking based on the methodology” (p. 1411), later curated by peer evaluation. It is important to note that Lee et al. (2010) define “heterodoxy” as the contrary to “mainstream”, however, the ranking of heterodox journals in this

Bounded Rationality and how. In Lee et al. (2010), the authors combine the ranking based on total citations with a social networking ranking to infer the impacts of a journal in the field of economics. Later the journals are peer-evaluated (see Lee et al., 2010, p. 1411 for further methodological details). This ranking of heterodox journals is based on Lee (2008), and the criteria for a journal to be considered heterodox are laid down in Lee (2009). As implied in Chapter II, the adoption of Bounded Rationality by heterodox approaches signifies a critical conceptual interpretation. An orthodox absorption, on the other hand, signifies an interpretation of Bounded Rationality closer to neoclassical economics.

All the dataset analyzed are composed of articles published between 2013 and 2023 in the area of economics¹⁰⁴, which contained the word ‘*Bounded Rationality*’ in the title, abstract, or in the keywords. This time frame was chosen because it encompasses, roughly, the last decade of developments¹⁰⁵. From this point, the resulting data was divided into two groups: the first group (Sample 1) is the search described above with no extra parameters. The second group (Sample 2) is identical to this first, with one exception: along with ‘*Bounded Rationality*’, this search also included the keyword ‘*Simon*’¹⁰⁶. Unfortunately, this would also include references to other authors with the same surname. The review of the abstracts, however, reveals that the articles tend to follow the desired theme of analysis and, as such, indeed provide a representative sample. That said, the reader should keep in mind that exceptions may have passed, as the systematic review would be impossible in Sample 1 due to the number of observations. As the objective of this bibliometric section is not to systematize the available knowledge but to assess the impact of an author in the literature, this methodology is sufficiently adequate.

Finally, we present two co-citation networks: 1) co-citation of authors and 2) co-citation of references. The co-citation of authors provides a broad, general view of the most co-cited authors in the samples – the main *corpus* of research. This, however, says little about the influence of an author in the intellectual structure of the field. With it in mind, the focus of

manner broadly coincides with the definitions presented in Section I.5. A close comparison to this system is CAPES’ *Qualis* system in Brazil.

¹⁰⁴ In Scopus, this area is identified as “*Economics, Econometrics and Finance*”.

¹⁰⁵ 2023 was chosen instead of 2024 because it is the last whole year while this research was being written.

¹⁰⁶ Due to reference conventions, which only require the surname of the author, we opted to use ‘*Simon*’ instead of ‘*Hebert Simon*’.

analysis will be the co-citation network of references. This allows a much more refined and precise analysis of the impact of an author on the different research programs that developed around a specific theme (Bounded Rationality, in this case). The information presented is compiled and further analyzed in Section III.4.

All of the data were processed in the *R* software (R Core Team, 2022) with the help of the *RStudio* interface (RStudio Team, 2020). The specific routines used to process the data are part of the *bibliometrix* package (Aria and Cuccurullo, 2017), and the majority of the analysis was done in the sub-interface *biblioshiny*, included in the *bibliometrix* package. The figures were also generated with *biblioshiny*.

III.2 – Data analysis of Sample 1

Main Information of Sample 1

From 2013 to 2023, a total of 707 articles in the field of economics indexed in *Scopus* included ‘*Bounded Rationality*’ in the paper’s title, abstract, or keywords. These papers were published in 274 different sources and included a total of 1,282 authors. The annual growth rate during the period was 5.38%, and the average number of citations of these articles per year is 13.68. Finally, a total of 32,906 references were cited in these articles. Table V summarizes the leading publishing journals.

Table V – Leading Publishers in Sample 1 (2013 – 2023)

Journal	N. of Papers	Intellectual Tradition	Additional Information
<i>Journal of Economic Behavior and Organization</i>	45	Heterodox*	Focus on economic decision-making, including human and computational cognition.
<i>Games and Economic Behavior</i>	25	Orthodox**	Focus on game theory. Also publishes in related areas, like mathematics and biology.
<i>Journal of Economic Theory</i>	21	Orthodox*	General interest, focuses on economic theory.
<i>Journal of Economic Dynamics and Control</i>	20	Orthodox*	General interest, highlights the use of computational methods.
<i>Theory and Decision</i>	18	Orthodox*	Focus on interdisciplinary decision-making.
<i>Journal of Mathematical Economics</i>	18	Orthodox**	Focus on formal mathematical models.
<i>Journal of Evolutionary Economics</i>	14	Heterodox*	Focus on the Schumpeterian evolutionary framework, including complexity.
<i>Computational Economics</i>	12	Orthodox**	Focus on computer sciences.
<i>Economic Modelling</i>	12	Orthodox*	Focus on large-scale macroeconomic models, including VAR and DSGE.
<i>Journal of Advanced Transportation</i>	12	Unrelated	Focus on multi-modal transport and traffic systems, unrelated to economics.

Source: own elaboration.

* Classification of Lee et al. (2010).

** Own classification, made with the review of the aims and scopes of each journal.

Table V the profile of publishing journals that publish the theme of Bounded Rationality – either heterodox or orthodox. Together, two heterodox journals were part of the top ten publishing sources (*Journal of Economic Behavior and Organization*, and *Journal of Evolutionary Economics*), representing almost 32% (59/185¹⁰⁷) of the total, including the leading publisher. The *Journal of Economic Behavior and Organization* wasn't only the journal that published the most research on Bounded Rationality, but was also the number one according to the h-index of 14¹⁰⁸, revealing an influence of heterodox journals in the research of Bounded Rationality. That being said, a general leaning towards mainstream orthodox economics is clear in the leading publishers. The remaining 68% of papers were published in journals that can be considered orthodox.

Thematic Map of Sample 1.

Here, we present the thematic map of Sample 1. The thematic map, originally developed by Callon, Courtial, and Laville (1991), is “a coordinate system consisting of centrality (x-axis) and density (y-axis)” (Yu and Muñoz-Justiticia, 2020, p. 4) used to analyze research subjects. Effectively, this is a visualization method for a research field, identifying themes by their relevance on the x-axis and their development as research themes on the y-axis. The higher a cluster is in the y-axis, the higher its density (degree of development). Its position in the x-axis, indicates a cluster's relevance to the field: the farther to the left a cluster is, the less central it is; the farther to the right, the more central (relevant) it is.

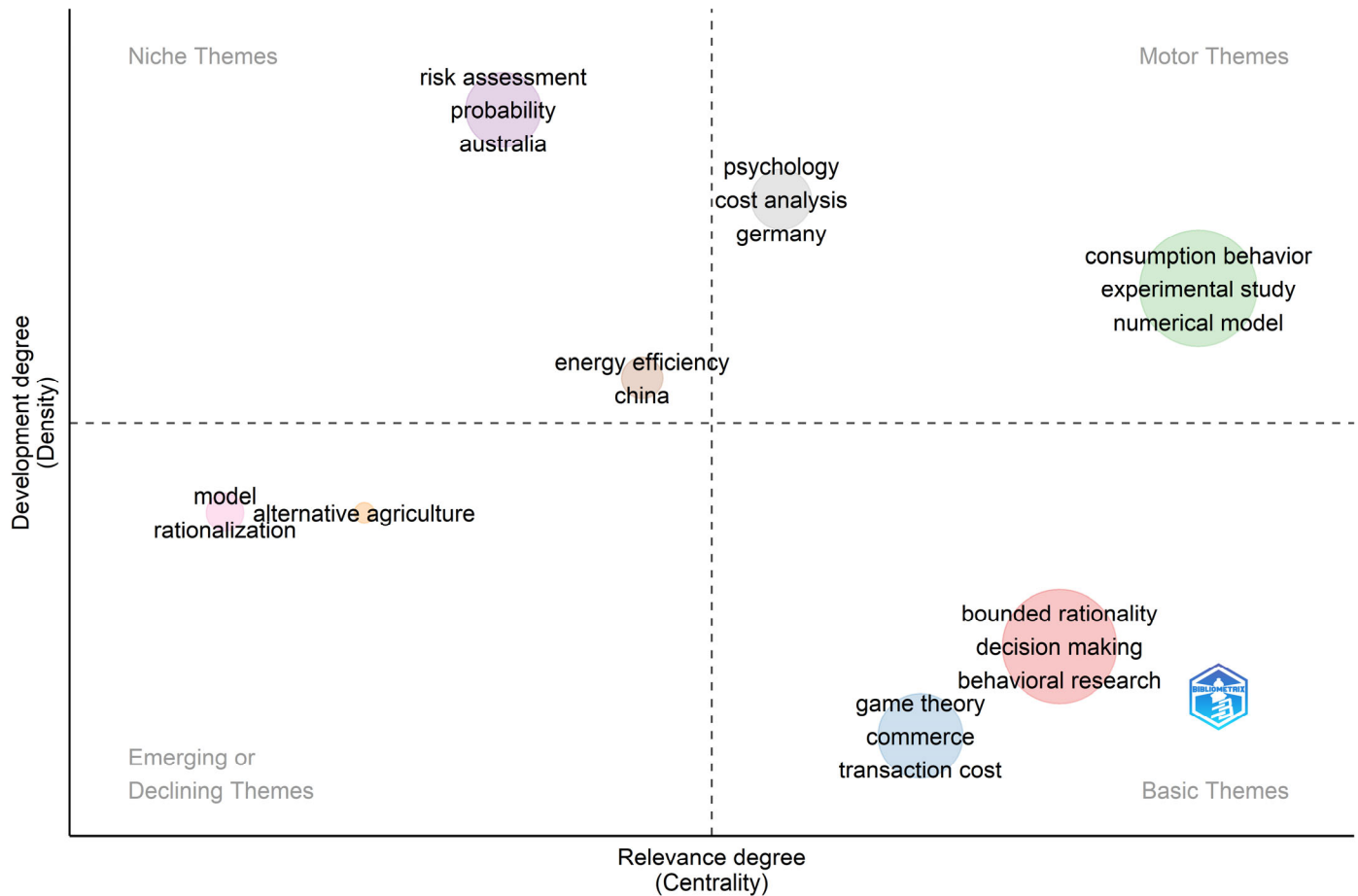
This visualization method was later implemented by Cobo et al. (2010). Quadrant 1 represents central and developed themes; these are the “motor themes” of a research subject (Cobo et al., 2010); quadrant 2 are central but undeveloped themes, these are the basic and transversal themes (*ibid.*); quadrant 3 consists of developed peripheral themes, the niche themes (*ibid.*); and quadrant 4 is peripheral and underdeveloped themes, these indicate either emerging

¹⁰⁷ This is the number excluding the *Journal of Advanced Transportation*.

¹⁰⁸ The h-index “is an author's number of articles (h) that have received at least h citations” (Engqvist and Frommen, 2008, p. 250) and is one of the leading metrics of impact. However, as Engqvist and Frommen (*ibid*) suggest, this index may be misleading, for it does not differentiate citation of other researchers from self-citation, making it unreliable in certain cases.

or declining themes of research (*ibid.*)¹⁰⁹. This is the version used in *bibliometrix*. Figure II is the thematic map of Sample 1:

¹⁰⁹ For a visualization of the interpretation of the thematic map, see Yu and Muñoz-Justicia (2020), reproduced in Appendix III.

Figure I – Thematic Map of Sample 1

The central motor themes of Sample 1 are consumption behavior, experimental studies, and numerical modeling (green cluster). The central basic, transversal themes are Bounded Rationality, decision-making, and behavioral research (red cluster), followed by game theory, commerce, and transaction costs (blue cluster). The niche themes are risk assessment, probability, and, curiously, Australia¹¹⁰ (purple cluster). Other highly developed themes represented in the figure are psychology and cost analysis (gray cluster) and energy efficiency (brown cluster). We can note, therefore, that highly developed themes of Sample 1 are tied with applications of Bounded Rationality – empirical research, for lack of a better term.

¹¹⁰ Australia, along with Germany (in the gray cluster), and China (in the brown cluster) are tied with case studies, such as Coggan et al. (2015), Mußhoff et al. (2018), and Zhao et al. (2023).

In the emerging or declining themes are rationalization modeling and alternative agriculture. ‘Rationalization’ (the act of rationalizing) is linked with efforts such as Manzini and Mariotti (2007), which propose a model of how agents choose progressively worse options – how they rationalize a worse decision. The agriculture cluster refers to articles such as Křečková and Brožová (2017) and is also mainly tied with case studies.

The classic themes of Herbert Simon that can be gleaned in this thematic map (based on the characterization in Chapter I) are in the red and green clusters, the central clusters of motor and basic themes. Therefore, the themes of Simon still remain relevant to this day. However, the thematic map does not show any detail on how these themes are being treated. Therefore, with this information, it is not possible to suggest that Simon’s approach remained relevant – and neither how he affected the application of Bounded Rationality.

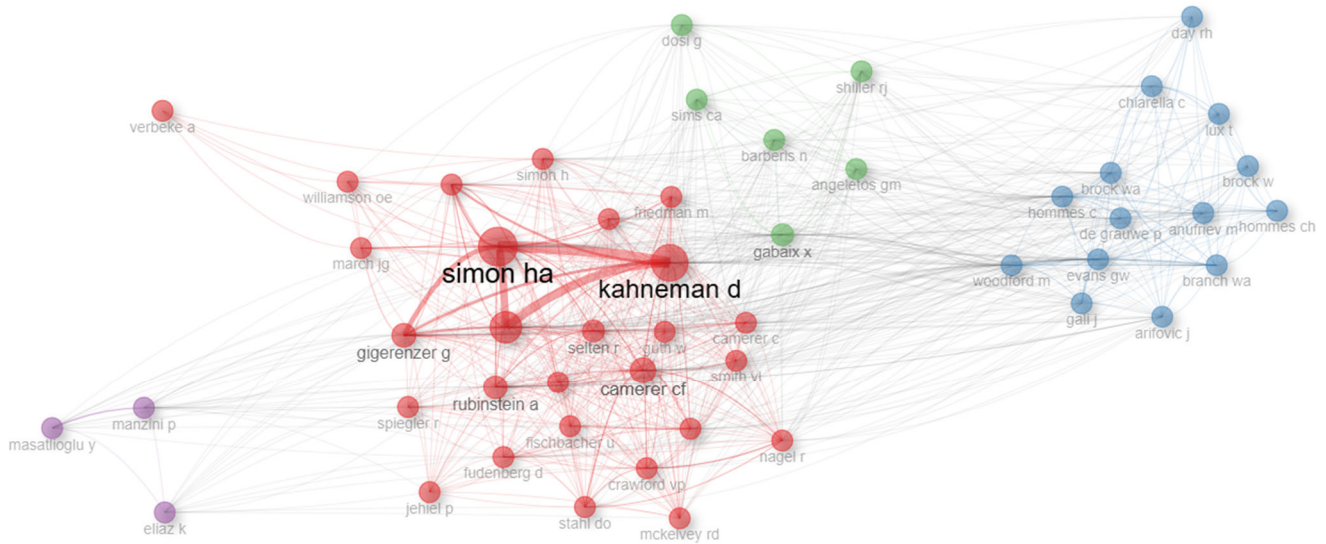
III.3 – Data visualization of Sample 1

This section presents a visualization of the data in Sample 1. Two styles of representation were used: authors co-citation and reference co-citation. Authors co-citation visually represents authors that are cited together. This reveals, roughly, the authors that compose the most cited references in a sample. This type of co-citation, however, does not differentiate the sources of the citations, nor their relevance to specific themes. As such, authors with little similarity may be bundled together due to a high number of citations or due to a diverse body of work. Simon is both much cited and diverse in his research. Therefore, this is a useful tool to perceive general trends – what are the ‘fundamentals’ of a field.

With this in mind, this section also analyzes co-citations of referenced articles. This tool is far more reliable than the first one, due to the fact that references, and not authors, are linked. As articles are much more specific than authors in their theme, the impact on different fields becomes much more discernible. Consequently, co-citation of references is the backbone of this analysis. This section is further subdivided to make the presentation of information simpler. Figure II is a representation of the co-citation networks of authors in Sample 1.

Co-citation Network of Authors

Figure II – Co-citation Network of Authors of Sample 1



Source: own elaboration, generated in *biblioshiny*.

The biggest nodule in this co-citation network is Herbert Simon, making him the most co-cited author in this sample. This does not signify that he is the most broadly cited author – only that many sources cite him¹¹¹. As such, little can be said about the “depths” of Simon’s influence based on this co-citation. In any occasion, this helps to illustrate the fact that Simon is still widely mentioned in the discussion of Bounded Rationality, independently of the degree that economics incorporated his science.

Other influential names that are linked with Bounded Rationality and are part of this main central cluster (red) are Daniel Kahneman, Amos Tversky, Gerd Gigerenzer, Colin Camerer, James March, Vernon Smith, Richard Thaler, Ariel Rubinstein, Oliver Williamson, and Milton Friedman. This cluster is the “fundamentals” of contemporary research in Bounded Rationality, congregating the most celebrated authors that developed this theme. However, little else unifies these names.

¹¹¹ Only one mention of Simon in the ‘references’ section of an article is sufficient for this article to be co-cited with another. Co-citation of authors, therefore, functions in a logic of “all-or-nothing” – it represents the articles that cite Simon *at all*.

The rightmost cluster (blue) reveals an interlinking with authors that are not part of the mainstream cluster. This means that these are authors working with Bounded Rationality, but their main references are not part of the main cluster (red). This cluster includes authors that research the mathematical modeling of Bounded Rationality, such as George W. Evans (a researcher focused on behavioral equilibrium dynamics), Cars Hommes (focused in behavioral macroeconomics and experimental economics), Michael Woodford (focused in price dynamics), William Brock (focused in behavioral modeling), among others. These authors are analyzed in more detail in the following co-citation network of reference.

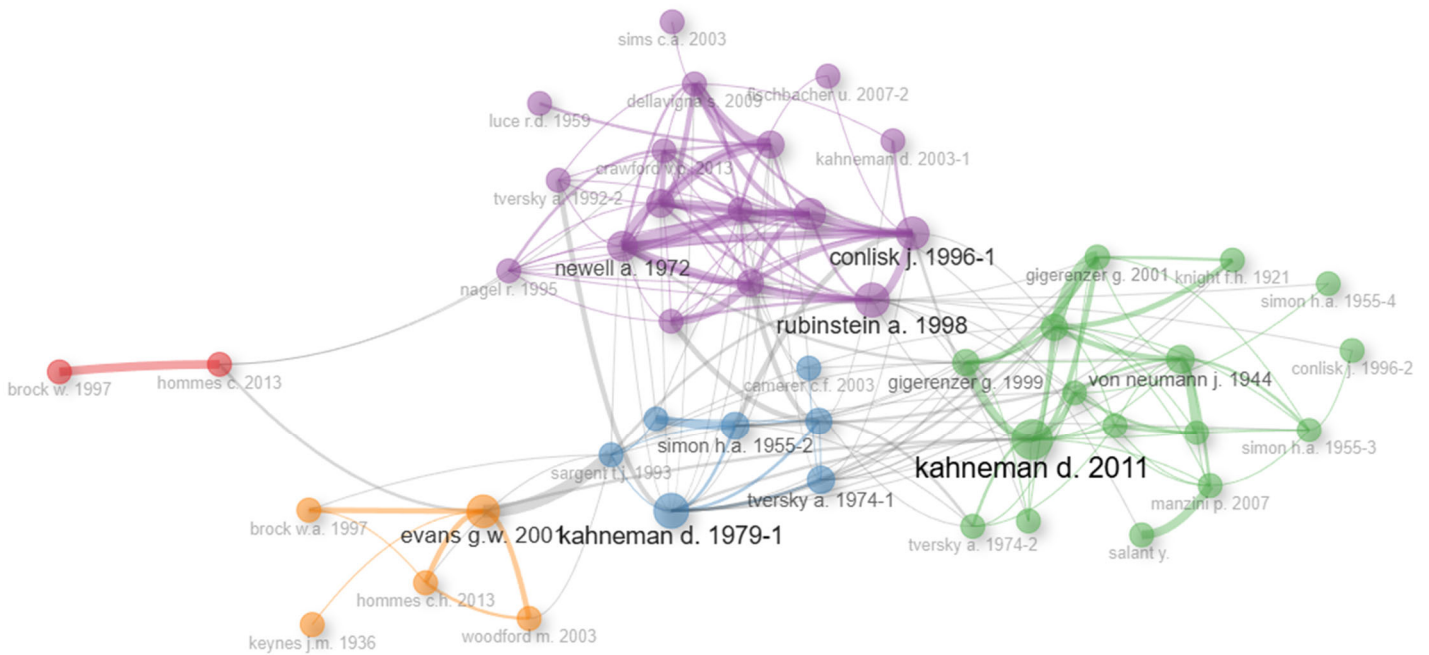
The small central cluster (green) includes authors like Giovanni Dosi (a leading researcher in many fields, such as economic innovation, evolutionary theory, ABM macroeconomics, and complexity economics), Robert Shiller (a researcher in Behavioral Finance), Xavier Gabaix (an author in Behavioral Finance) and Nicholas Barberis (focused on behavioral theories of investment trading and financial assets). These authors do not appear as main references in the co-citation network of references and, therefore, do not compose a discernible orientation of research in Bounded Rationality. As such, these are the least central authors in the theme of Bounded Rationality, associated with either niche, emerging, or declining themes in the sample.

Finally, the small leftmost cluster (purple) includes only three authors: Paola Manzini, Kfir Eliaz, and Yusufcan Masatlioglu. As in the previous two cases, this cluster is either niche, emerging or declining. Here there is a predominance of applied, context-specific individual behavior modeling, but little more unifies them. Therefore, it is probable that this is the cluster that describes authors specialized in the applications and case studies of Bounded Rationality.

In the following, we present the co-citation of references of Sample 1.

Co-citation Network of References

Figure III – Co-citation Network of References of Sample 1



Source: own elaboration, generated in *biblioshiny*.

Co-citation of references functions much like co-citation of authors. However, the analysis is much more specific. Co-citation of references, as the name suggests, displays the clusters that form around references associated with specific themes. As such, by congregating clusters of references, research themes also become discernible.

Five clusters can be identified: cluster 1.1 (purple) is centered around the paper Rubinstein (1998) and its main theme is the study of models of Bounded Rationality; cluster 1.2 (green) is centered around Kahneman (2017 [2011]) and its main theme is the study of uncertainty in decision-making; cluster 1.3 (blue) is centered around Kahneman and Tversky (2013 [1979]) and reflects the basic themes of BE; cluster 1.4 (yellow) is centered around Evans (2001b), and cluster 1.5 (red), centered around Hommes (2013), discusses the theme of ‘Behavioral Macroeconomics’.

Table VI – Clusters in the Co-citation of References in Sample 1

Cluster	Main Theme	Central Reference	References to Herbert Simon	Predominant Form of Bounded Rationality Discussed
<i>Cluster 1.1 (purple)</i>	Modeling Bounded Rationality	Rubinstein (2011)	Simon, 1955; Newell and Simon, 1972.	Optimizing Bounded Rationality
<i>Cluster 1.2 (green)</i>	Decision Uncertainty	Kahneman (2017 [2011])	Simon 1955	Biased Bounded Rationality
<i>Cluster 1.3 (blue)</i>	Basic Themes	Kahneman and Tversky (2013 [1979])	Simon, 1955; Simon, 1956; March and Simon, 1993 [1958])	Undetermined
<i>Cluster 1.4 (yellow)</i>	Behavioral Macroeconomics	Evans and Honkapohja (2001b)	None.	Optimizing Bounded Rationality
<i>Cluster 1.5 (red)</i>	Behavioral Macroeconomics	Hommes et al. (2013)	None.	Optimizing Bounded Rationality

Source: Own elaboration.

In Sample 1, *A Behavioral Model of Rational Choice* (Simon, 1955) is the key reference to Herbert Simon. In this paper, the author proposes a critique of Substantive Rationality in line with the concept described in Section I.4: “Broadly stated, the task is to replace the global rationality of economic man with a kind of rational behavior that is compatible with the access to information and the computational capacities that are actually possessed by organisms” (Simon, 1955, p. 99). This critical vein, along with the methodology applied in this paper, will define much of the evolution of Bounded Rationality, for even in this early moment, we can glean the first steps of the methodology that would become Optimizing Bounded Rationality – mathematical modeling based on rational choice theory, but changing key aspects of it. If one paper can be pointed to as a precursor to Optimizing Bounded Rationality, it is this. However, we should note that this was only *one* paper of Simon, in a vast array of other works of his. Therefore, it is correct to point out that this was *one* of his methodologies – not *the* methodology.

Human Problem Solving (Newell and Simon, 1972), another key reference, is also worthy of note, for it develops a general framework for decision-making. In line with the

explanation of Section I.2, including mention to ‘task environments’, ‘symbolic logic’, and ‘AI’¹¹². Other references to Simon include *Organizations* (Simon, 1993 [1958]) and *Rational Choice and the Structure of the Environment* (Simon, 1956).

What is most notable, however, in this table is that no article of Simon’s is in the center of any cluster. The main cluster (1.1), for instance, is centered around Rubinstein’s (1998) *Modeling Bounded Rationality*, a proposal that is very different from Simon’s. In Rubinstein’s words: “The choice of the title ‘modeling Bounded Rationality’ rather than ‘models of Bounded Rationality’ or ‘economic models of Bounded Rationality’ emphasizes that the focus is not on substantive conclusions derived from the models but on the tools themselves” (Rubinstein, 1998, p. 1). In a way, Rubinstein’s subverts the ‘Adaptable Conceptual Definitions’ trait of Simon’s hardcore, proposed in Section I.6.

Daniel Kahneman (2017 [2011]) and Amos Tversky (Kahneman and Tversky, 2013 [1979]) are also central to the current research in Bounded Rationality. This contribution, therefore, is one of the byproducts of the H&B research program, specially in clusters 1.2 and 1.3. Other notable nodules here are Kahneman (2003) and Tversky and Shafir (1992).

The dissent of Gerd Gigerenzer and co-authors also make its mark in Sample 1 in Gigerenzer and Todd (1999) and Gigerenzer and Selten (2001). The critique here follows the basic tenements presented earlier: Biased Bounded Rationality depends on normative neoclassical theory and underplays the role of the environment. The experimental approach is also criticized by Gigerenzer. As Dekker and Remic (2019) summarize: to Gigerenzer, “experiments need to be performed ‘outdoors’ and not in the context-less lab”¹¹³ (p. 297).

Cluster 1.2 congregates three classic works: *Risk, Uncertainty and Profit* by Frank Knight (1921), *Theory of Games and Economic Behavior* by John von Neumann and Oliver Morgenstern (1944), and *The Foundations of Statistics* by Leonard Savage (1972 [1952]). Both Knight and Savage famously differentiated between the situation of “knowable” and “unknowable” uncertainty. This is formalized by Dequech (2011) as, respectively, “weak” and “strong” uncertainty. The famous “Von Neumann-Morgenstern Utility” is the result of this

¹¹² “One must work with task environments in which artificial intelligence has provided the requisite array of plausible mechanisms” (p. 6).

¹¹³ Curiously, Gigerenzer does not comment on Vernon Smith’s experimental methodology, which is also closely related with laboratory experiments.

investigation, where an agent chooses taking account of *expected* utility – not the actual final utility of a decision. In a sense, these references may be understood as the “pre-history” of Bounded Rationality, the conceptual foundations that will be mobilized to form it.

Clusters 1.4 and 1.5 are interesting for its focus on equilibrium macroeconomics (Woodford, 2003; Brock and Hommes, 1997; Hommes et al., 2013). Curiously, however, no work of Akerloff, associated with Behavioral Macroeconomics (Tomer, 2007), was part of this cluster. In line with Optimizing Bounded Rationality, these researchers are forming clusters outside the dominant approach, even if they are methodologically consistent with them.

In Sample 1, Cars Hommes was the most relevant author (with the higher number of publications), with a total of eleven articles published between 2013 and 2023, almost double that of the second places¹¹⁴. The relevance of Hommes and the fact that works linked to his were not part of the main clusters may suggest that a new approach to BE is taking shape – one focused on macroeconomics and Optimizing Bounded Rationality. It remains to be seen if these clusters will evolve in discernible research orientations, or, in time, they will be congregated into the main clusters.

III.4 – Summary of Sample 1

Sample 1 reveals a perceivable influence of Herbert Simon in current literature, even if not a decisive one. Simon is the most co-cited author in Sample 1, but the analysis of references indicates that his influence was overshadowed by other authors. Names like Daniel Kahneman, Ariel Rubinstein, George Evans, Seppo Honkapohja, and Cars Hommes are the central references in the clusters that discuss Bounded Rationality in the current literature. Consequently, Simon’s presence is felt, but his influence is more tangential than the common narrative suggests.

The most influential work of Herbert Simon in current literature is *A Behavioral Model of Rational Choice*, present in clusters 1.1, 1.2, and 1.3. This central paper determines the tonality of the integration of Bounded Rationality into economics, as his chosen methodology there is the formalization of a mathematical model of Bounded Rationality. This model alleviated some of the more strict elements of traditional rationality, but the general neoclassical

¹¹⁴ These are Victor Hugo Aguiar, C. N. McKinney Jr., and Josep Lusthenhouwer, with 6 articles each.

framework remained (roughly) the same as before. *A Model of Rational Choice*, therefore, can be seen as the basis in which Optimizing Bounded Rationality, such as Rubinstein's (1998), will be developed.

Organizations also was an influential work, but only in cluster 1.3 (basic themes). As such, the influence of Simon's administrative theories is not seen in the "specialized" clusters. This indicates that *Organizations* and organizational behavior are two of the more "transversal" themes of Simon, along with AI and cognitive psychology. Unlike *A Model of Rational Choice*, no specialized literature sprung from these themes, and, consequently, no discernible cluster based on them formed. Due to limitations of the sample (specifically the sample's focus restriction to Bounded Rationality), we cannot conclusively say that the themes of organizational behavior were not developed somewhere. In any occasion, we can conclude that the methodology of Bounded Rationality was not applied to study the themes of organizational behavior in current literature.

The leading publishing journals demonstrate a clear tendency toward orthodoxy, even if the leading journal is heterodox. This is also reflected in the co-citation of references – no cluster presented 'Adaptive Bounded Rationality' as their main conceptual framework. Consequently, the form of Bounded Rationality that more closely relates with Simon (and the work of Gigerenzer, Vernon Smith, Peter Earl, etc.) is also tangential in the sample.

If Sample 1 presents a "leading figure", this would be Daniel Kahneman. Even if Kahneman is less co-cited than Simon (Figure II), his influence is both wider and deeper than Simon's in current literature (Figure III). Kahneman is closely tied with the H&B approach, and his version of Bounded Rationality is the 'Biased' one. This concept of Bounded Rationality is not properly "neoclassical", but as Gigerenzer and Berg (2010) suggests, they share a "common language" that makes it possible for Kahneman's Bounded Rationality to be capable of being applied to neoclassical economics.

However, even if Kahneman is the most influential author, the approach of 'Optimizing Bounded Rationality' was the leading concept: from the five clusters of Sample 1, three mainly focused in this version of Bounded Rationality. The chief figure here is Ariel Rubinstein, followed by the researchers who congregate around the work of Cars Hommes, possibly a "rising star" in the research of Bounded Rationality. This shift towards orthodoxy, however, cannot be interpreted as a phenomenon exclusive to Bounded Rationality or BE – this is part of a slow process, initiated in the 2010s, which affected the entirety of economics.

In summary, the data presented in this chapter permits us to affirm that the “Bounded Rationality” nowadays is very different from that of Simon. More specifically, the bibliometric evidence allows us to infer how economics is discussing Bounded Rationality nowadays, if compared to Simon’s original concept. Sample 1 indicates that few elements of Simon’s hardcore were maintained. Only the focus on decision-making and the (declared) focus on “empirical research” decisively remained. Besides small hints, the discussion of ‘artificiality’ is invisible in Sample 1. The matters of interconnection (or ‘separateness’ in Tomer’s, 2007, terms) among other sciences are wholly dependent on the approach itself – there are approaches connected to other sciences; there are approaches that do not attempt anything of the sort – but a tendency towards more insulated research approaches can be gleamed.

The biggest change, however, is the gradual substitution of the ‘decision-making relativism’ by a framework based on neoclassical economics. This, we argue, is the byproduct of Simon’s lack of a cohesive normative theory (refer to Section II.1) and a general leaning of economics towards the mainstream, alluded to in Section I.5. In short, if no alternative normative theory presents itself, economists need either to strive to formulate them (this is Gigerenzer’s and Smith’s approach) or to adopt a familiar theory as a normative basis (this is Kahneman’s and Rubinstein’s approach). The last option, Sample 1 suggests, was the more popular one.

Far from irrelevant, however, Simon remains a distinguished figure, but his methodological influence is shallower than some of the literature suggests. Running the risk of excessively simplifying the matter, we propose an analogy: similar to Kahneman and Tversky’s ‘nod’ at Simon (Lopes, 1992) – a recognition of a “distinguished figure [rather] than an acknowledgment of a significant intellectual link” (Gigerenzer, 2020, p. 60) – the economics of the last decade seem also to ‘nod’ to Simon, recognizing the importance of his Bounded Rationality, but preferring other versions of it.

III.5 – Data analysis of Sample 2

Sample 1 analyzed the broad use of Bounded Rationality; Sample 2 is more focused, assessing the research that studied both Bounded Rationality and Simon. Consequently, Sample 2 is a subset of Sample 1. The aim here is to identify the discussion around Herbert Simon himself, only glimpsed in Sample 1.

From the 707 documents of Sample 1, only 28 included the term “Simon” in the paper’s title, abstract, or in the keywords – approximately 4% of Sample 1. These papers were collected from 22 different sources and included a total of 39 authors. This sample presented a 0% annual growth rate, and the average citation by paper is 6.71, almost half of the 13.68 from Sample 1¹¹⁵. Table VII presents the leading publishing journals.

¹¹⁵ Due to the low number of articles, some journals in the top ten only published one article. Taking it into consideration, another form of ranking was considered, but ultimately not used. CAPES (*Centro de Aperfeiçoamento de Pessoal de Nível Superior*) is a Brazilian institution that provides a ranking of journals, the *Qualis* system. It was considered to use *Qualis* as a criterion to define the top ten; however, some important rankings were missing, and this alternative would make for an incomplete ranking. *Qualis* is a ranking that is divided hierarchically into nine categories: A1, A2, A3, A4, B1, B2, B3, B4, and C, similar to the ranking system described by Lee et al. (2010) and, indeed, is cited in it. *Mind and Society*, *Economic Thought*, *Economics and Sociology* and *German Economic Review*, some of the leading publishers in this sample, were not evaluated by CAPES, making it difficult to articulate both the number of articles (*bibliometrix* criteria) and their *Qualis* ranking. It also should be noted that the *Journal of Economic Methodology*, the second most relevant source, was evaluated as an A3 journal. This would put the *Journal of Economic Methodology* behind journals with fewer papers but a higher ranking, like the *European Journal of the History of Economic Thought* (A2). Due to these reasons, and to avoid unnecessary confusion, the ranking was kept as generated by *bibliometrix*.

Source: <https://www.gov.br/capes/pt-br/centrais-de-conteudo/documentos/avaliacao/avaliacao-quadrienal-2017/DocumentotecnicoQualisPeridicosfinal.pdf>

Table VII – Leading Publishers in Sample 1 (2013 – 2023)

Journal	N. of Papers	Intellectual Tradition	Additional Information
<i>Mind and Society</i>	4	Orthodox**	Journal founded by Herbert Simon. Closely follows his approach.
<i>Journal of Economic Methodology</i>	3	Heterodox**	Focused on ‘Big-M Methodology’ (philosophy of science).
<i>Computational Economics</i>	2	Heterodox**	Focus on computer sciences.
<i>Brazilian Journal of Political Economy</i>	2	Heterodox*	Focused on the scientific development of Latin America, but not limited to it in theme.
<i>Economic Thought</i>	1	Heterodox**	Interdisciplinary journal that incentivizes publications outside the mainstream.
<i>Economics and Sociology</i>	1	Heterodox*	Focused on political sciences, cultural sociology, gender studies, philosophy of economics, etc.
<i>Estudos Econômicos</i>	1	Pluralist**	“[...] this journal presupposes the respect to the plurality of ideas, themes, arguments, theories, and methods shared by the economic scientific community” ¹¹⁶ .
<i>European Journal of the History of Economic Thought</i>	1	Pluralistic**	General interest and pluralistic, but favors research in the history of economic thought.
<i>Forum for Social Economics</i>	1	Heterodox**	Focused on socioeconomics and promotes alternative research to orthodoxy.
<i>German Economic Review</i>	1	Pluralistic**	General interest, but favors Political Economy.

Source: own elaboration.

* Classification of Lee et al. (2010).

** Own classification, made with the review of the aims and scopes of each journal. Journals that were explicitly open to either orthodox or heterodox, or described themselves as ‘pluralistic’, were considered so.

¹¹⁶ Source: <https://www.scielo.br/journal/ee/about/#item-1>

Mind and Society is one of the most notable sources of this sample, for this journal was founded by Herbert Simon himself¹¹⁷. As such, *Mind and Society* is closely related to the characterization of Simon proposed in Chapter I. The aims and scope of this journal were reproduced below:

*“Mind & Society is a peer-reviewed journal founded by Herbert Simon and other cognitive scientists, for ideas, analyses, discussions and investigations on the interaction between the human mind and the societal and economic environment. Minds and environments are broadly constructed to include entities such as AIs, businesses and governments. The focus of the journal is on adaptive behaviours in situations characterized by risk and uncertainty, as well as complexity and unpredictability. Priority themes are bounded and ecological rationality, problem solving and embodied cognition”*¹¹⁸.

Fittingly, this journal is one of the leading publishers of specialized literature in both Bounded Rationality and Herbert Simon in Sample 2. *Mind and Society* was also part of Sample 1, but stayed immediately behind the *Journal of Advanced Transportation*, in eleventh place, and, consequently, was left out of that analysis. *Computational Economics* and *Mind and Society*, therefore, were the only journals in both samples that mention computer sciences or AI. Besides the occasional mention of Agent Based Models (ABMs), little attention is given to the central role that AI played in Simon’s version of Bounded Rationality.

Due to its closeness to Simon’s research interests, *Mind and Society* is a special case in this classification as either ‘orthodox’ or ‘heterodox’, similar to the difficulty in classifying Simon in those two categories. We can classify *Mind and Society* in an analogous manner to Simon – negatively heterodox but closer to the mainstream in a positive classification. The classification of Lee et al. (2010), therefore, will tend to approximate *Mind and Society* with orthodox economics¹¹⁹ and, as such, we can classify *Mind and Society* as an orthodox journal, at least in the sense of not explicitly contributing to heterodox economics, proposed in Lee (2009) and Lee et al. (2010).

In any occasion, Sample 2 reveals almost a complete inversion of the results of Sample 1: eight journals in the sample are heterodox or pluralistic journals, while two journals lean closer to mainstream economics. While Sample 1 reveals that orthodoxy was interested in

¹¹⁷ Source: <https://link.springer.com/journal/11299/aims-and-scope>

¹¹⁸ *Ibid.*

¹¹⁹ The same method was applied to *Computational Economics*.

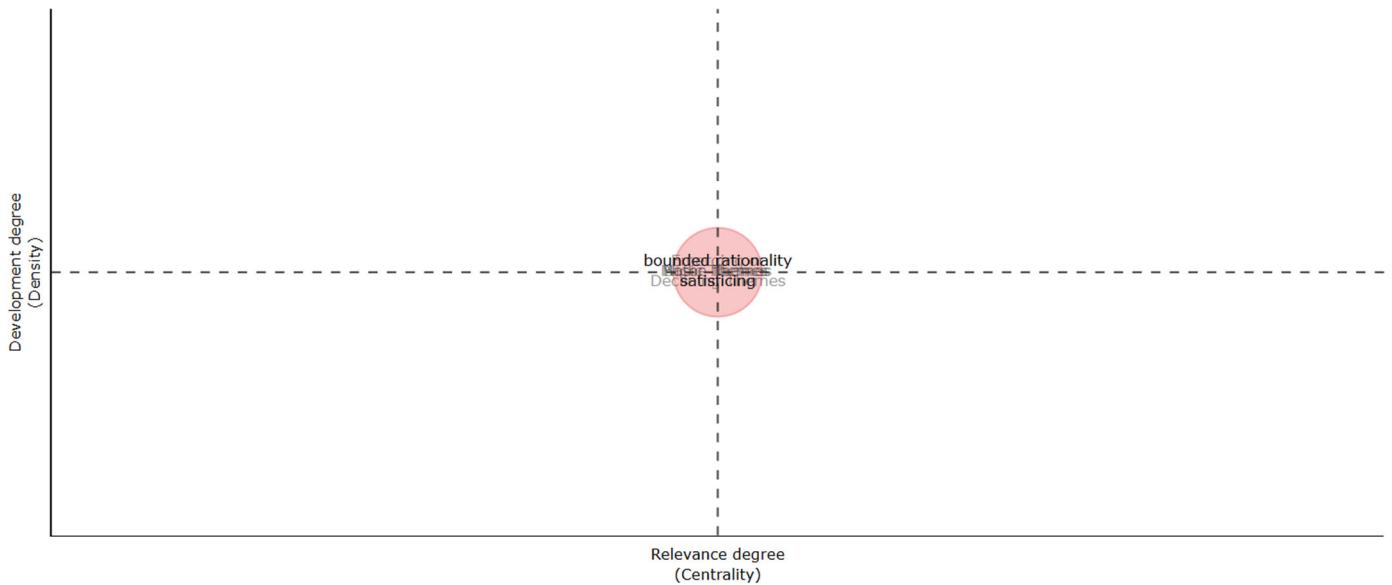
Bounded Rationality, Sample 2 demonstrates that heterodoxy has a greater interest in Herbert Simon himself. The ranking of *Mind and Society* as the top publisher in Sample 2 indicates the editorial body is closer to the work of Simon, but is, nonetheless, a small minority in the current broad discussion around Bounded Rationality.

We also notice a stark contrast between the discussion assessed in Sample 1. While Sample 1 was closely tied with behavioral modeling, Sample 2's journals showed a profile of interest in pluralistic economics, economic philosophy, and scientific methodology. The inclusion of Herbert Simon in many journals specialized in Political Economy also are worthy of mention, due to Political Economy's tendency towards historical methods and a greater adherence of political economists to heterodoxy.

We also notice a stark contrast between the discussion assessed in Sample 1. While Sample 1 was closely tied with behavioral modeling, Sample 2's journals showed a profile of interest in pluralistic economics, economic philosophy and scientific methodology. The inclusion of Herbert Simon in many journals specialized in Political Economy also are worthy of mention. Studies in Political Economy tend to be historical and adhere more closely to heterodoxy. In summation, the tendency towards orthodoxy in Sample 1 is reversed in Sample 2. Greater importance is placed in the methodological discussion and economic social aspects, a shallow theme in Sample 1.

Thematic Map of Sample 2

The analysis of the thematic map of Sample 2 is analogous to that of Sample 1 (refer to Section III.2).

Figure IV – Thematic Map of Sample 2

Source: own elaboration, generated in *bibliometrix*.

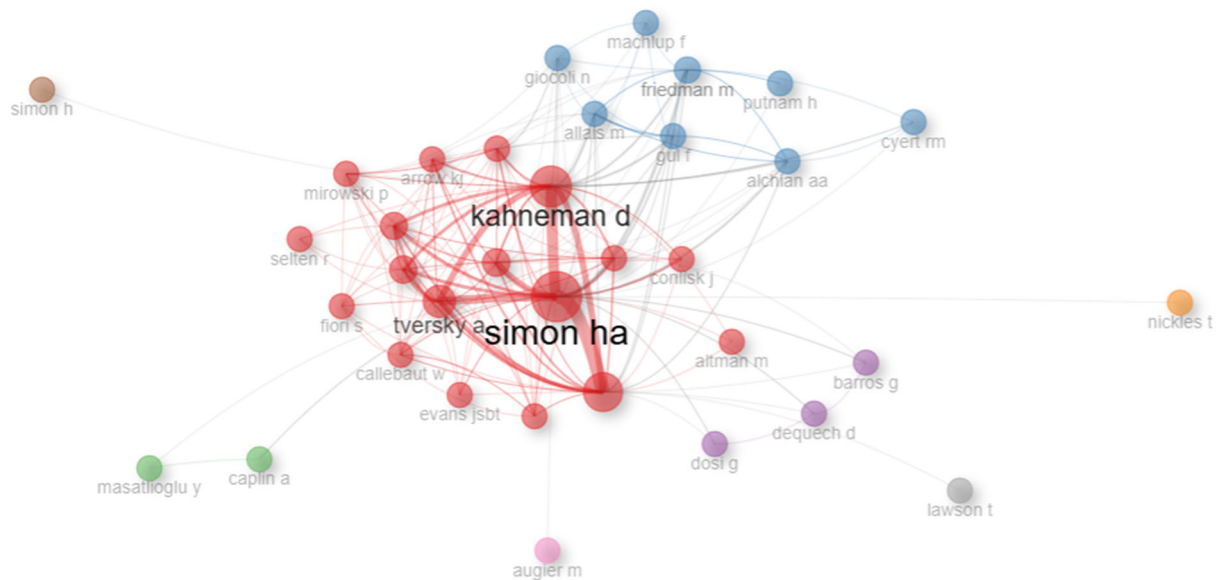
Differently from the thematic map of Sample 1, Sample 2 only presents one cluster with the words ‘bounded rationality’ and ‘satisficing’ at its center. This is due to the low number of articles present in this sample (only 28). The fact that there is only one cluster impedes any meaningful assessment of the themes of Sample 2. The localization of the cluster in the center is due to a lack of other clusters for it to be compared to. Consequently, the cluster of the thematic map of Sample 2 is in a neutral “zero-space”. By all means, this is a spurious result: the sample that discusses Bounded Rationality has one cluster centered around Bounded Rationality. The thematic map was included to maintain the structure of the bibliometric analysis, but it does little to help us interpret Sample 2.

III.6 – Data visualization of Sample 2

The visualizations are analogous to those of Section III.3.

Co-citation Network of Authors of Sample 2

Figure V – Co-citation Network of Authors of Sample 2



Source: own elaboration, generated in *bibliometrix*.

As in Sample 1, the main nodule of the main cluster (red) is also Herbert Simon. Also, like Sample 2, Simon is the most co-cited author (but not necessarily the most influential). Some important names are also repeated from the co-citation of authors in Sample 1: Daniel Kahneman, Amos Tversky, Gerd Gigerenzer, Ariel Rubinstein, and James March. This cluster also includes Alan Newell (one of Simon's main co-authors), Esther-Mirjam Sent (a behavioral economist closer to heterodoxy), D. Wade Hands (a methodologist of economics), Hunter Crowther-Heyck (one of the few economists interested in the history of Simon's economic thought), Stefano Fiori (a main reference to this research and one of the few researchers that stressed the role of AI in Simon's work), Phillip Mirowski (another main reference of this dissertation and the proponent of Simon as a 'cyborg scientist'), among other scientists. The renowned economist Kenneth Arrow is also part of this cluster, even if his contribution is not directly linked with BE.

The upper-right cluster (blue) includes names such as Milton Friedman and Richard Cyert, already mentioned earlier. This cluster also includes researchers such as Faruk Gul (a mainstream-aligned economist who researches behavior modeling), Armen Alchian (a classical

author of microeconomics, credited as one of the precursors to the New Institutional Economics¹²⁰), Maurice Allais (an influential macroeconomist and Nobel Prize winner in 1988¹²¹) and Nicola Giocoli (a researcher in behavioral modeling). This cluster, therefore, is closer to traditional micro and macroeconomics than to BE itself.

Two smaller clusters also can be identified. The lower-right cluster (purple) includes three authors that were already cited: Giovanni Dosi, David Dequech, and Gustavo Barros. Dosi is a famous researcher in evolutionary economics. These authors are interested in the effects of uncertainty in decision-making, specifically on procedural rationality (Dosi, 1982; Dosi and Egidi, 1991; Dosi, Marengo, and Fagiolo, 2001), as is Dequech (1997; 2001; 2006; 2011¹²²) and Barros (2010). The lower-left cluster (green) includes Andrew Caplin (also preoccupied with behavioral modeling) and the already mentioned Yusufcan Masatlioglu. Analogous to the “applications cluster” of the co-citation of authors in Sample 1, this cluster also congregates researchers mainly engaged with the application of Bounded Rationality in case studies.

Finally, four outliers were generated: Mie Augier (pink), Tony Lawson (gray), Thomas Nickles (yellow)¹²³. Mie Augier was already cited in this dissertation and was one of the main influences in this research. Her interests are in the bibliographical details of Herbert Simon (Augier and Kreiner, 2000; Augier, 2000; 2001a; 2001b; Augier and March, 2003; 2004), and her relative isolation from the other clusters indicates that her work impacted little on the overall research effort. Tony Lawson is a famous philosopher of economics, engaged with research in critical realism, but is not particularly interested in Bounded Rationality. Therefore, Lawson integrates this debate in the broader sense of assessing neoclassicism’s “realism”, and what are the implications of it are for economics science. Thomas Nickles, the final author, is a philosopher of science, not specialized in economics, engaging instead in the broad themes of scientific development. Like Lawson, however, his work is not directly linked with the question at hand.

¹²⁰ Source: <https://www.corecon-rj.org.br/portal/interna2.php?i=1485356410/armen-a-alchian-1914>

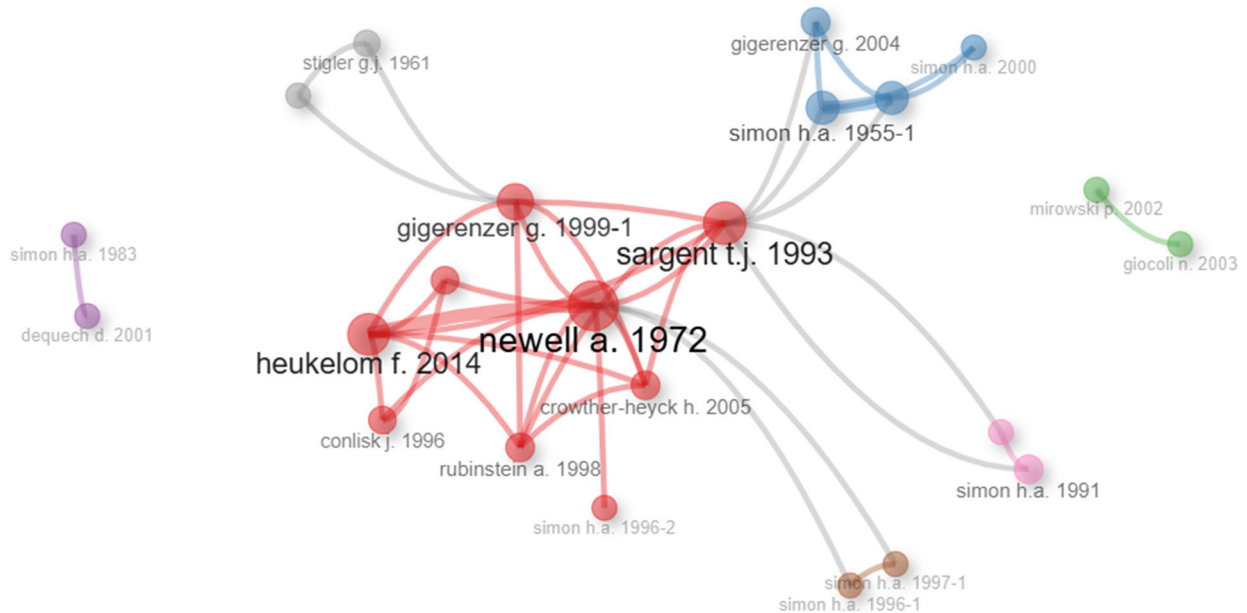
¹²¹ Source: <https://www.britannica.com/money/Maurice-Allais>

¹²² Dequech (2011) is not cited in the sample, but goes in the same direction as the other references.

¹²³ The “simon h” cluster (brown) also refers to Herbert Simon, but groups articles that cited “H. Simon” instead of the more correct “H. A. Simon”. As such, this cluster is part of the central “simon ha” nodule, separated due to a software limitation. Hence, it requires no further analyses.

Co-citation of References of Sample 2

Figure VI – Co-citation Network of References of Sample 2



Source: own elaboration, generated in *biblioshiny*.

Figure VI represents the co-citation network of Sample 2. Seven clusters can be identified. Two main clusters formed: cluster 2.1 (red), centered around Newell and Simon's (1972) *Human Problem Solving*, and cluster 2.2 (blue), centered around *A Behavioral Model of Rational Choice* (Simon, 1955). Other three small clusters were generated, all linked to the main one: cluster 2.3 (grey) centered around Kahneman and Tversky (2013 [1979]) and Stigler (1961); cluster 2.4 (pink), centered around *Organization and Markets* (Simon, 1991); cluster 2.5 (brown), centered around *Models of My Life* (Simon, 1996 [1991]) and the 1997 edition of *Administrative Behavior* (this book is cited as Simon 1976 [1947] in the references). Finally, two outlier clusters, not linked to the main clusters, can be identified: cluster 2.6 (purple), centered around *Why Should Machines Learn* (Simon, 1983) and Dequech (2001); and cluster 2.7 (green), which includes Mirowski (2002) and Giocoli (2003)¹²⁴.

¹²⁴ *Bibliometrix* generated an eight cluster, containing references to Esther-Mirjam Sent and Herbert Simon. However, this cluster was left out of the analysis due an error in the *Scopus* database: the year was not specified

Table VIII – Clusters in the Co-citation of References in Sample 1.

Cluster	Main Theme	Central Reference	Additional References to Herbert Simon	Predominant Form of Bounded Rationality Discussed
<i>Cluster 2.1 (red)</i>	Basic Themes	Newell and Simon, 1972	Simon, 1996 (cited as Simon, 1969 in the references).	Adaptive Bounded Rationality
<i>Cluster 2.2 (blue)</i>	Original Bounded Rationality	Simon, 1955	Simon, 1956; Simon, 2000	Adaptive Bounded Rationality
<i>Cluster 2.3 (grey – niche)</i>	Mainstream impact of Bounded Rationality	Kahneman and Tversky (2013 [1979])	None.	Biased Bounded Rationality
<i>Cluster 2.4 (pink – niche)</i>	Simon's Administrative Theory	Simon (1976 [1947])	Simon (1991)	Adaptive Bounded Rationality
<i>Cluster 2.5 (brown – niche)</i>	Simon's Bibliography	Simon (1996 [1991])	Newell and Simon (1972)	Adaptive Bounded Rationality
<i>Cluster 2.6 (purple – outlier)</i>	Unable to be classified.	Simon (1983)	None.	Adaptive Bounded Rationality
<i>Cluster 2.7 (green) – outlier</i>	History of Bounded Rationality	Mirowski (2002)	None.	Adaptive Bounded Rationality

Source: Own elaboration.

While the greater influence of Simon in Sample 2 is a spurious result (the search was specifically tailored to identify his influence), the profile of the research work is notable. Firstly, the relevance of *A Behavioral Model of Rational Choice* is considerably diminished. In its place, there is a more diverse set of influences, like *Sciences of the Artificial* (Simon, 1969), *Administrative Behavior* (Simon, 1976 [1947]), and even *Models of My Life* (Simon, 1996 [1991]). Even the obscure article *Why Should Machines Learn?* (Simon, 1983) is part of Sample 2. This shows that while in Sample 1, the influence of Simon (1955) was symptomatic of the

(it only showed “simon h.a.”, “sent e.m.”). For this reason, it was impossible to identify which of Sent's or Simon's articles this cluster was referring to.

research focus on models of Bounded Rationality, the specialized literature is much more diverse. Sample 2 reveals a more pluralistic set of preoccupations, including Political Economy, philosophy of economics, socioeconomics, etc. In fact, curiously, the effort to ‘model Bounded Rationality’ becomes secondary when the search is focused in Herbert Simon. This is unexpected, due to Simon’s focus in models.

Secondly, the influence of both Kahneman and Rubinstein is much less felt. Given the retrospect made in Chapter II, this result is not surprising; however, it should be noted that the distance between the Bounded Rationalities of Kahneman and Simon is not common sense in economics. As such, the literature specialized in Simon does not echo the common narrative of a straight lineage between his concept of Bounded Rationality and Kahneman’s¹²⁵. The same can be said about Rubinstein and Hommes¹²⁶, relevant authors in Sample 1, but they almost don’t show up in Sample 2.

Thirdly, a much more diverse assortment of authors is part of Sample 2. Examples that were mentioned previously in this dissertation are Augier (2000), Dequech (2001), Mirowski (2002), etc. Other authors worthy of note are Heukelom (2014), who writes a book that chronicles the history of BE (although not focusing on Simon); Petracca (2017), which explores the link between Simon’s *cognitivism* (symbolic information processing) and his concept of Bounded Rationality – the “road not taken” (Petracca, 2017, p. 33) in his words; and Giocoli (2003), which, like Mirowski (2002), analyzes the role that the Cold War played in the intellectual development of Herbert Simon, but focusing on the impacts of Bounded Rationality in game theory.

III.7 – Summary of Sample 2

In zooming in on the mentions of Herbert Simon in Sample 1 (Sample 2), we can observe how specific research in Simon himself is being made. Contrary to the results of Sample 1, Sample 2 was mainly published in heterodox and pluralistic journals, often focused in either the philosophy of science, philosophy of economics, or correlated themes. The presence of a discussion around Simon in journals of Political Economy is of special note, for Simon himself

¹²⁵ Kahneman, along with Stigler (1961), only show up in cluster 2.3 and nowhere else in Sample 2.

¹²⁶ Rubinstein only appears tangentially in cluster 2.1. Hommes does not appear in Sample 2.

contributed little to this field. Moreover, we note a steep decline into the two main research focuses from Sample 1: economic modeling of Bounded Rationality and applications of Bounded Rationality. In their place, the discussion was dominated by theory and historical analysis.

The presence of other central authors of Sample 1, such as Kahneman, Tversky and Rubinstein, is felt much less in Sample 2, while Simon himself was both central and influential in Sample 2 – the opposite of Simon’s position in Sample 1. There, Simon was co-cited a high number of times, but his influence wasn’t as deep as expected, favoring other interpretations. Sample 2, on the other hand, gives a central role to Simon’s interpretations of Bounded Rationality. As a chief evidence of this shift, we can cite the influence of *Mind and Society* in this sample, a journal that closely follows the research program of Herbert Simon himself, which did not made the top 10 in Sample 1, but was the publishing leader in Sample 2.

No one work of Simon’s can be said to be the most influential in Sample 2. As such, *A Behavioral Model of Rational Choice*, the most influential paper of Simon in Sample 1, was only one among many. Consequently, the “tonality” of Sample 1 – a discernible effort to model and “mathematize” Bounded Rationality – shifts towards a more wide preoccupation with history, methodology, and Political Economy in Sample 2.

III.8 – Interpretation of the results

Samples 1 and 2 tell two very different stories. The first, as a wide sweep of research in Bounded Rationality, favors versions of this concept that are different from those of Simon himself. Taking the characterization presented in Chapter II, recent mainstream favors either Optimizing Bounded Rationality or Biased Bounded Rationality. In any occasion, the Optimizing variant seems to be gaining traction in recent years.

For the time, however, the most relevant author to BE is still Daniel Kahneman, even if Simon is still the most recognizable. We make this affirmation mainly observing the contrast between co-citation of authors and co-citation of references: Simon is the most co-cited author in both samples, but his influence is less apparent; Kahneman is less co-cited, but his influence is perceivable in the three main clusters of Sample 1. Ariel Rubinstein is also a promising contender for the rank that Kahneman now occupies, already being the most influential author

in the main cluster of Sample 1. His work, however, is more recent than Kahneman's, having less time to mature in the economic literature. What is undeniable is that Rubinstein's formula of mathematical modeling of Bounded Rationality became very popular in the mainstream.

In a similar and related note, this bibliometric review may indicate the starting steps of a new approach. While embryonic as a discernible approach, the works of George Evans, Cars Hommes, Michael Woodford, and William Brock form discernible clusters in many of the co-citation maps. Cars Hommes is of special note, being the researcher with the highest number of published articles in the period analyzed. The characteristics of this group, so far, are a focus on macroeconomics and a close relation with the work of Thomas Sargent (1993). This, we argue, possibly reveals a growing interest in 'Behavioral Macroeconomics', even if this is not linked to Akerlof's approach. Note also that these are not emerging or low-developed themes – in ten years, this new literature established itself and began to form discernible clusters of research. Further details can only be uncovered in time.

The results of this analysis, therefore, suggest that Simon plays a far smaller role in research in Bounded Rationality than the literature suggests. Simon, while being much mentioned, has little actual influence on the current research of Bounded Rationality. That, again, is not to say that Simon lost its relevance in the current mainstream, but the bibliometric data reveals that his version of Bounded Rationality is not the one being used in economics – even if some economists think that this is the case.

This can be attributed to a "veil" around the concept of Bounded Rationality: economics does not fully recognize the characteristics of each type of Bounded Rationality, which gives the impression that Simon, Kahneman, and Rubinstein are all talking about the same thing – even when, in further inspection, they are not. As such, we favor the alternative hypothesis (H1): Current literature in Bounded Rationality does not closely reflect the work of Herbert Simon.

CONCLUSION OF CHAPTER III

Should we be impressed by the fact that Simon's original approach of Bounded Rationality is not the main focus of economics in recent years? The answer probably is 'no'. Simon's original concept is almost 70 years old at this point – there is no surprise in the fact

that this concept changed through all this time. What is worthy of note in this history, however, is not that the concept “changed” – but how it changed. Effectively, Chapter II suggests that very little of Simon remained in Bounded Rationality after the 1970s – a process that is still happening today. Simon wasn’t forgotten and this is far from the reality of facts – but it is widely clear that economics remember Simon selectively: he is the father of Bounded Rationality, but to some, little else.

It is correct to assume that economics incorporated the work of Herbert Simon into the mainstream? It is hard to argue that this wasn’t the case, but the evidence presented in this chapter reveals that the impact of Simon in the discussion of economic behavior is far more modest than some may seem to assume. In the Kuhnian sense, Simon did not engender a scientific revolution – he was critical of the mainstream, but not radically so. Simon’s contribution also can be appropriated by heterodoxy, as it was, but what becomes clear is that Bounded Rationality is not a revolutionary concept by itself.

Finally, with the “inevitable” return of orthodoxy, preconceived by Davis (2008) and Bianchi (2010), should we be surprised that the more “heterodox” version of Bounded Rationality would lose relevance? Again, probably not. That being said, Bounded Rationality was a pluralistic concept in its origin, but became gradually less so with its conceptual development. As such, Bounded Rationality does not need to be revolutionary to be valuable to pluralism and heterodoxy. The “lost potential” of Bounded Rationality, therefore, is not that it developed approaches that are (more) in line with neoclassicism – but that it almost exclusively was used in this manner. There is potential for Bounded Rationality to be an important tool in a pluralistic economic science – but the untapped potential of alternative concepts of Bounded Rationality deprives this concept of this important role. It remains to be seen if Simon’s legacy can be once again a beacon of pluralism, or if its fate is otherwise.

5. CONCLUSION – A Map to the Maze?

I

In this dissertation we analyzed the inception of Bounded Rationality in the work of Herbert Simon and its evolution in the economic sciences. Chapter I presented a lengthy reconstruction of Simon's development as a scientist, focusing on the author's idiosyncrasies that differentiate them from other behavioral scientists and economists. We focused on two main conceptions that permeated effectively all of his work: artificiality and his 'monomania' about decision-making. In line with the 'cyborg sciences', Simon's conception of artificiality favored the interpretation that human behavior could be understood in similar grounds to machine behavior. AI and computer science are not the sole focus of the economic discussion in Herbert Simon, but they play an important role in his sciences – a role that often is overlooked.

We argued that Bounded Rationality is an objective-centered style of behavior that follows four main steps: 1) the definition of the criteria for "satisfaction" described by Simon as the process of 'satisficing', 2) the definition of a problem space (relation between agents cognitive capabilities and the environment), 3) the implementation of a plan of action, chosen by the agent in view of the specific problem space and 4) the emerging outcome, that takes into consideration the results between the problem space and the chosen plan of action. Simon's theory heavily emphasized the observations of the results of behavior. This, we argue, is due to the influence of the empiricism of logical positivism. Even so, Simon adopted a specific interpretation of this empiricism, which relaxed some of the more strict elements of this type of positivism.

Simon's Bounded Rationality was critical of neoclassical economics, but neither the critique or the repair of neoclassicism was the focus of Simon. Instead, Simon proposed an alternative interpretation of rationality, which differed from neoclassical Substantive Rationality both in its intensity (from perfect to imperfect information, from fully rational to limited rationality, etc.) and its structure (behavior is explained different in both models). This brought important implications for economics, however, other economists, such as Kenneth Arrow, Sydney Winter, Richard Nelson and Oliver Williamson were to be remembered as the

economists that implemented Bounded Rationality into economics proper. Other developments in Behavioral Economics, such as the contributions of Daniel Kahneman, Amos Tversky and Gerd Gigerenzer also tend to be bundled with the contributions of Simon, making it difficult to identify how him specifically contributed in this discussion.

Simon's style of economics landed him in a "superposition" in the orthodoxy-heterodoxy debate. We proposed an interpretation: in the negative definition he is a heterodox economist – but not in the positive definition; he was not an orthodox economist, but also did not made attempts to come closer to heterodoxy. This duality, we hope, is useful to interpret the impact of this author, keeping in mind the influence that he acquired both in orthodox and heterodox economics.

Chapter I ends with a summary of what we propose as the *hardcore* of Simon's economic sciences. The elements are: 1) indistinction between 'natural' and 'artificial'; 2) focus on 'problem-solving'; 3) decision-making relativism; 4) focus on "empirical" observation; 5) takes inspiration on other sciences, but only in matters of decision-making; and 6) tends to adopt flexible and adaptable conceptual definitions. This, we propose, is the 'hardcore' of Herbert Simon.

Based on this definition, we analyzed the differences between Simon's conception of Bounded Rationality, and others conceptions. Chapter II comments some of the definitions. Three main contributions were articulated: the 'two periods' interpretation, Tomer's classification, and the 'three Bounded Rationalities' of Gigerenzer. With this articulation, we identified three main concepts used in current literature: Adaptive Bounded Rationality, Biased Bounded Rationality and Optimizing Bounded Rationality.

With little in common, however, no significant methodological commonalities were found among the three concepts. What they most decisively shared is a 'research ethos': "clearly people are not rational, therefore...". The explanation that comes next depends on which form of Bounded Rationality the researcher in question adheres to.

As such, we conclude that Simon was one of the few, if not the only, behavioral scientist associated with the development of BE that presents explicitly the characteristics of the focus on artificiality and his special interest in computer sciences. In other words, the interdisciplinary elements are not reminiscent of computer sciences. Instead, a greater focus on psychology can be observed. Some examples are the influence of Cognitive Sciences in the works of Daniel

Kahneman and Amos Tversky and the Brunwickian psychology in the works of Gerd Gigerenzer. Moreover, a great degree of endogeneity can be observed in contemporary use of Bounded Rationality – some behavioral economists, such as Ariel Rubinstein, are familiar only with the psychology of BE, not the field of psychology.

The 1970s saw an upsurge of Biased Bounded Rationality due to the Heuristics and Biases program, but this is not the case anymore. The last chapter provided bibliographical evidence about the state of the studies of Bounded Rationality in (roughly) the last decade. Data from 2013 and 2023 reveals much about the studies in Bounded Rationality. Kahneman, and no longer Simon, is the main reference on Bounded Rationality, even if H&B did not incorporate Bounded Rationality at first. Therefore, in this period, ‘Bounded Rationality’ mainly meant ‘Biased Bounded Rationality’ – not ‘Adaptive Bounded Rationality’, which is closer to Simon’s original concept.

Nowadays, however, this is not longer the case. Rather, we can observe an upsurge in ‘modeling Bounded Rationality’, an effort in line with the neoclassical paradigm. Even if Kahneman still is the most relevant author, Ariel Rubinstein, along with his modeling methodology, is slowly rising in relevance. It would not surprise if, in the future, he assumes the position that Kahneman now occupies. Arguably, this phenomenon of “neoclassization” started in the 1970s, but, as Gigerenzer correctly identified, it never ended. Therefore, in recent literature, the “dominant” concept is ‘Optimizing Bounded Rationality’ – also not in line with Simon’s original preoccupations.

As such, this dissertation concludes the following: **contemporary research on Bounded Rationality is not reminiscent of Simon’s methodology.**

II

While commenting on his professor, Alfred Marshall, John Maynard Keynes (1924) defines the “rare combination of gifts” (p. 322) that make a ‘master-economist’:

“He must reach a high standard in several different directions and must combine talents not often found together. He must be mathematician, historian, statesman,

philosopher-in some degree. He must understand symbols and speak in words. He must contemplate the particular in terms of the general, and touch abstract and concrete in the same flight of thought. He must study the present in the light of the past for the purposes of the future. No part of man's nature or his institutions must lie entirely outside his regard. He must be purposeful and disinterested in a simultaneous mood; as aloof and incorruptible as an artist, yet sometimes as near the earth as a politician" (Keynes, 1924, p. 322).

A paradoxical figure, this 'master' exists in "An easy subject, at which very few excel!" (p. 321-322). As Heilbroner (2011) suggest, this 'master-economist' is a mundane thinker (or worldly philosopher), that gazes profundity in the trivial matters of life. The list of 'master-economists' is incredibly short – but is Herbert Simon part of it?

Simon thought big about little things. This dissertation mentioned only briefly his incursion on the cognition to play chess, but he dedicated much to understand this game. He could see great complexity in the simple rules of this. Not only that, he also attempted to systematize his specific findings, even if no attempt to underlie a 'general theory' was made. Unlike Benjamin Franklin, however, Simon did not consider the "morals of chess"¹²⁷, or its capability to elevate humankind. Simon, therefore, saw chess in a mundane manner: a problem to be solved. Unlike Franklin, he sought no greater truth in the chessboard, no hidden meaning in the pieces. Simon more closely resembles the phrase attributed to Millôr Fernandes: "chess is the Chinese game that increases one's capacity to play chess" (own translation). Could there ever be someone as committed to the ordinary as Simon?

That is not to say that Simon was an ordinary thinker – this is not the case. Simon was a deep thinker, yet, his commitment to practicality impeded him to propose a more general, all-encompassing theory of behavior. He, like Marcus Aurelius, "practiced what he preached", as became clear with the comic recollection of Franco Modigliani (2011), when Simon, asked why he always ordered the same lunch, explained that it was to produce less stress on his decision-making apparatus.

Simon was a man of *models*, just as he as a man that studied *behaviors*. The evocative nature of the concepts that he proposed, therefore, seem to be one of the key elements to his legacy. He, we argued, is somewhat of a unique specimen of economist – one so focused in *behavior*, proper, that important economic matters are relegated to footnotes or passing

¹²⁷ Source: founders.archives.gov/documents/Franklin/01-29-02-0608#BNFN-01-29-02-0608-fn-0001

commentaries. Moreover, Simon presented a complex and multifaceted intellectual development. Forged in a unique blend of political sciences, traditional and heterodox economics, empiricism and a deep interest in computer sciences, Simon articulated the many “knowledges” which convinced him.

In a manner of speaking, the interest in many areas that Simon uphold through his life suggest that he never lost the quality of the boy that was fascinated with “idea that human behavior could be studied scientifically” (1996, p. 18). Ever so curious, Simon’s tool of choice to interpret the world was science. Therefore, we should not ask if Simon was a ‘master-economist’, but rather if he was an ‘economist’ at all. Does the designation of ‘cognitive scientist’, perhaps, fit him better than the other term?

III

In any occasion, his contribution to economics was and is still felt – but not in the way often believed. Bounded Rationality changed permanently after the 1970s and, after that, resembles little the original concept of Simon. This is still the case in contemporary economics, but Bounded Rationality further evolved, acquiring meanings other than those that its creator intended.

Bounded rationality is a simple, yet powerful concept. The simplicity gives the impression that Bounded Rationality is “obviously true”, however, this apparent simplicity is due to an ease to attribute meanings to both the ‘bounded’ and the ‘rational’ – not a conceptual lack of depth. In other words, everyone “intuitively” understands what being ‘rational’ is, and can, consequently, imagine what ‘limiting’ this rationality means for behavior. However, this intuition has little bearing on the concept (or concepts) of Bounded Rationality. Its “obvious”, axiomatic nature, therefore, is both a great strength, for it makes the concept adaptive, but also blur the line of what this scientific concept mean.

The ‘cacophony’ that we mentioned in the beginning of this dissertation comes, we think, from this ‘obvious’ nature of Bounded Rationality. In a manner of speaking, ‘Bounded Rationality’ already was born with its many interpretations “baked into it” – when an economist hear ‘Bounded Rationality’ he may think that it entails systematic errors, like Kahneman

suggest; or another form of maximization, as Rubinstein argues. What these key-researchers do, therefore, is to formalize this “intuition” around Bounded Rationality. This effort, as we saw, has a discernible orthodox tint to it. Probably, the biggest casualty that Bounded Rationality suffered is tied to this orthodox appropriation of the concept, for it lost its role as a pluralistic concept. It remains to be seen if heterodox economists can reestablish this important role that Bounded Rationality originally had.

IV

In short, Simon is part of a select group of authors that are known for eye-catching and easily synthesized ideas. Other such authors are Niccolo Machiavelli, Jean-Jacques Rousseau or Adam Smith; reduced to “ends justifying means”, “man being naturally good”, or the “invisible hand”, Simon is commonly reduced to “Bounded Rationality”. He was (and still is) a scholar which has ideas that are easily “borrowed”, for Bounded Rationality no longer comes the philosophical or methodological baggage described in Chapter I. This “intuitive” appropriation makes it easy to implement Bounded Rationality in a number of theories, however, this proves to be a forked road. On the one hand, Bounded Rationality is one of the most successful concepts in the history of economics; on the other hand, it resembles little the original methodology of Simon, even if economists still pay lip-service to him – a practice that does not seem to weaken, even when the influence of Simon keeps being silently replaced.

What we tried to argue in this dissertation is that Bounded Rationality comes with its fair share of ‘baggage’. Applying the concept of Simon with no regard to its history or methodological underpinnings creates the illusion that we, economists that study Bounded Rationality, are part of an intellectual lineage that starts with Herbert Simon. The reality of facts, however, is much more complicated and, we venture saying, more interesting than the common narrative. The point here is not to recollect the “good old days” when Bounded Rationality was Simon’s, but to elucidate the complex heritage of this concept – its many forms, many applications, and many theoretical consequences.

So what is the “true” version of Bounded Rationality? “All of them; for the “real” self is an illusion” (Simon, 1996 [1991], p. XVI).

REFERENCES

- Akerlof, George A. Behavioral macroeconomics and macroeconomic behavior. *American Economic Review*, v. 92, n. 3, p. 411-433, 2002.
- Almeida, Felipe; Paula, Luis Gustavo de. The place of uncertainty in heterodox economics journals: a bibliometric study. *Journal of Economic Issues*, v. 53, n. 2, p. 553-562, 2019.
- Almeida, Felipe; Cruz-e-Silva, Victor; Brites, Maríndia. A bibliometric analysis of the possible convergence of heterodox associations. *Journal of Economic Issues*, v. 57, n. 1, p. 218-240, 2023.
- Angner, Erik. We're all behavioral economists now. *Journal of Economic Methodology*, v. 26, n. 3, p. 195-207, 2019.
- Aoujil, Zakaria et al. Artificial intelligence and behavioral economics: A bibliographic analysis of research field. *IEEE Access*, 2023.
- Aria, M.; Cuccurullo, C. bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, v. 11, n. 4, p. 959-975, 2017.
- Arrow, Kenneth J. Rationality of self and others in an economic system. *The Journal of Business*, v. 59, n. 4, p. S385-S399, 1986. Disponível em: <http://www.jstor.org/stable/2352770>. Acesso em: 24 nov. 2024.
- Augier, Mie. Models of Herbert A. Simon. *Perspectives on Science*, v. 8, n. 4, p. 407-443, 2000.
- Augier, Mie. Sublime Simon: The consistent vision of economic psychology's Nobel laureate. *Journal of Economic Psychology*, v. 22, n. 3, p. 307-334, 2001a.
- Augier, Mie. Simon says: Bounded rationality matters: Introduction and interview. *Journal of Management Inquiry*, v. 10, n. 3, p. 268-275, 2001b.
- Augier, Mie. Will the real Herbert Simon please stand up? Or, Behavioral economics: hopes for the past—lessons from the future. In: Eastern Economic Association Conference in New York City. *Perspectives on Science*, v. 23, 2003.
- Augier, Mie; Kreiner, Kristian. Rationality, imagination and intelligence: some boundaries in human decision-making. *Industrial and Corporate Change*, v. 9, n. 4, p. 659-681, 2000.

Augier, Mie; March, James G. The economic psychology of Herbert A. Simon: Introduction to a special issue. *Journal of Economic Psychology*, v. 24, n. 2, p. 135-141, 2003.

Augier, Mie; March, James G., eds. Models of a man: Essays in memory of Herbert A. Simon. MIT Press, 2004.

Baars, Bernard. The cognitive revolution in psychology. New York: Guilford Press, 1986.

Barros, Gustavo. Herbert A. Simon e o conceito de racionalidade: limites e procedimentos. 2007.

Barros, Gustavo. Herbert A. Simon and the concept of rationality: boundaries and procedures. *Brazilian Journal of Political Economy*, v. 30, p. 455-472, 2010.

Becker, Gary S. The economic approach to human behavior. The University of Chicago Press, 1976.

Brette, Olivier; Lazaric, Nathalie; Vieira da Silva, Victor. Habit, decision-making, and rationality: comparing Thorstein Veblen and early Herbert Simon. *Journal of Economic Issues*, v. 51, n. 3, p. 567-587, 2017.

Blaug, Mark. The methodology of economics: Or, how economists explain. Cambridge University Press, 1992.

Bianchi, Ana Maria. Reflexões sobre o passado e especulações sobre o futuro da metodologia econômica. *Revista Economia*, 2010.

Bianchi, Ana Maria. A ética na economia comportamental: uma breve incursão. *Guia de Economia Comportamental e Experimental*. São Paulo, 2019. Disponível em: <https://www.economiacomportamental.org/guia/>.

Bianchi, Ana Maria; Silva Filho, Geraldo Andrade da. Economistas de avental branco: uma defesa do método experimental na economia. *Revista de Economia Contemporânea*, v. 5, n. 2, p. 129-154, 2001.

Bianchi, Marina. Tibor Scitovsky as behavioral economist. *Journal of Behavioral Economics for Policy*, v. 2, n. 1, p. 39-43, 2018.

Brites, Maríndia; Almeida, Felipe. A economia institucional original no Brasil. *História Econômica & História de Empresas*, v. 26, n. 1, p. 235-260, 2023.

Brock, William A.; Hommes, Cars H. A rational route to randomness. *Econometrica: Journal of the Econometric Society*, p. 1059-1095, 1997.

Callon, Michel; Courtial, Jean Pierre; Laville, Francoise. Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics*, v. 22, p. 155-205, 1991.

Camerer, Colin. Behavioral economics: Reunifying psychology and economics. *Proceedings of the National Academy of Sciences*, v. 96, n. 19, p. 10575-10577, 1999.

de Castro, Alex Sandro Rodrigues. Economia comportamental: caracterização e comentários críticos. Dissertação. [s.n.], 2014.

Crowther-Heyck, Hunter. Herbert A. Simon: The bounds of reason in modern America. JHU Press, 2005.

Cunha, Elcemir Paço. O itinerário epistemológico de Herbert Simon aos limites da racionalidade. *Organizações & Sociedade*, v. 29, p. 614-637, 2022.

Cobo, M. J.; Lopez-Herrera, A. G.; Herrera-Viedma, E.; Herrera, F. An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field. *Journal of Informetrics*, v. 5, n. 1, p. 146-166, 2011.

Colander, David; Holt, Richard; Rosser Jr., Barkley. The changing face of mainstream economics. *Review of Political Economy*, v. 16, n. 4, p. 485-499, 2004.

Commons, John R. Institutional economics: Its place in political economy. v. 1. Routledge, 1990 [1934].

Conlisk, John. Why bounded rationality? *Journal of Economic Literature*, v. 34, n. 2, p. 669-700, 1996.

Costa, Daniel Fonseca; Carvalho, Francisval de Melo; Moreira, Bruno César de Melo. Behavioral economics and behavioral finance: A bibliometric analysis of the scientific fields. *Journal of Economic Surveys*, v. 33, n. 1, p. 3-24, 2019.

Coggan, Anthea et al. Private transaction costs of participation in water quality improvement programs for Australia's Great Barrier Reef: Extent, causes and policy implications. *Australian Journal of Agricultural and Resource Economics*, v. 59, n. 4, p. 499-517, 2015.

Davidson, Paul. Behavioral economists should make a turn and learn from Keynes and Post Keynesian economics. *Journal of Post Keynesian Economics*, v. 33, n. 2, p. 251-254, 2010.

Davis, John B. The turn in recent economics and return of orthodoxy. *Cambridge Journal of Economics*, v. 32, n. 3, p. 349-366, 2008.

Dequech, David. Uncertainty in a strong sense: meaning and sources. *Economic Issue*, 1997.

Dequech, David. Bounded rationality, institutions, and uncertainty. *Journal of Economic Issues*, v. 35, n. 4, p. 911-929, 2001.

Dequech, David. The new institutional economics and the theory of behaviour under uncertainty. *Journal of Economic Behavior & Organization*, v. 59, n. 1, p. 109-131, 2006.

Dequech, David. Neoclassical, mainstream, orthodox, and heterodox economics. *Journal of Post Keynesian Economics*, v. 30, n. 2, p. 279-302, 2007.

Dequech, David. Uncertainty: a typology and refinements of existing concepts. *Journal of Economic Issues*, v. 45, n. 3, p. 621-640, 2011.

Dekker, Erwin; Remic, Blaž. Two types of ecological rationality: or how to best combine psychology and economics. *Journal of Economic Methodology*, v. 26, n. 4, p. 291-306, 2019.

Dosi, Giovanni. Technological paradigms and technological trajectories: a suggested interpretation of the determinants and directions of technical change. *Research Policy*, v. 11, n. 3, p. 147-162, 1982.

Dosi, Giovanni; Egidi, Massimo. Substantive and procedural uncertainty: an exploration of economic behaviours in changing environments. *Journal of Evolutionary Economics*, v. 1, p. 145-168, 1991.

Dosi, Giovanni; Marengo, Luigi; Fagiolo, Giorgio. Learning in evolutionary environments. *LEM Working Paper Series*, n. 2003/20, 2001.

Dow, Sheila. Economic methodology, the philosophy of economics and the economy: another turn? *Journal of Economic Methodology*, v. 28, n. 1, p. 46-53, 2021.

Earl, Peter E. Economics fit for the Queen: a pessimistic assessment of its prospects. *Prometheus*, v. 28, n. 3, p. 209-225, 2010.

Engqvist, Leif; Frommen, Joachim G. The h-index and self-citations. *Trends in ecology & evolution*, v. 23, n. 5, p. 250-252, 2008.

Evans, George W.; Honkapohja, Seppo. *Economics of*. In: International Encyclopedia of the Social & Behavioral Sciences. Pergamon, 2001a. p. 5060-5067.

Evans, George W.; Honkapohja, Seppo. *Learning and expectations in macroeconomics*. Princeton: Princeton University Press, 2001b.

Feigenbaum, Edward A. What hath Simon wrought?. In: *Complex Information Processing*. Psychology Press, 2013 [1989]. p. 185-202.

Fiori, Stefano. Forms of Bounded Rationality: The reception and redefinition of Herbert A. Simon's perspective. *Review of Political Economy*, v. 23, n. 4, p. 587-612, 2011.

Friedman, Milton. *The Methodology of Positive Economics*. Essays in Positive Economics. Chicago: University of Chicago Press, 1953.

Gigerenzer, Gerd. *Adaptive thinking: Rationality in the real world*. Oxford: Oxford University Press, 2002.

Gigerenzer, Gerd. Striking a blow for sanity in theories of rationality. 2004.

Gigerenzer, Gerd. What is Bounded Rationality?. In: *Routledge Handbook of Bounded Rationality*. Routledge, 2020. p. 55-69.

Gigerenzer, Gerd; Berg, Nathan. As-if behavioral economics: Neoclassical economics in disguise?. *History of economic ideas*, v. 18, n. 1, p. 1000-1033, 2010.

Gigerenzer, Gerd; Gaissmaier, Wolfgang. Heuristic decision making. *Annual review of psychology*, v. 62, n. 1, p. 451-482, 2011.

Gigerenzer, Gerd; Goldstein, Daniel G. Reasoning the fast and frugal way: Models of Bounded Rationality. *Psychological review*, v. 103, n. 4, p. 650, 1996.

Gigerenzer, Gerd; Todd, Peter M. Fast and frugal heuristics: The adaptive toolbox. In: *Simple heuristics that make us smart*. Oxford: Oxford University Press, 1999. p. 3-34.

Gigerenzer, Gerd; Selten, Reinhard, eds. *Bounded rationality*. CogNet, 2001.

Giocoli, Nicola. *Modeling rational agents: From interwar economics to early modern game theory*. Cheltenham: Edward Elgar Publishing, 2003.

Gohman, Timothy. Economia comportamental e psicologia na prática. In: Guia de economia comportamental e experimental. São Paulo, 2019. Disponível em: <https://www.economiacomportamental.org/guia/>. Acesso em: 1 jan. 2025.

Haraway, Donna. Simians, cyborgs, and women: The reinvention of nature. Routledge, 2013 [1991].

Heukelom, Floris. Behavioral economics: A history. Cambridge University Press, 2014.

Heilbroner, Robert L. The worldly philosophers: The lives, times and ideas of the great economic thinkers. Simon and Schuster, 2011.

Hodgson, Geoffrey M. Evolutionary and institutional economics as the new mainstream? *Evolutionary and Institutional Economics Review*, v. 4, p. 7-25, 2007.

Hicks, John R. Mr. Keynes and the "classics"; a suggested interpretation. *Econometrica: Journal of the Econometric Society*, v. 5, p. 147-159, 1937.

Holt, C.C.; Modigliani, F.; Simon, H.A. Linear decision rule for production and employment scheduling. *Management Science*, v. 2, p. 1-30, 1955.

Holt, C.C.; Modigliani, F.; Muth, J.F.; Simon, H.A. Planning production, inventory and work force. Englewood Cliffs, NJ: Prentice Hall, 1960.

Hommes, Cars. Behavioral rationality and heterogeneous expectations in complex economic systems. Cambridge University Press, 2013.

Huang, Anna. Similarity measures for text document clustering. In: PROCEEDINGS OF THE SIXTH NEW ZEALAND COMPUTER SCIENCE RESEARCH STUDENT CONFERENCE (NZCSRSC 2008), Christchurch, New Zealand, v. 4, 2008.

Jain, Jinesh, et al. Mapping the field of behavioural biases: A literature review using bibliometric analysis. *Management Review Quarterly*, v. 71, p. 1-33, 2021.

Kahneman, Daniel. Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, v. 93, n. 5, p. 1449-1475, 2003.

Kahneman, Daniel. Thinking, fast and slow. 2017 [2011].

Kahneman, Daniel; Tversky, Amos. On the study of statistical intuitions. *Cognition*, v. 11, n. 2, p. 123-141, 1982.

Kahneman, Daniel; Tversky, Amos. On the reality of cognitive illusions. 1996. p. 582.

Kahneman, Daniel; Tversky, Amos. Prospect theory: An analysis of decision under risk. In: HANDBOOK OF THE FUNDAMENTALS OF FINANCIAL DECISION MAKING: PART I, 2013 [1979], p. 99-127.

Keynes, John Maynard. Alfred Marshall, 1842–1924. *The Economic Journal*, v. 34, n. 135, p. 311-372, 1924.

Keynes, John Maynard. Teoria geral do emprego, do juro e da moeda. São Paulo: Saraiva Educação SA, 2017 [1936].

Knight, Frank H. Risk, uncertainty and profit. Hart, Schaffner and Marx, 1921.

Kirman, Alan. Complex economics: Individual and collective rationality. Routledge, 2010.

Kirzner, Israel M. Rational action and economic theory. *Journal of Political Economy*, v. 70, n. 4, p. 380-385, 1962.

Křečková, J.; Brožová, H. Agricultural insurance and bounded rationality. *Agris On-line Papers in Economics and Informatics*, v. 9, n. 1, p. 91-97, 2017.

Lakatos, Imre. Criticism and the methodology of scientific research programmes. *Proceedings of the Aristotelian Society*, v. 69, 1968.

Langlois, Richard, ed. *Economics as a process: Essays in the new institutional economics*. Cambridge University Press, 1986.

Lavoie, Marc. Theory of choice. In: *POST-KEYNESIAN ECONOMICS*, 2022, p. 75-127. Edward Elgar Publishing.

Lee, Frederic S. A case for ranking heterodox journals and departments. *On the Horizon*, v. 16, n. 4, p. 241-251, 2008.

Lee, Frederic S. *A history of heterodox economics: Challenging the mainstream in the twentieth century*. Routledge, 2009.

Lee, F. S.; Cronin, B. C.; McConnell, S.; Dean, E. Research quality rankings of heterodox economic journals in a contested discipline. *American Journal of Economics and Sociology*, v. 69, n. 5, p. 1409-1452, 2010.

Loasby, B. Herbert Simon's human rationality. *Research in the History of Economic Thought and Methodology*, v. 6, p. 1-17, 1989.

Lopes, Lola L. Three misleading assumptions in the customary rhetoric of the bias literature. *Theory & Psychology*, v. 2, n. 2, p. 231-236, 1992.

Maltaca, Jose; Almeida, Felipe. *Journal of Economic Issues e Journal of Institutional Economics: uma análise bibliométrica sobre o tipo de institucionalismo promovido por suas principais revistas*. In: *ANAIS DO 50º ENCONTRO NACIONAL DE ECONOMIA*, Fortaleza, 2022.

March, James G.; Simon, Herbert A. *Organizations*. John Wiley & Sons, 1993 [1958].

Marx, Karl. *O capital - Livro 1: Crítica da economia política*. Livro 1: O processo de produção do capital. São Paulo: Boitempo Editorial, 2015 [1867].

Mäki, Uskali. The field: tasks, pasts, futures. *Journal of Economic Methodology*, v. 28, n. 1, p. 3-13, 2021.

Manzini, Paola; Mariotti, Marco. Sequentially rationalizable choice. *American Economic Review*, v. 97, n. 5, p. 1824-1839, 2007.

McCloskey, Deirdre Nansen. *The rhetoric of economics*. University of Wisconsin Press, 1985.

Melo, Tatiana Massaroli; Possas, Mario Luiz; Dweck, Esther. Um modelo setorial baseado na abordagem kaleckiana da distribuição setorial funcional da renda e na teoria schumpeteriana da concorrência. *Economia e Sociedade*, v. 25, n. 01, p. 109-145, 2016.

Mirowski, Philip. *Machine dreams: Economics becomes a cyborg science*. Cambridge University Press, 2002.

Mitchell, Wesley C. The rationality of economic activity. *Journal of Political Economy*, v. 18, n. 3, p. 197-216, 1910.

Miller, George A. The cognitive revolution: A historical perspective. *Trends in Cognitive Sciences*, v. 7, n. 3, p. 141-144, 2003.

Mußhoff, Oliver, et al. Bounded rationality and the adoption of weather index insurance: Evidence from an extra-laboratory experiment with farmers in Germany. *Agricultural Finance Review*, v. 78, n. 1, p. 116-134, 2018.

Modigliani, Franco. *Adventures of an economist*. 2001.

Nagatsu, Michiru. Behavioral economics, history of. In: *INTERNATIONAL ENCYCLOPEDIA OF THE SOCIAL & BEHAVIORAL SCIENCES*, v. 2, n. 2, 2015.

Nelson, Richard R.; Winter, Sidney G. The Schumpeterian tradeoff revisited. *The American Economic Review*, v. 72, n. 1, p. 114-132, 1982.

Nelson, Richard R.; Winter, Sidney G. *An evolutionary theory of economic change*. Harvard University Press, 1985.

Newell, Allen; Simon, Herbert Alexander. *Human problem solving*. v. 104, n. 9. Englewood Cliffs, NJ: Prentice-Hall, 1972.

Pessali, H. F. Teoria do custos de transação: Uma avaliação à luz de diferentes correntes do pensamento econômico. 1998. 155 f. Dissertação (Mestrado) - Curso de Desenvolvimento Econômico, Universidade Federal do Paraná (UFPR), Curitiba, 1998.

Petracca, Enrico. A cognition paradigm clash: Simon, situated cognition and the interpretation of bounded rationality. *Journal of Economic Methodology*, v. 24, n. 1, p. 20-40, 2017.

R Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Viena, Áustria, 2022. Disponível em: <https://www.R-project.org/>. Acesso em: 1 jan. 2025.

RStudio Team. *RStudio: Integrated Development for R*. RStudio, PBC, Boston, MA, 2020. Disponível em: <http://www.rstudio.com/>. Acesso em: 1 jan. 2025.

Raposo, Elsa Simões. George Katona (1901-1981): Precursor da psicologia económica. *Análise Psicológica*, 1983, p. 299-307.

Robbins, Lionel. An essay on the nature and significance of economic science. Ludwig von Mises Institute, 2007 [1932].

Rubinstein, Ariel. Modeling bounded rationality. MIT Press, 1998.

Rutherford, Malcolm. The institutionalist movement in American economics, 1918–1947: Science and social control. Cambridge University Press, 2011.

Sargent, Thomas J. Bounded rationality in macroeconomics: The Arne Ryde Memorial Lectures. Oxford University Press, 1993.

Savage, Leonard J. The foundations of statistics. Courier Corporation, 1972 [1952].

Schultz, Duane P.; Schultz, Sydney Ellen. A history of modern psychology. Wadsworth, Cengage Learning, 2011.

Schumpeter, Joseph A. The theory of economic development. Routledge, 2021 [1911].

Sent, Esther-Mirjam. Sargent versus Simon: Bounded rationality unbound. *Cambridge Journal of Economics*, v. 21, n. 3, p. 323-338, 1997.

Sent, Esther-Mirjam. Behavioral economics: How psychology made its (limited) way back into economics. *History of Political Economy*, v. 36, n. 4, p. 735-760, 2004.

Sent, Esther-Mirjam. Simplifying Herbert Simon. *History of Political Economy*, v. 37, n. 2, p. 227-232, 2005.

Sposito, Theodoro Cesar de Oliveira. Pluralism in Brazilian economics: History, institutions, and implications. Campinas State University, 2024.

Simon, Herbert A. A behavioral model of rational choice. *The Quarterly Journal of Economics*, 1955, p. 99-118.

Simon, Herbert A. Rational choice and the structure of the environment. *Psychological Review*, v. 63, n. 2, p. 129, 1956.

Simon, Herbert A. Models of man; social and rational. 1957a.

Simon, Herbert A. A behavioral model of rational choice. In: *MODELS OF MAN, SOCIAL AND RATIONAL: Mathematical essays on rational human behavior in a social setting*, v. 6, n. 1, p. 241-260, 1957b.

Simon, Herbert A. Theories of bounded rationality. 1964.

Simon, Herbert A. The sciences of the artificial. 1969.

Simon, Herbert A. From substantive to procedural rationality. 1976.

Simon, Herbert A. Administrative behavior. New York, NY: MacMillan, republished in 1957, 1976, 1996. 1976 [1947].

Simon, Herbert A. Rationality as process and as product of thought. The American Economic Review, v. 68, n. 2, p. 1-16, 1978.

Simon, Herbert A. Rational decision making in business organizations. The American Economic Review, v. 69, n. 4, p. 493-513, 1979.

Simon, Herbert A. Why should machines learn? In: MACHINE LEARNING, Morgan Kaufmann, 1983, p. 25-37.

Simon, Herbert A. Invariants of human behavior. Annual Review of Psychology, v. 41, n. 1, p. 1-20, 1990a.

Simon, Herbert. Reason in human affairs. Stanford University Press, 1990b.

Simon, Herbert A. Bounded rationality. In: Eatwell, J.; Milgate, M.; Newman, P. (eds). Utility and Probability. The New Palgrave. Palgrave Macmillan, London, 1990c.

Simon, Herbert A. Organizations and markets. Journal of Economic Perspectives, v. 5, n. 2, p. 25-44, 1991.

Simon, Herbert A. Decision making: Rational, nonrational, and irrational. Educational Administration Quarterly, v. 29, n. 3, p. 392-411, 1993.

Simon, Herbert A. Artificial intelligence: An empirical science. Artificial Intelligence, v. 77, n. 1, p. 95-127, 1995.

Simon, Herbert A. Models of my life. MIT Press, 1996 [1991].

Simon, Herbert Alexander. Models of bounded rationality: Empirically grounded economic reason. v. 3. MIT Press, 1997 [1982].

Simon, Herbert A. Bounded rationality in social science: Today and tomorrow. Mind & Society, v. 1, p. 25-39, 2000.

Simon, Herbert A. Models of discovery: And other topics in the methods of science. v. 54. Springer Science & Business Media, 2012 [1977].

Smith, Vernon L. Constructivist and ecological rationality in economics. American Economic Review, v. 93, n. 3, p. 465-508, 2003.

Steingraber, Ronivaldo; Garcia Fernandez, Ramon. A racionalidade limitada de Herbert Simon na microeconomia. *Revista da Sociedade Brasileira de Economia Política*, 2013.

Stigler, George J. The economics of information. *Journal of Political Economy*, v. 69, n. 3, p. 213-225, 1961.

Stigler, G. The theory of price. New York: Macmillan, 1966.

Stiglitz, Joseph E. Freefall: America, free markets, and the sinking of the world economy. New York: Norton, 2010.

Thaler, Richard. Toward a positive theory of consumer choice. *Journal of Economic Behavior & Organization*, v. 1, n. 1, p. 39-60, 1980.

Thaler, Richard H. Irving Fisher: Modern behavioral economist. *The American Economic Review*, v. 87, n. 2, p. 439-441, 1997.

Thaler, Richard H. Behavioral economics: Past, present, and future. *American Economic Review*, v. 106, n. 7, p. 1577-1600, 2016.

Thaler, Richard H.; Sunstein, Cass R. *Nudge: The final edition*. Yale University Press, 2021 [2008].

Tobón, Santiago. The Rational Expectations Hypothesis: An assessment on its real world application. *Ecos de Economía*, v. 18, n. 39, p. 37-47, 2014.

Tomer, John F. What is behavioral economics?. *The Journal of Socio-Economics*, v. 36, n. 3, p. 463-479, 2007.

Truc, Alexandre. Interdisciplinary influences in behavioral economics: A bibliometric analysis of cross-disciplinary citations. *Journal of Economic Methodology*, v. 29, n. 3, p. 217-251, 2022.

Tversky, Amos; Shafir, Eldar. The disjunction effect in choice under uncertainty. *Psychological Science*, v. 3, n. 5, p. 305-310, 1992.

Tversky, Amos; Kahneman, Daniel. Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, v. 185, n. 4157, p. 1124-1131, 1974.

Tversky, Amos; Kahneman, Daniel; Slovic, Paul. *Judgment under uncertainty: Heuristics and biases*. 1982.

Veblen, Thorstein. *The theory of the leisure class*. Routledge, 2017 [1899].

Veblen, Thorstein. Why is economics not an evolutionary science?. The Quarterly Journal of Economics, v. 12, n. 4, p. 373-397, 1898.

Viskovatoff, Alex. Rationality as optimal choice versus rationality as valid inference. Journal of Economic Methodology, v. 8, n. 2, p. 313-337, 2001.

Von Neumann, John; Morgenstern, Oskar. Theory of games and economic behavior, 2nd rev. 1944.

Williamson, Oliver E. Transaction cost economics. Handbook of Industrial Organization, v. 1, p. 135-182, 1989.

Woodford, Michael. Optimal interest-rate smoothing. The Review of Economic Studies, v. 70, n. 4, p. 861-886, 2003.

Yu, Jingyuan; Muñoz-Justicia, Juan. A bibliometric overview of twitter-related studies indexed in Web of Science. Future Internet, v. 12, n. 5, p. 91, 2020.

Zhao, Meng et al. Attention-based constraint to MNC coevolution in China's changing stakeholder environment. Journal of Business Ethics, v. 186, n. 4, p. 797-814, 2023.

APPENDIX I – Details of the Bibliographical Model

This Section presents in greater mathematical detail the process of clustering analyzed in Chapter III.

Document Representation

According to Huang (2008), a set of documents can be grouped in a set $D = \{d1, \dots, dn\}$ in which the set of terms included in these documents are represented as $T = \{t1, \dots, tm\}$. Documents from set D are represented as an m-dimensional. This representation of vectors permits the comparison of documents. The element $d1$, from set D , for instance, can be compared with element $d2$ from the same set by analyzing the vectors of the two elements¹²⁸. The amount of times a term is used defines its weight in the document (Sposito, 2024).

Two further precautions are taken to identify the vectors that describe elements from set D (Huang, 2008): 1) the suppression of stop-words; 2) the grouping of words that share a suffix (for instance, rational, rationality, rationally, rationale, etc., share the suffix “rational” and thus are grouped together); 3) the suppression of infrequent words (Huang, 2008 argues that these words are probably not representative of the contents of the document and add noise to the clustering). After it, the similar documents are closely linked by their vector coordinates and are clustered.

Clusters are positioned according to “centroid”, defined as:

$$\vec{tc} = 1/|C| \sum \vec{td}$$

Source: Huang, 2008.

Where $\vec{tc} \in C$, C is the set of documents, considering $C \in D$ and $\vec{tc}$ is the m-dimensional vector of documents, considering $td \in T$. The default distance is measured by the K-mean algorithm, which measures the Euclidean distance between documents (Aria and Cuccurullo, 2017).

¹²⁸ Another popular coupling algorithm is the Cosine Similarity. As *bibliometrix* uses the K-mean algorithm by default, this measurement was not used, for K-mean is based on the Euclidean distance between documents.

This distance is defined by the formula:

$$De(\vec{t}_a, \vec{t}_b) = \left(\sum (w_{t,a} - w_{t,b})^2 \right)^{(1/2)}$$

Source: Huang, 2008.

The K-means algorithm aims to minimize the least squares error and generate locally optimal solutions for the data set¹²⁹ (Huang, 2008), similar to the Ordinary Minimum Square method, popular in econometrics. A consequence of the K-means algorithm is the minimization of the distance inside clusters.

Co-citation analysis

Co-citation is a type of bibliographic coupling, defined by Aria and Cuccurullo (2017) as when two articles are both cited in a third article. The *bilbiometrix* tool functions by connecting a document's information with itself, forming connections. Aria and Cuccurullo (2017) define as matrix *A* as the product of the matrices *Document* and *Attribute*. Therefore:

$$A = Document \times Attribute$$

Source: Aria and Cuccurullo, 2017

Articles “are said to be bibliographically coupled if at least one cited source appears in the bibliographies or reference lists of both articles” (Aria and Cuccurullo, 2017, p. 968). A matrix, B_{coup} , is defined as the product of matrix *A* and its transposed matrix. The co-citation analysis, the central key of the bibliometric method of this dissertation, is a subcase of this coupling matrix, (B_{cocit}), defined as:

¹²⁹ This starts with the random selection of a centroid of the document, defining an initial seed. The locally optimal solution occurs taking account of both the data set and the initial seed. See Huang (2008) for further detail.

$$B_{cocit} = A \times A'$$

Source: Aria and Cuccurullo, 2017

This is a symmetric matrix that “indicates how many co-citations exist between documents” (Aria and Cuccurullo, 2017, p. 969). Both the co-citation network and thematic maps are visualizations of this matrix B_{cocit} .

For a demonstration of this bibliographical statistical model, see Sposito (2024, p. 75-77).

References

Aria, M. & Cuccurullo, C. (2017) bibliometrix: An R-tool for comprehensive science mapping analysis, *Journal of Informetrics*, 11(4), pp 959-975, Elsevier.

Huang, Anna. "Similarity measures for text document clustering." *Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008)*, Christchurch, New Zealand. Vol. 4. 2008.

Sposito, Theodoro Cesar de Oliveira. *PLURALISM IN BRAZILIAN ECONOMICS: HISTORY, INSTITUTIONS, AND IMPLICATIONS*. Campinas State University. (2024).

APPENDIX II – R script used to process and present the collected data.

This appendix presents the code that was used in the bibliographical analysis. This code was processed through the software R, with the assistance of the interface RStudio (the methodology is specified in Topic III.1). Note that the results were obtained through the interface *biblioshiny*, included in the *bibliometrix* package.

For more information, go to: <https://www.bibliometrix.org/home>

Code for the Bibliographical Analysis.

```
rm(list=ls())

#Configurar diretório.
setwd("C:\\Users\\Bruno\\Desktop\\Dissertação\\Dados Racionalidade Limitada\\Dados I - Bounded Rationality")
getwd()
dir()

{# install.packages("knitcitations")
# install.packages("bibliometrix")}

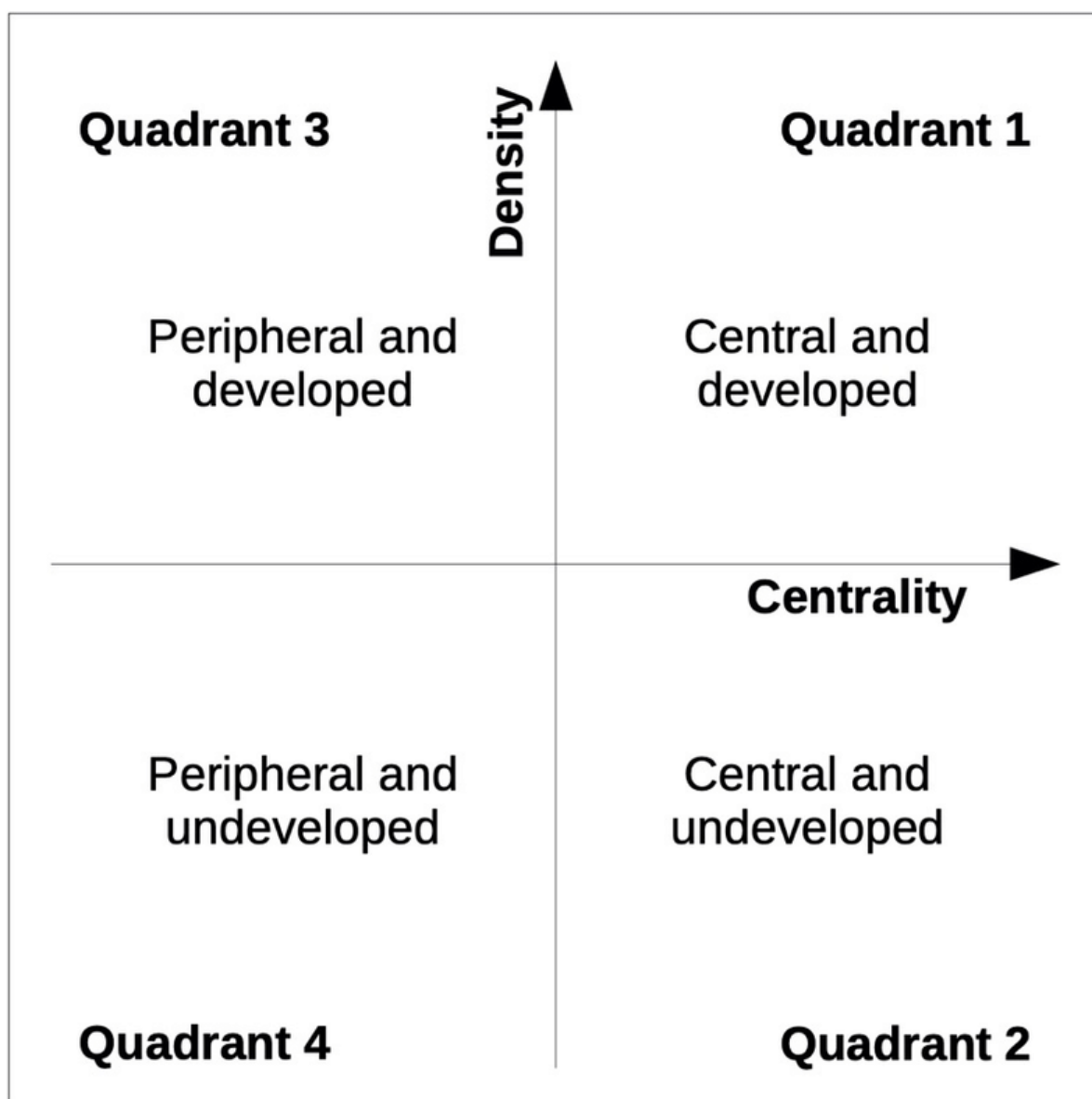
library(bibliometrix)
#####

(I) Bounded Rationality
M = convert2df("scopus_I_1_ESSEAQUI.csv", dbsource = "scopus", format = "csv", remove.duplicates = TRUE)

write.bibtex(M, file = "knitcitations.bib", append = FALSE,
             verbose = TRUE, create_key = TRUE)
write.csv(M, file = "dados", row.names = 1)

biblioshiny()
```

APPENDIX III – Interpretation of the Thematic Map



Source: Yu and Muñoz-Justicia (2020).

References

Yu, Jingyuan, and Juan Muñoz-Justicia. "A bibliometric overview of twitter-related studies indexed in web of science." *Future Internet* 12.5 (2020): 91.

APPENDIX IV – Complete list of Co-citations

Complete List of Co-citation of References in Sample 1

Co-citation of Authors of Sample 1

Node	Cluster	Betweenness	Closeness
Simon Ha	1 (red)	68,79	0,01
Kahneman D	1 (red)	41,32	0,01
Tversky A	1 (red)	17,13	0,01
Gigerenzer G	1 (red)	3,24	0,01
Selten R	1 (red)	3,51	0,01
Camerer Cf	1 (red)	8,57	0,01
Rubinstein A	1 (red)	12,81	0,01
Crawford Vp	1 (red)	1,29	0,01
Fehr E	1 (red)	2,38	0,01
Fudenberg D	1 (red)	2,4	0,01
Thaler Rh	1 (red)	2,68	0,01
Spiegler R	1 (red)	1,03	0,01
Camerer C	1 (red)	5,44	0,01
Rabin M	1 (red)	2,11	0,01
Simon H	1 (red)	8,45	0,01
Williamson Oe	1 (red)	1,11	0,01
March Jg	1 (red)	0,15	0,01
Guth W	1 (red)	0,44	0,01
Stahl Do	1 (red)	0,56	0,01
Conlisk J	1 (red)	3,71	0,01
Fischbacher U	1 (red)	0,6	0,01
Nagel R	1 (red)	4,07	0,01
Smith Vl	1 (red)	0,79	0,01
Jehiel P	1 (red)	0,1	0,01
Friedman M	1 (red)	0,92	0,01
Mckelvey Rd	1 (red)	0,17	0,01
Verbeke A	1 (red)	0	0
Evans Gw	2 (blue)	24,94	0,01
Hommes C	2 (blue)	26,03	0,01
Branch Wa	2 (blue)	0,88	0,01
Chiarella C	2 (blue)	1,08	0,01
Woodford M	2 (blue)	44,82	0,01
Hommes Ch	2 (blue)	1,06	0,01

Brock Wa	2 (blue)	6,45	0,01
De Grauwe P	2 (blue)	5,82	0,01
Anufriev M	2 (blue)	3,48	0,01
Lux T	2 (blue)	1,85	0,01
Day Rh	2 (blue)	0,83	0,01
Gali J	2 (blue)	9,5	0,01
Arifovic J	2 (blue)	11,7	0,01
Brock W	2 (blue)	1,84	0,01
Gabaix X	3 (green)	252,42	0,02
Dosi G	3 (green)	85,03	0,01
Barberis N	3 (green)	80,48	0,01
Shiller Rj	3 (green)	47,96	0,01
Sims Ca	3 (green)	29,28	0,01
Angeletos Gm	3 (green)	26,39	0,01
Manzini P	4 (purple)	16,39	0,01
Masatlioglu Y	4 (purple)	5,19	0,01
Eliaz K	4 (purple)	11,84	0,01

Source: own elaboration, generated in *bibliometrix*.

Co-citation of References of Sample 1

Node	Cluster	Betweenness	Closeness
Rubinstein A. 1998	1.1 (purple)	170,65	0,01
Conlisk J. 1996-1	1.1 (purple)	46,31	0,01
Simon H.A. 1955-1	1.1 (purple)	53,49	0,01
Spiegler R. 2011	1.1 (purple)	67,72	0,01
Eyster E. 2005	1.1 (purple)	0	0,01
Nelson R.R. 1982	1.1 (purple)	0,04	0
Cyert R.M. 1963	1.1 (purple)	118,07	0,01
Dellavigna S. 2009	1.1 (purple)	45	0,01
Kahneman D. 2003-1	1.1 (purple)	0	0
Luce R.D. 1959	1.1 (purple)	0	0
Newell A. 1972	1.1 (purple)	46,87	0,01
Fischbacher U. 2007-2	1.1 (purple)	0	0
Sims C.A. 2003	1.1 (purple)	0	0
Crawford V.P. 2013	1.1 (purple)	0	0
Nagel R. 1995	1.1 (purple)	0,01	0
Tversky A. 1992-2	1.1 (purple)	0	0,01
Fehr E. 1999	1.1 (purple)	17,21	0,01
Kahneman D. 2011	1.2 (green)	176,13	0,01
Salant Y.	1.2 (green)	2	0,01
Gigerenzer G. 1999	1.2 (green)	65,46	0,01
Tversky A. 1974-2	1.2 (green)	38,5	0,01
Von Neumann J. 1944	1.2 (green)	74,11	0,01
Gigerenzer G. 2001	1.2 (green)	33,49	0,01
Savage L.J. 1954	1.2 (green)	79,11	0,01
Manzini P. 2007	1.2 (green)	0	0
Simon H.A. 1955-3	1.2 (green)	7,65	0
Tversky A. 1992-1	1.2 (green)	0	0
Kahneman D. 1982	1.2 (green)	90,28	0,01
Kahneman D. 1979-2	1.2 (green)	0	0
Knight F.H. 1921	1.2 (green)	1,1	0,01
Simon H.A. 1955-4	1.2 (green)	0	0,01
Conlisk J. 1996-2	1.2 (green)	0	0,01
Kahneman D. 2003-2	1.2 (green)	0	0,01
Kahneman D. 1979-1	1.3 (blue)	351,79	0,01
March J.G. 1958	1.3 (blue)	122,4	0,01
Tversky A. 1974-1	1.3 (blue)	33,77	0,01
Simon H.A. 1955-2	1.3 (blue)	157,51	0,01
Sargent T.J. 1993	1.3 (blue)	245,73	0,01
Simon H.A. 1956	1.3 (blue)	2,99	0,01

Camerer C.F. 2003	1.3 (blue)	7,43	0,01
Evans G.W. 2001	1.4 (yellow)	210,56	0,01
Woodford M. 2003	1.4 (yellow)	0	0,01
Keynes J.M. 1936	1.4 (yellow)	0	0
Brock W.A. 1997	1.4 (yellow)	0	0,01
Hommes C.H. 2013	1.4 (yellow)	0	0,01
Hommes C. 2013	1.5 (red)	130,57	0,01
Brock W. 1997	1.5 (red)	0	0

Source: own elaboration, generated in *bibliometrix*.

Complete List of Co-citation of References in Sample 2

Co-citation of Authors of Sample 2

Node	Cluster	Betweenness	Closeness
Simon Ha	1 (red)	241,14	0,01
Kahneman D	1 (red)	114,75	0,01
Gigerenzer G	1 (red)	75,89	0,01
Tversky A	1 (red)	8,5	0,01
Sent Em	1 (red)	6,15	0,01
Newell A	1 (red)	8,56	0,01
Altman M	1 (red)	0	0
Rubinstein A	1 (red)	1,5	0,01
Evans Jsbt	1 (red)	0	0
March Jg	1 (red)	2,7	0,01
Selten R	1 (red)	0	0
Conlisk J	1 (red)	0	0
Mirowski P	1 (red)	0,48	0,01
Arrow Kj	1 (red)	0,23	0,01
Hands Dw	1 (red)	15,12	0,01
Callebaut W	1 (red)	0	0,01
Crowtherheyck H	1 (red)	0	0,01
Fiori S	1 (red)	0	0,01
Friedman M	2 (blue)	57,57	0,01
Gul F	2 (blue)	50,66	0,01
Alchian Aa	2 (blue)	27,32	0,01
Allais M	2 (blue)	44,3	0,01
Machlup F	2 (blue)	0,48	0,01
Putnam H	2 (blue)	0	0,01
Cyert Rm	2 (blue)	2,14	0,01
Giocoli N	2 (blue)	40,21	0,01
Caplin A	3 (green)	0	0,01
Masatlioglu Y	3 (green)	0	0,01
Dosi G	4 (purple)	1,11	0,01
Dequech D	4 (purple)	1,11	0,01
Barros G	4 (purple)	3,11	0,01
Nickles T	5 (yellow)	0	0,01
Simon H	6 (brown)	0	0,01
Augier M	7 (pink)	0	0,01
Lawson T	8 (grey)	0	0,01

Source: own elaboration, generated in *bibliometrix*.

Co-citation of References of Sample 2

Node	Cluster	Betweenness	Closeness
Newell A. 1972	2.1 (red)	60	0,01
Gigerenzer G. 1999-1	2.1 (red)	33	0,01
Heukelom F. 2014	2.1 (red)	12	0,01
Sargent T.J. 1993	2.1 (red)	85	0,01
Conlisk J. 1996	2.1 (red)	0	0,01
Crowther-Heyck H. 2005	2.1 (red)	0	0,01
Petracca E. 2017	2.1 (red)	0	0,01
Rubinstein A. 1998	2.1 (red)	0	0,01
Simon H.A. 1996-2	2.1 (red)	0	0,01
Simon H.A. 1955-1	2.2 (blue)	8	0,01
Simon H.A. 1956	2.2 (blue)	8	0,01
Gigerenzer G. 2004	2.2 (blue)	0	0,01
Simon H.A. 2000	2.2 (blue)	0	0,01
Stigler G.J. 1961	2.3 (grey)	0	0,01
Kahneman D. 1979	2.3 (grey)	0	0,01
Simon H.A. 1991	2.4 (pink)	0	0,01
Simon H.A. 1976	2.4 (pink)	0	0,01
Simon H.A. 1997-1	2.5 (brown)	0	0,01
Simon H.A. 1996-1	2.5 (brown)	0	0,01
Dequech D. 2001	2.6 (purple)	0	0,2
Simon H.A. 1983	2.6 (purple)	0	0,2
Giocoli N. 2003	2.7 (green)	0	0,2
Mirowski P. 2002	2.7 (green)	0	0,2

Source: own elaboration, generated in *bibliometrix*.