Does Language Shape Our Reasoning?
A Critical Analysis of Rothschild's Theory on Al and Domain-General Reasoning

Regina Kukuh P Ati / 3212507

ABSTRACT

Rothschild proposes the view that language plays a role in shaping the way we think, particularly in the context of artificial intelligence (AI) systems and domain-general reasoning. This study examines the relationship between language and human reasoning, focusing on the theory put forward by Rothschild regarding the role of language in shaping how we think and argue. Rothschild's theory on AI and domain-general reasoning will be critically analyzed with the aim of understanding the extent to which language influences our ability to think rationally and develop solutions to complex problems. Through an analytical approach, this research presents a more comprehensive view of how language can shape our thinking framework and how this can be translated into the development of more advanced AI. This study contributes to a better understanding of the interaction between language and reasoning, as well as the challenges faced in replicating human thinking processes through machines. It is hoped that this research will open new discourse in the fields of AI and cognitive psychology, emphasizing the importance of language as a key factor in rational and effective human reasoning.

Keywords: *Language, Rothschild's Theory, Artificial Intelligence, Domain-General Reasoning*

INTRODUCTION

Reasoning is one of the fundamental elements of human intelligence, enabling individuals to understand, analyze, and solve problems logically. This ability forms the basis for various intellectual activities, including decision-making, problem-solving, and the development of new ideas (Minda, 2020). Reasoning plays a role in everyday life and is also at the core of the development of science, technology, and human culture as a whole (Viale, 2001).

In reasoning, language is the medium for organizing, articulating, and communicating human thoughts. Language is not only a tool for communication but also functions as a structure that facilitates cognitive processes such as information organization and drawing conclusions (Carruthers, 2002). Research shows that the development of reasoning ability often aligns with language skills, indicating a close relationship between the two (Leroy et al., 2012).

Moreover, language enables humans to conceptualize abstract ideas and communicate them to others, which in turn fosters collaboration and innovation. This suggests that language is not only an expression tool but also an important instrument in building and expanding human intellectual horizons (Pinker, 2010). By using language, individuals can formulate arguments, evaluate evidence, and generate creative solutions, all of which are vital components in the reasoning process (Kaijanaho, 2015).

Understanding the role of language in reasoning is increasingly important in the modern era, especially with the emergence of artificial intelligence (AI) designed to mimic human cognitive abilities. AI systems, such as large language models, highlight the importance of language as the foundation for building artificial intelligence capable of deeply understanding and processing information (Kolides et al., 2023). Therefore, the study of the relationship between reasoning and language is not only relevant in cognitive science but also has broad practical implications in the development of modern technology.

Artificial intelligence (AI) has developed rapidly in recent decades and is now capable of handling various cross-domain tasks with ever-increasing sophistication. One significant advancement is AI's ability to perform domain-general reasoning, which is the ability to understand, analyze, and solve problems across different contexts without being limited to a specific domain (Russell & Norvig, 2020).

Generative AI models such as Large Language Models (LLMs), including GPT, have become the center of attention due to their ability to process natural language and generate human-like responses (Hadi et al., 2024). In this context, LLMs demonstrate how language use not only serves as a communication tool but also plays a crucial role in supporting cross-domain reasoning.

Daniel Rothschild, in his academic work titled *Language and Thought in Large Language Models*, highlights the critical role of language in shaping and expanding the domain-general reasoning abilities of AI. Rothschild argues that language is not just a medium for transferring information but also a cognitive framework that enables AI to understand the relationships between concepts across various domains (Rothschild, 2024). This perspective suggests that the use of language helps AI models identify patterns, make inferences, and even formulate new hypotheses relevant to different contexts.

By leveraging the structure of natural language, LLMs like GPT can represent complex knowledge and suggest creative solutions, even though they are limited to the data they have learned. Rothschild notes that language acts as a "shortcut" for AI to overcome the domain-specific limitations that typically constrain traditional AI systems (Rothschild, 2024). This indicates that human language, rich in meaning and context, plays a crucial role in supporting high-level reasoning, both for humans and artificial intelligence systems.

Advancements in LLM technology have not only technical but also epistemological implications, where the relationship between language and thought becomes a key topic in the philosophy of artificial intelligence. Rothschild emphasizes that this research is not only relevant for technological development but also provides new insights into how language shapes the way AI, and even humans, process information (Rothschild, 2024). This perspective positions LLMs as a meeting point between linguistics, cognition, and technology, opening new opportunities to better understand the relationship between language and thought more broadly (Bender et al., 2021).

The Sapir-Whorf Hypothesis, also known as the linguistic relativity hypothesis, is a theoretical framework that highlights the influence of language on human thought patterns and perception. This theory proposes that the structure of the language used by individuals can shape the way they understand the world, influence cognitive processes, and guide decision-making patterns (Perlovsky, 2009).

In the context of this research, the concept becomes relevant when connected to how artificial intelligence (AI) systems, such as large language models (LLMs), "think" and "understand" through the language used to train these models. By comparing human and AI cognition patterns, this hypothesis provides a theoretical foundation for exploring the extent to which language affects reasoning abilities in both entities.

One of the key propositions of this hypothesis is that language differences will result in differences in how individuals perceive reality. In other words, language is not only a tool for communication but also a framework that shapes one’s worldview (Ali, 2023). This concept becomes particularly interesting in the AI era, where language-based models like GPT-4 are designed to understand and generate contextually appropriate text. Does AI, trained using human language, also get "shaped" by the structure of that language in the same way humans do? This research seeks to answer that question using a comparative approach, contrasting the effects of language on humans with its influence on AI systems.

Although the development of artificial intelligence (AI) has shown tremendous potential in mimicking human thinking patterns, there is significant disagreement among researchers about whether the way AI performs reasoning truly aligns with human reasoning processes. Some researchers argue that large language models (LLMs) like GPT-4 can represent human thought due to their ability to understand and generate natural language (Kalyan, 2023).

However, this view is contested by others who claim that AI's success is more based on statistical capabilities than actual cognitive reasoning (Felin & Holweg, 2024). Thus, this debate highlights a research gap in understanding the extent to which AI reasoning truly resembles human reasoning and whether language is a key element in this process.

The primary focus of this gap is the lack of evidence supporting a direct relationship between language and domain-general reasoning abilities in AI. In Rothschild's theory, language is seen as a crucial foundation in supporting universal reasoning that is not tied to a specific context (Rothschild, 2024). However, this theory has not been adequately tested in the context of modern AI models. Much of the previous research has only emphasized the role of language in text-based learning without deeply exploring how language influences the flexibility of cross-domain reasoning in AI systems.

The main objective of this paper is to critically analyze the theory proposed by Rothschild and its implications for understanding reasoning processes in artificial intelligence (AI). In this context, the research aims to explore how Rothschild's theory can provide insights into the connection between language and thought, as well as its relevance to the development of AI technology. The central thesis of this paper is that while Rothschild's theory offers valuable insights into the relationship between language and reasoning in AI, a deeper evaluation of the role of language in reasoning is necessary to understand the implications of the theory both in the context of AI and in relation to human cognition.

METHOD

This study use a literature review method to explore whether language shapes our reasoning, with a particular focus on a critical analysis of Rothschild's theory on artificial intelligence (AI) and cross-domain reasoning (Rothschild, 2024). The literature review approach was chosen as it enables the author to gather information from various credible sources, including scientific journals, books, articles, and online publications relevant to the research topic (Snyder, 2019). This approach aims to examine in-depth the relationship between linguistic structures and human reasoning abilities, as well as to understand the extent to which Rothschild's theory can be applied in the context of modern AI.

The data collection process began with identifying relevant literature through online searches using keywords such as "language and reasoning", "Rothschild's theory", "domain-general artificial intelligence", and "Sapir-Whorf hypothesis". Once the relevant literature was gathered, content analysis was conducted to identify connections between specific linguistic characteristics and observed reasoning patterns in both humans and AI. This analysis also explored how machine learning algorithms can be designed to replicate or extend human cross-domain reasoning processes.

This research employs data analysis methods adapted from Miles et al. (2014), which consist of four main stages: data collection, data reduction, data presentation, and conclusion drawing. In the first stage, relevant literature was collected through academic databases such as ScienceDirect as well as additional searches on Google Scholar for open-access sources. The search results focused on publications that specifically discuss the relationship between language and reasoning, Rothschild's theory, and domain-general approaches in AI.

The second stage involved data reduction, where the obtained information was filtered to ensure that only the most relevant and significant data were analyzed. This process included a critical reading of the collected literature to identify patterns, key concepts, or significant findings pertinent to the research questions. For example, particular attention was given to studies discussing the relationship between language's influence on cross-cultural cognition and its implications for designing modern AI systems.

The third stage was data presentation, where the analyzed information was systematically organized based on the predetermined research framework. The data were presented in the form of analytical narratives to facilitate the reader's understanding of the relationships between the research findings (Creswell, 2014). This presentation included comparisons between Rothschild's theory and other cognitive models, as well as their relevance to the implementation of domain-general AI.

The final stage was conclusion drawing, where the results of the analysis were used to address the main research question. Based on the presented data, this study concludes the extent to which Rothschild's theory explains the relationship between language and reasoning, while also evaluating its potential application in future AI technology development. Furthermore, this study identifies the limitations of the theory and proposes directions for future research to delve deeper into the connections between language, reasoning, and technology.

ANALYSIS

Rothschild argues that artificial intelligence (AI) leverages the structure of language to support domain-general reasoning, the ability to think in ways not confined to a specific field (Rothschild, 2024). This is based on the premise that language, as a symbolic system, provides a framework for AI to construct flexible conceptual models. However, this analysis has faced criticism, particularly because modern AI models like GPT tend to rely on statistical correlations rather than deep semantic understanding. This suggests that while language may offer an initial structure, AI does not fully grasp the concepts behind the words it uses.

Human reasoning is often influenced by language, but this does not directly translate into the AI context. Humans tend to use linguistic structures to build narratives or logical patterns when solving problems. In contrast, AI generates responses based on the probabilistic sequence of words without understanding their true contextual meaning. Thus, while AI may appear capable of "thinking", the linguistic structures it employs are more reflective of algorithmic mechanisms than domain-general reasoning capacity.

Furthermore, Rothschild's approach also faces challenges from cross-linguistic cognitive perspectives. Language is believed to influence human thought patterns, as described by the Sapir-Whorf Hypothesis, wherein linguistic differences can shape experiences and understanding of the world. However, AI trained across multiple languages does not exhibit evidence that linguistic differences influence its reasoning models. Instead, AI tends to normalize these differences into uniform patterns, highlighting its limitations in replicating human cognitive flexibility.

The implications of this analysis suggest that while language may provide an initial structure for AI to develop reasoning, this ability remains limited. AI lacks the "linguistic awareness" that enables humans to contextually connect abstract concepts. This reinforces the view that although AI can assist in domain-general reasoning, it is still unable to emulate humans' ability to adapt to truly novel situations.

From a practical perspective, these limitations have significant consequences for AI applications. For example, in ethics-based decision-making, AI often fails to capture moral nuances intrinsically tied to human experiences framed by language. Thus, Rothschild's claim that AI can fully replace human reasoning in a domain-general context appears overly optimistic.

RESULTS AND DISCUSSION

The main finding from large-scale artificial intelligence (AI) experiments is the revelation of the significant role of language in thought and general reasoning. Large Language Models (LLM), such as GPT-4, demonstrate that natural language-based systems possess extraordinary domain-general reasoning capabilities, surpassing other AI systems in terms of inference breadth (Devlin et al., 2018). This indicates that language, as a symbolic system, plays a central role in supporting cognitive processes.

The surprising performance of LLM reveals that natural language-based representations are highly data-efficient. Language enables systems to make abstract inferences with fewer computational resources compared to other approaches. For instance, the ability of LLM to generate complex, profound, and coherent text proves that linguistic representation is a powerful tool for symbolic processing in cognitive reasoning (Fedorenko et al., 2024).

These experiments also highlight the strength of the subsymbolic architecture underpinning LLM. Before the era of LLM, artificial neural network (ANN)-based systems were often regarded as limited to specific patterns, such as visual recognition or board games (Rumelhart et al., 1986). However, advancements in LLM have changed this perspective, showing that ANN can achieve complex reasoning levels by integrating subsymbolic and symbolic approaches (Devlin et al., 2018).

While the success of subsymbolic architectures is evident, the results of these AI experiments also affirm the importance of symbolic systems, particularly natural language, in supporting domain-general reasoning capabilities. As a symbolic representation system, natural language enables LLM to process data efficiently and generalize across various contexts. For example, the ability of LLM to comprehend and respond to abstract questions reflects the transformative nature of language in enhancing cognitive processing capabilities (Fedorenko et al., 2024).

The results of extensive experiments in AI development reveal that only AI systems extensively trained with natural language, such as LLM, are capable of demonstrating significant general-domain reasoning abilities. In contrast, non-language-based AI systems lack comparable capabilities in performing such tasks (Rothschild, 2024). General-domain reasoning refers to the ability to integrate information from various domains, such as agents, locations, or quantities, into a single coherent conclusion. In this context, LLM, such as GPT-4 or Claude, have shown superior cross-domain reasoning compared to other AI systems (Mahowald et al., 2024).

Language-based AI systems not only outperform other models but also surpass humans in certain tasks involving general reasoning. Nonetheless, LLM still face limitations, particularly in reasoning that requires precise logic, mathematics, and handling long arguments (Dziri et al., 2024). However, overall, their ability to perform general inference far exceeds that of ANN models that do not utilize natural language (Rothschild, 2024). This indicates that the success of

LLM is not solely dependent on design or training for limited purposes but also on the central role of natural language in supporting general reasoning.

One of the main reasons LLM excel in general-domain reasoning is the capability of natural language to simplify complex information. Natural language allows for dense yet relevant descriptions, reducing the computational burden in the inference process (Rothschild, 2024). For instance, a low-quality 30-second video requires approximately 2.5 megabytes of storage space, while the entire novel *War and Peace* takes up only 3 megabytes (Mahowald et al., 2024).

This highly compressed and abstract nature of natural language enables AI, particularly LLM, to manage vast amounts of data and make efficient token-based predictions (Dziri et al., 2024). In contrast, AI systems operating in domains like video face significant challenges in performing equivalent inferences because visual data lacks a similar level of abstraction and simplification (Rothschild, 2024). Thus, natural language functions not only as a communication tool but also as a highly efficient medium for information processing in AI systems.

LLM, such as GPT-4, stand out as AI systems that demonstrate significant success in general-domain reasoning. This success is evident in their ability to regularly and consistently make inferences across various domains, including language, logic, mathematics, causality, and decision-making (Mahowald et al., 2024). This achievement is not merely the result of large-scale data training but also evidence of these models' ability to generate original texts based on user prompts (McCoy et al., 2023).

LLM have succeeded in surpassing merely repeating their training data. Although they are trained using nearly all publicly available texts, the process of selecting relevant texts to answer questions constitutes a form of complex problem-solving (Millière & Buckner, 2024). In scenarios where the training data does not include both prompts and responses, LLM demonstrate their ability to make inferential decisions rather than simply memorizing data (McCoy et al., 2023).

The ability of LLM to infer can be observed through how they respond to hypothetical situations. For instance, when given a scenario involving a roll of toilet paper on an uneven table, GPT-4 can provide a logical explanation of the object's behavior based on basic principles of physics. While these inferences may appear simple, the results reflect a type of non-deductive reasoning often required by humans in everyday life (Rothschild, 2024). This ability indicates that LLM are capable of handling a wide range of cross-domain problems.

In contrast, non-linguistic AI systems, such as those used in facial recognition, autonomous driving, and strategic games, exhibit narrow and domain-specific capacities. For example, AI systems designed to play board games like Go require thousands of hours of training to achieve high performance. This contrasts with humans, who can learn more quickly by leveraging general knowledge (Lake et al., 2017).

Analysis suggests that the domain-general reasoning ability of LLM is strongly tied to the representational efficiency enabled by natural language. Natural language facilitates the efficient abstraction of information, including only details relevant for drawing conclusions. For example, a simple message like “your bike was stolen” can convey essential information without the excessive details that might be included in a video recording of the theft (Silver et al., 2018). This abstraction reduces the complexity of the data to be processed, making inference tasks more accessible to neural networks.

Natural language, with its symbolic nature, offers a unique advantage over non-linguistic representational media such as video. Videos carry an excess of information, making the inference process more complex. In contrast, language-based representations include only the information significant for decision-making. For instance, predicting the next event in a text involves far fewer possibilities for the subsequent word compared to the number of possible frames in a video, simplifying prediction and reasoning processes (Shanahan, 2016).

LLM demonstrate extensive inferential capabilities that exceed initial expectations. This ability is largely attributed to the efficiency of language in presenting information in a relevant and structured manner. By encoding information in natural language, models do not merely predict the next word but also exhibit cross-domain reasoning abilities akin to human logical argumentation. For example, complex problems such as strategizing in chess games or predicting textual continuations have been successfully addressed using efficient linguistic representations (Silver et al., 2016).

The ability of natural language to perform symbolic abstraction also plays a critical role in inference-making. An AI trained on linguistic data does not require additional abstraction processes to extract relevant information, as would be necessary in video-based systems. This reduces computational burden and allows neural networks to focus more on inference tasks. This explains why natural language-based models like LLM excel at inference tasks compared to non-linguistic models (Shanahan, 2016).

Experiments in AI have demonstrated the remarkable inferential potential of LLM-based systems. These systems, trained on vast amounts of linguistic data, exhibit significant cross-domain inferential capacities. This supports the view that language serves as an abstraction medium facilitating general human thought, rather than merely being a communication tool. While human language is designed for communication efficiency, studies on LLM provide insights that the abstract structures in language also play a crucial role in enhancing cognitive abilities, particularly in solving non-deductive problems (Frank, 2023).

This perspective challenges the traditional assumption proposed by Pinker, which emphasized that language has little instrumental value for thought. Instead, AI experiments show that language can be an effective tool for creating inferential structures that cognitive systems can leverage (Carruthers, 2002). Therefore, these findings lay a new foundation for understanding the role of language in human and AI cognition, showing that language is more than a mere reflection of human cognition; it can also shape it.

Empirical research shows that many abstractions in natural language not only reflect human cognitive capacity but are also directly encoded in the language itself. LLM provide evidence that language can create pathways for strong inferential abilities without relying on inherent cognitive structures. This affirms the possibility that language learning can provide significant cognitive transformation, both in humans and AI systems, as observed in LLM experiments (Fedorenko et al., 2024).

However, it is important to note that humans have access to vast non-linguistic data, both through individual learning processes and evolution, which LLM do not have. Therefore, human inferential abilities may be a unique combination of linguistic and pre-linguistic processing. The abstractions made by the human brain often reflect innate processes, such as the ability to understand objects, agents, and causality, which have been shown to exist even in non-linguistic beings.

One of the key findings from LLM experiments is that training on large linguistic data not only influences linguistic processing but also enhances domain-general inferential abilities. This strengthens the argument that language plays a transformative role in shaping human thought, both through its communicative abilities and its capacity to enable cross-domain inference (Frank, 2023).

These results provide important insights into the role of language in thinking, both in humans and AI. While many classical arguments, such as the Language of Thought (LOT) hypothesis, suggest the existence of a pre-linguistic representational system underlying human thought, empirical evidence from LLM experiments shows that language itself can drive cognitive abilities (Quilty-Dunn et al., 2023). Thus, this research strengthens the argument that language is not just a communication tool but also a key element in the development of human cognitive abilities.

CONCLUSION

In this study, Rothschild's theory on the role of language in shaping our reasoning has been discussed, providing significant insights into understanding both artificial intelligence (AI) and human intelligence. Rothschild's argument suggests that AI's inferential capabilities, when trained using language and leveraging large linguistic datasets, outperform other types of AI. This is due to the abstract nature of natural language, which allows us to convey what we care about with minimal data. This phenomenon opens up opportunities for AI systems to receive more useful training in a shorter period compared to AI trained on more naturalistic data. This approach also illustrates that the efficiency of linguistic representation may have a profound impact on our mental development as humans. Language enables us to represent abstract and complex concepts more efficiently than other media, such as visual or numerical representations. In this context, language not only functions as a means of interacting with the external world but also as an internal tool that shapes how we think and process information. When we connect this understanding to AI, the role of language in enhancing AI's inferential capabilities becomes clearer. AI systems trained using linguistic data can mimic how we structure reasoning and make decisions, allowing them to learn from more varied and relevant examples. Therefore, this research not only offers a new perspective on how language plays a role in the development of human intelligence but also opens further potential for the advancement of more sophisticated AI. Rothschild's theory provides an important foundation for understanding the relationship between language and reasoning, both in humans and AI systems. The efficiency of linguistic representation found in natural language is key to enhancing inferential abilities, which in turn leads to advancements in AI technology and a better understanding of human mental processes. Moving forward, further research is needed to explore more deeply how language shapes cognitive processes and to develop AI applications that increasingly resemble the way humans think and act.

REFERENCES

Ali, A. (2023). The power of language: Exploring how language shapes our understanding of the world. *Policy Research Journal, 1*(1), 9-14.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021, March). On the dangers of stochastic parrots: Can language models be too big? In *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency* (pp. 610-623).

Carruthers, P. (2002). The cognitive functions of language. *Behavioral and Brain Sciences, 25*(6), 657-674.

Creswell, J. W. (2014). *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*. Thousand Oaks, CA: SAGE Publications.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2018). BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.

Felin, T., & Holweg, M. (2024). Theory is all you need: AI, human cognition, and decision-making. *Human Cognition and Decision Making*.

Fedorenko, E., Piantadosi, S. T., & Gibson, E. A. (2024). Language is primarily a tool for communication rather than thought. *Nature, 630*(8017), 575–586.

Frank, M. C. (2023). Bridging the data gap between children and large language models. *Trends in Cognitive Sciences*.

Hadi, M. U., Al Tashi, Q., Shah, A., Qureshi, R., Muneer, A., Irfan, M., & Shah, M. (2024). Large language models: A comprehensive survey of its applications, challenges, limitations, and future prospects. *Authorea Preprints*.

Kaijanaho, A. J. (2015). *Evidence-based programming language design: A philosophical and methodological exploration* (Doctoral dissertation, University of Jyväskylä).

Kalyan, K. S. (2023). A survey of GPT-3 family large language models including ChatGPT and GPT-4. *Natural Language Processing Journal, 100048*.

Kolides, A., Nawaz, A., Rathor, A., Beeman, D., Hashmi, M., Fatima, S., & Jararweh, Y. (2023). Artificial intelligence foundation and pre-trained models: Fundamentals, applications, opportunities, and social impacts. *Simulation Modelling Practice and Theory, 126*, 102754.

Leroy, S., Parisse, C., & Maillart, C. (2012). Analogical reasoning in children with specific language impairment. *Clinical Linguistics & Phonetics, 26*(4), 380-396.

McCoy, R. T., Smolensky, P., Linzen, T., Gao, J., & Celikyilmaz, A. (2023). How much do language models copy from their training data? Evaluating linguistic novelty in text generation using Raven. *Transactions of the Association for Computational Linguistics, 11*, 652–670.

Millière, R., & Buckner, C. (2024). A philosophical introduction to language models – part I: Continuity with classic debates. *arXiv preprint arXiv:2401.03910*.

Minda, J. P. (2020). *The Psychology of Thinking: Reasoning, Decision-Making and Problem-Solving*. London: Sage Publication.

Perlovsky, L. I. (2009). Emotions, language, and the Sapir-Whorf hypothesis. In *2009 International Joint Conference on Neural Networks* (pp. 2501-2508). IEEE.

Pinker, S. (2010). The cognitive niche: Coevolution of intelligence, sociality, and language. *Proceedings of the National Academy of Sciences, 107*(Supplement_2), 8993-8999.

Quilty-Dunn, J., Porot, N., & Mandelbaum, E. (2023). The best game in town: The reemergence of the language-of-thought hypothesis across the cognitive sciences. *Behavioral and Brain Sciences, 46*, e261.

Rothschild, D. (2024). Language and thought: The view from LLMs. In Lepore, E., & Sosa, D. (Eds.), *Oxford Studies in Philosophy of Language*.

Russell, S., & Norvig, P. (2020). *Artificial Intelligence: A Modern Approach* (4th ed.). Pearson.

Shanahan, M. (2016). The frame problem. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 2016 edn). Metaphysics Research Lab, Stanford University.

Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and go through self-play. *Science*. [science.org/doi/abs/10.1126/science.aar6404]

Silver, D., Huang, A., Maddison, C. J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., & Lanctot, M. (2016). Mastering the game of Go with deep neural networks and tree search. *Nature, 529*(7587), 484–489.

Viale, R. (2001). Reasons and reasoning: What comes first. In *L'explication des normes sociales*. Paris: Presses Universitaires de France.