# Exploring Education Cyborg Space:

# Bibliographic and Metaphor Analysis of Educational Psychology and Artificial Intelligence Studies

by

Byoung-gyu Gong

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved April 2021 by the
Graduate Supervisory Committee:

Iveta Silova, Chair
Ruth Wylie
Sherman Dorn
Yi Zheng

ARIZONA STATE UNIVERSITY

May 2021

ABSTRACT

The emergence of machine intelligence, which is superior to the best human talent in some problem solving tasks, has rendered conventional educational goals obsolete, especially in terms of enhancing human capacity in specific skills and knowledge domains. Hence, artificial intelligence (AI) has become a buzzword, espousing both a crisis rhetoric and an ambition to enact policy reforms in the educational policy arena. However, these policy measures are mostly based on an assumption of a binary human-machine relations, focusing on exploitation, resistance, negation, or competition between human and AI due to the limited knowledge and imagination about human-machine relationality. Setting new relations with AI and negotiating human agency with the advanced intelligent machines is a non-trivial issue; it is urgent and necessary for human survival and co-existence in the machine era. This is a new educational mandate.

In this context, this research examined how the notion of human and machine intelligence has been defined in relation to one another in the intellectual history of educational psychology and AI studies, representing human and machine intelligence studies respectively. This study explored a common paradigmatic space, so-called ‘cyborg space,’ connecting the two disciplines through cross-referencing in the citation network and cross-modeling in the metaphorical semantic space. The citation network analysis confirmed the existence of cross-referencing between human and machine intelligence studies, and interdisciplinary journals conceiving human-machine interchangeability. The metaphor analysis found that the notion of human and machine intelligence has been seamlessly interwoven to be part of a theoretical continuum in the most commonly cited references. This research concluded that the educational research and policy paradigm needs to be reframed based on the fact that the underlying knowledge of human and machine intelligence is not strictly differentiated, and human intelligence is relatively provincialized within the human-machine integrated system.

Keywords: Cyborg, Artificial Intelligence and Education, Citation Network, Metaphor Analysis, Educational Paradigm

DEDICATION

*For my lovely wife Yuseung, and*
*my son Jason Sehyun*

ACKNOWLEDGEMENT

This paper would not have been possible without the support of my professors, colleagues, and friends. First of all, I would like to express my deep gratitude to my dissertation committee chair and advisor Dr. Iveta Silova. She inspired and encouraged me to continue this study, sharing thoughtful advice and feedbacks all the time. This research was solely possible under the unique academic culture she has cultivated with her students at Arizona State University (ASU), experimenting with diverse methods and perspectives to illuminate our education from a different angle. We discovered and nurtured our Cyborg, the core concept of this dissertation research, within this academic culture, and through constant engagement in deep dialogue and thought experiments. My dissertation committee member, Professors Sherman Dorn, Ruth Wylie, and Yi Zheng, were a great support team who contributed to my research through inspiring feedback. Dr. Sherman Dorn encouraged me to develop this research topic in the Research Design class. His introduction of the book *An Elusive Science: The Troubling History of Education Research* written by Ellen C. Lagemann was a decisive trigger to develop my research idea. Dr. Ruth Wylie provided a thoughtful reflection on cognitive science and educational psychology's intellectual trends and history. Her creative and imaginative activities in the Center for Science and the Imagination at ASU were a source of motivation to pursue my research topic. Dr. Yi Zheng has always been supportive of my various statistical research. She helped me analyze a bulk of bibliographic network data and carefully reviewed my code line by line. We have also discussed the rise of machine learning in psychometric research and the implications of this new trend.

Besides, this research was possible only in the ASU Mary Lou Fulton Teachers College (MLFTC) and its dedication to providing sufficient financial support and intellectual resources to the doctoral students. I can confidently say that I could not have finished this research in other institutions because the help and support of ASU and MLFTC were truly enormous and decisive to my study in many ways. Many other colleagues and friends read, reviewed, and listened to my research and gave me supportive feedback and encouraging comments. I am thankful to Woo Yeoung Kim, Dr. Sung Sang Yoo who shared their speculations on the scientification of education. I am also grateful to the Class of 2016 of the Educational Policy and Evaluation Program who gave me feedback in our classwork.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

# CHAPTER 1. INTRODUCTION

## Background

*Human-like artificial intelligence (AI) is not sci-fi anymore. It is already here.* Google kicked off the year of 2020 with the announcement that their advanced AI program had surpassed human doctors in predicting breast cancer with a mammography screening (McKinney et al., 2020). As the AI technology becomes widely adopted by the business sector, companies use the machine to automatically detect less productive workers and fire them (Jee, 2020). In terms of language ability, the Chinese AI engine, Baidu, scored 90 out of 100 in the General Language Understanding Evaluation (GLUE) benchmark test, surpassing the average human score of 87 (Hao, 2019; The General Language Understanding Evaluation, 2020). It is a non-trivial achievement considering that the GLUE includes human-level reading comprehension such as identifying a pronoun *it* from multiple candidate nouns, as well as detecting topic, logic, and common sense and knowledge. Much more substantive progress has been unraveling in the self-driving auto technology area. The Tesla company promotes its full self-driving package equipped with almost all human drivers' features and has recently announced its full self-driving plan. In 2019, autonomous trucks crossed the US from west to east in three days as they could run day and night without sleep. It is not even eye-catching news anymore that AI programs defeated human players in various games such as chess, Go, and Starcraft. AI robots now train themselves to find rules and patterns without any human input, with emergent intelligence technology. For instance, the OpenAI company created a robot hand that detects Rubic’s cubic problem solution and plays with it without any given initial clue of this object. Nasr et al. (2019) found that the deep neural network designed for visual recognition learned the concept of quantification and abstract mathematical sense without explicit programming. Despite much skepticism about AI's true potential, it seems clear that living in a world transformed by AI technology is simply inevitable today.

AI signals a technological advancement instrumental to our comfortable life and transforms our world from the fundamental level. More and more people believe that AI would bring cascade effects to industrial transformation and our society more broadly. Either with a positive or negative outlook, most scholars agree that the emergence and broad application of automated robots will mark a new revolution in human history, commensurate with the previous agricultural and industrial revolution (Bostrom, 2014; Harari, 2014; Schwab, 2016). From the perspective of big history, Harari (2014) saw that with the invention of AI, human evolution does not rely on natural selection anymore but more on intelligent design, which can create another independent life with a digitized human mind. Brynjolfsson and McAfee (2014) insisted that we had enhanced and augmented our physical body for the last two centuries since the industrial revolution initiated by the invention of a steam engine, which he named as *the first machine age* and then now is the moment of *the second machine age* wherein the intelligent machines enhance our mental force. Schwab (2016), a founder of the World Economic Forum (WEF), called this change the fourth industrial revolution, defining it as a megatrend driven by industrial automation. He saw that our next generation's significant challenges would come from this fundamental change in the industrial system. Bostrom (2014) considered that human civilization would reach a testing moment sooner or later due to the enhanced AI that outperforms humans in most intellectual domains, shifting the ontological ground of human-being. Despite small variations in their perspectives, AI scientists and scholars commonly assume that there will be a point-of-no-return where the ontological ground of human-being will be significantly changed due to the outperforming AI.

AI's invention also marks the end of human centrism, a continued epistemological legacy since the renaissance. We, humans, self-nominated ourselves as the most intelligent beings on earth, calling ourselves Homo Sapiens, the wise man. In this way, intelligence superior to other beings was a strong basis for recognizing ourselves as privileged beings. However, as the machinic intelligence emerges with AI technology, this presupposition becomes no longer tenable. Mazlish (1993) argued that the discontinuity assumed between humans and machines was demolished with the advent of intelligent machines. He recognized that any refusal to accept the machine as an intelligent being, whether it is fear or resistance, stemmed from our centrism sustaining a belief that we, humans, are unique beings not replicable by the other creatures. As an extension to Jerome Bruner, who insisted on three discontinuities assuming human as a pure, rational, unique, and superior creature, he called this persisting perception denying a connection between human and machine as *the fourth discontinuity*. He predicted that we would plan our future with these machines only after overcoming the discontinuity assumed between humans and machines. In a very similar manner, Haraway agreed that the development of machine intelligence is “the fourth wound” (Gane, p.141) to narcissist humanity. According to her statement, this new invention of machines implies that even machines can be lively like humans. She captured this point very succinctly by saying, “Our machines are disturbingly lively, and we ourselves frighteningly inert” (Haraway, 2016, p. 152). Like Mazlish (1993), she wanted people to accept the continuity between human and machine, being playful with the boundary to envision our future in the more flexible way  where we are decentered from our universe and more open to other beings.

As robots proved their agility and intellectual superiority over humans, people started recognizing the disruptive AI technology's possible impact. The economists were one of the first to notice this change and explored its disruptiveness. The National Bureau of Economic Research (NBER) invited economists to the first conference on AI economics in 2017, and the proceedings were published as a book of *The Economics of Artificial Intelligence: An Agenda*. This book included the most extensive and systematic reviews that have been made so far by social scientists concerning the impact of AI technology on our society. In this book, the economists discussed the future social transformation based on expert-level knowledge of the most recent AI technology among prominent AI scholars. The papers covered the mechanism of AI, its impact on innovation practice, and its socio-economic impact.

Although pessimists do not believe in dramatic technology-led economic development, this book's authors generally agreed that innovative technology could open up the next chapter of human history, boosting the quality of life and economic productivity. For these economists, AI, particularly the most recent development of machine learning (ML) algorithm revolutionizing AI mechanism, is the general-purpose technology (GPT), a game-changer of business, industry, and research mechanism, a significant upheaval to our existing social order. The GPT is a technology of technology, that can set new paradigm, and has high applicability across many different sectors. Examples of GPT are steam engines, electric motors, computer chips, and even human intelligence. The economists said that whenever there was an invention of GPT, it revolutionized our civilization, bringing industrial revolutions. The AI is even regarded as the last invention because it imitates human intelligence, one of the most versatile GPTs, and even aims to surpass it.

The economists believe AI to be the GPT for several reasons. Brynjolfsson et al. (2019) insisted that AI can be applied to a wide range of tasks in the current human occupation, more than 45% of the total 2,000 distinctive human tasks across occupations, and cultural and institutional resistance to the change is the only last barrier remaining at this moment. Taddy (2019) classified ML a GPT because it is getting cheaper and faster with a linear performance increase over time. Cockburn et al. (2019) saw that the invention of ML would bring a fundamental change in scientific inquiry tradition**,** shifting its focus away from mini-scale causal effect to massive multi-causal effect. They predicted that the ML would reshape the research and development (R&D), previously intensive but inefficient search tasks. Agrawal et al. (2019) insisted that given the fact that innovation occurs with a combination of existing knowledge, the ML, an algorithm good at combining accumulated big knowledge would take over human researchers' role in creating innovation.

Despite skepticism about AI's revolutionary impact, another consensus among the economists in this book was that the policy intervention should be ready for future economic transition led by automation. Although the authors generally agreed that AI would have a positive impact in the long term, they unanimously subscribed to the short-term projection that social instability is inevitable due to massive job loss. They predicted that the consequence of automation in the short and medium-term would be a widening gap between winners and losers. Firstly, there was a concern over the speed of automation in the industry, assuming that overhype of automation in the business and public sector in a relatively short period would bring a social crunch. The economists warned that the possible mismatch between labor demand and workers skills would translate into a vast and painful social burden (Acemoglu & Restrepo, 2019; Bessen, 2019). Secondly, some of them were concerned about AI technology as it tends to replace human labor rather than enhance human capability. They suggested an extreme scenario that the AI would possibly replace human work almost 100%, considering its speed of improvement and sophistication at the current speed (Korinek & Stiglitz, 2019; Trajtenberg, 2019). This group of economists warned that our society should be ready for a fundamental restructuring of our job, economy, and life, whatever the speed of change toward automation we have.

*Now, governments and international organizations scrutinize AI in order to predict our future in the longer term.* Recently**,** the US government embarked on a more systemic AI technology study to envision its broader social impact. The special report of the Executive Office of the President (2016) set the tone of policy intervention to prepare for the AI-triggered social transformation. This report predicted that 47% of current occupations would be destroyed by automation in a decade in its worst-case scenario. The report also saw that the most affected by this change would be low-income and low-skilled workers. Based on this argument, the report suggested three policy directions, including enhancing research and development (R&D) capacity for AI, supporting education and training of the workers, and securing a social safety net for the massive job transition. According to this report, the US government scaled up its effort to be ready for the AI technology transformation. The White House launched the American AI Initiative with a clear intention to maintain its industrial and research strength against the other international competitors. This initiative pledged to prioritize the AI sector and AI workforce development with a long-term investment in education and training. The federal government pledged to harness AI technology for government service and policy mechanisms. The impact of the US federal government's automation would be immense considering its power to set the standard of state bureaucracy and grant funding to research.

At the international level, the Organization of Economic Cooperation and Development (OECD) has provided a serious policy and strategic approach in the advent of AI technology. In 2019, the OECD member countries declared the OECD Principles of AI, emphasizing that AI development should be responsible for its impact on human society (OECD, 2020). The principles focused on promoting democracy, humanitarianism, and social justice through AI technology. The OECD also mentioned that increasing transparency and explicability of the AI outcomes and mechanisms is the key to deploying AI technology with enhanced human-machine interaction. Also, G7 Summit released the Statement on Artificial Intelligence in 2018, highlighting human-centric, inclusive, and trustworthy AI innovation.

In East Asian countries such as China, South Korea, and Japan, the divide between sci-fi and reality has already collapsed as the Google engineered AI program AlphaGo defeated the world human Go champion with a considerable margin in 2016 South Korea. It was indeed a Sputnik moment for the East Asian countries where the people have played the Go more than 2,000 years and top players have been revered as geniuses (Perez, 2017). Since then, AI technology has begun to dominate public discourse.. Also, people in East Asian countries have requested substantive government intervention plans to address the AI technology gap with the Western countries and enhance the social safety net in preparation for the expected massive job loss. In reaction to the AlphaGo shock, the East Asian country governments allocated a considerable budget to boost AI technology and established a governance system to support a social transition during the next industrial revolution (Perez, 2017, Asia Pacific Foundation of Canada, 2019).

**Problem Statement**

Despite the growing demand to study AI technology's social impact, there is not enough research in the education field on this issue. Such inattention is at odds with the emphasis on educational policy intervention in most current AI-related policy recommendations. Only recently, some scholars in education embarked on serious scrutiny of the intersection between disruptive AI technology and educational policy and practice. There are possibly three dimensions of the interplay between education and AI: (1) paradigmatic interplay at the research level, (2) practical application of automated machines for teaching, learning, and decision-making process, and (3) educational policy reform to prepare for automated industry. Nonetheless, most researches focused on the second and third dimensions of interplay, leaving  the first one relatively unexplored.

Firstly, studies explore AI technology's practical application for teaching, learning, and educational decision-making process, estimating the effect of automation and digitization of educational processes. On the one hand, some studies explore AI in education with the utilitarian perspective, emphasizing this technology's effectiveness in relation to the students' learning and the school administrative system. An intelligent tutoring system (ITS) runs on AI engines customizing and individualizing students' learning at the class level. At the administrative level, there are attempts to adopt a data dashboard, achievement prediction modeling, and automated student selection system to replace humans with algorithmic decision-making (Gulson & Webb, 2017a; 2017b; Marcinkowski et al., 2020). With this new AI-driven educational technology, scholars are interested in the effectiveness of this AI learning and administrative platforms compared to the conventional human-based system (see Corbett et al., 1997; Kulik, 2016; Ma et al., 2014). One of these studies' focus is identifying the comparative advantage of machine-human interaction against human-human interaction to enhance students' learning. There has been still controversy over the effectiveness of machine-human interaction. There have been tensions between enthusiasts who advocate widespread technology diffusion in education and pessimists who criticize education technology as overhyping or a waste of money.

On the other hand, some studies examine AI technology's application in education with a more critical perspective. These studies focus on the expansion of datafication and computational rationality of educational policy. This group of critical studies considers the increasing usage of data science, learning applications, an AI engine for the prediction of students' performance, and rationalization of administrative allocation system as a symptom of big-brother governance (Decuypere, 2019; Gulson & Webb, 2017a; 2017b; Means, 2019; Sellar & Gulson, 2019; Williamson, 2016). They insist that the black-box type of AI mechanism, an unexplainable decision-making procedure, can cause ethical problems or even problems beyond our expectations. Also, they are concerned about the machines' biased decisions as they learn from human input, including a lot of noise and bias in it. The monopoly of the data infrastructure by the tech giants is also a concern for their critical research.

Secondly, future work and labor are also foci of research interest, as the automated industry can significantly change our society in general. Thus, studies concentrate on workforce development concerning the replacement of human labor with the machinic forces. These studies emphasize educational policy reform to prepare for the automated industry considering education as an essential component to enhance STEM and R&D workforce development and instrumental to transforming society into a whole new era. However, there are two contrasting opinions surrounding the real impact of industrial automation. The educational policy reform scenario is substantially different depending on the belief in the potential power of AI. One group of scholars has serious concerns about the rising intelligent machines estimating their capacity to closely approximate human capacity. Economists such as Sachs (2019), Korinek and Stiglitz (2019), and Cockburn et al. (2019) consider automation as a real challenge as it can substitute a significant portion of human labor, and even high-skilled workers are not an exception.

On the other hand, some studies consider AI technology as not being near to the human level capacity. They believe that humans will find their niche protected against the intelligent machines despite their increasing capacity. The belief in such a scenario is only possible as they assume fundamental discontinuity between human and machine intelligence. For instance, Aoun (2017), the author of *Robot-proof: Higher Education in* *the Age of Artificial Intelligence*, insisted that future education should be designed to enhance “uniquely human cognitive capacity” (p. xviii), and raise our students to be creators. Similarly, Trajtenberg (2019) emphasized creative skill as the key for the workforce adaptation to the machine age, insisting on an educational revolution toward enhancing uniquely human skills. Oleinik (2019) insisted that the current AI model of neural networks is inherently limited to catch up with human creativity as it cannot generate metaphorical expressions, interact with other entities creating social connections, and predict unexpected patterns that the existing data do not provide.

Despite this recent research boom exploring AI’s impact on education, a more fundamental level of intersection between AI and education at the paradigmatic and epistemological level remains unexplored. The scholars exploring AI's practical application in education tend to lean toward either of the opposite extremes: technophilic or technophobic perspectives. The former emphasizes the instrumental and utilitarian value of the AI for educational enhancement, while the latter frames human and machine relationships in terms of a dualistic rivalry. It reflects the fact that we perceive this technology as somewhat alien to ourselves. However, it seems misleading to consider the permanent divide between machine and human intelligence to be taken for granted, given the intellectual history wherein human and machine intelligence studies have been interrelated. As Mazlish pointed out, machine and human science is a “theoretical continuum”(Mazlish, 1993, p.6), cross-referencing and imitating each other. It is not a coincidence that the human brain and computing machines' comparative studies came to the fore at every critical historical juncture of the computing machines. For instance, John Von Neuman, a father of the modern computer, explained how the human brain and computing machine could be explained with the unified theory of information processing in his famous book, *The Computer and the Brain*. At the dawn of machine learning (ML), Jeffrey Hinton and his colleagues, fathers of machine learning, published a book of *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. This book suggested parallel distributed processing (PDP) as a new architecture of machine intelligence imitation of the human brain process. In human science studies, the machine metaphor has been used to understand the human mind already for more than a century of its history. The physiological laboratory tradition established by Wilhelm Wundt (1832-1920) made people conceive of the human mind as a materially constructed being. The computer machine's invention made scholars in human science imagine the human brain as an advanced information processor, which can process abstract symbols (Abrahamsen & Bechtel, 2012; Glaser, 1984; 1991).

In that account, AI technology is a tool, paradigm, and epistemology of educational studies, especially for learning science and educational psychology interconnected with brain science. The intelligent machines reflect our zeitgeist, believing that the full features of human intelligence are based on materiality to be imitated mechanically. Unfortunately, the existing literature about AI technology, including education studies, have largely overlooked an intellectual history of the interplay between human and machine science studies, profoundly engaging with each other in creating a common knowledge of intelligence. Sciences of a machine and human intelligence are extensions of the human desire to strengthen our control and prediction power over the world by programming intelligence. They use the same language to represent information processing in biological and mechanical circuits. Thus, as Graham (2002) insisted, such advanced human science and technology made us reconsider a conventional way of defining a human as a purely natural organism, while the mechanical device of high intelligence was gradually "assimilated into nature as a fully functioning component of organic life itself" (p. 53).

**Research Purpose**

This research lies where the modern conception of human and machine intelligence is co-constructed or at least emerged as a twin birth. There is a growing need to explore this grey area where human and machine intelligence converge and are not conceptually differentiated as such hybridity was a force of framing our education as an education science, scientific modeling of human learning. This research hypothesizes that the underlying assumption on the human mind, intelligence, and behavior framing our modern education shares a common knowledge with machine science, which tested various hypothetical models of human-like intelligence and its optimization.

Given this hypothesis, this research intends to describe how the paradigmatic hybridity between human and machine science studies is textured. This research examines how the studies of artificial intelligence and education science have interacted with each other by analyzing (1) the number of standard references they share as its paradigmatic commonplace and (2) the main conceptual anchors connecting these two seemingly distanced disciplines. With this extended recognition of the convergence between the two, this research ultimately strives to get a new sense of reality: Our modern education has been situated in parallelism between "mechanization of mind and humanization of machine" (Dupuy, 2010b, p.229), thus ambiguously assuming the human mind as something in-between machine-like mind and human-like machine.

Accordingly, the research questions are set as follows: What is the interplay between the studies of education science and AI studies at the paradigmatic level?

- To what extent and how do they converge with each other? (Quantitative study)
- What do they share as standard features in their basic conceptualization? (Qualitative study)
- What are the implications of such convergence for the future of the education field?

**Definitions**

In this research, the education science means a subfield of education studies having mechanical perspective in the human mind and possibly including educational psychology, learning science, and cognitive science in education studies. These studies predominantly use scientific and mathematical models, mostly borrowed from the hard sciences such as biology, physics, and statistics, to explain humans' mental phenomena and improve human learning. The pursuance of this scientific feature in the education science studies is well described in the paper of the National Research Council (NRC, 2000, p. 3):

> The essence of the matter, the origins of the universe, the nature of the human mind—these are the profound questions that have engaged thinkers through the centuries. Today, the world is amid an extraordinary outpouring of scientific work on the mind and brain, on the processes of thinking and learning, on the neural processes that occur during thought and learning, and on competence

development. The revolution in the study of the mind that has occurred in the last three or four decades has important implications for education.

This study specifically selected educational psychology to represent the education science because its size and impact are immense in educational studies' history (Lagemann, 2000; Mayer, 1992). From behaviorism to cognitive psychology, educational psychology has always been at the center of the education studies' research paradigm and continuously influenced educational policy and practice (Anderman, 2011; Mayer, 1992). Educational psychology scholars have been pushing forward scientification and rationalization of educational research and practice by framing educational psychologists as scientist-practitioners (Hagstrom et al., 2016) and diffusing medical discourse into the education sphere using diagnostic terms such as learning disability (Mehan, 2014).

On the one hand, many people believe that AI targets human like cognitive behavior. Khakurel et al. (2018) define AI in this way: "AI can be described as a cluster of technologies and approaches, that is, statistical and symbolic that aim at mimicking human cognitive functions or exhibiting aspects of human intelligence by performing various tasks…" (p. 2). However, people often misunderstand one of the specific AI techniques machine learning (ML), as the entire AI technology community. A deep neural net, the core of the ML, is a massive parallel distributed model processing information in a complex network of superficial multilayered nodes and their connections exchanging input and feedback. Although the human neuron initially inspired the deep neural net, the technology evolved into much more abstract models combined with various statistics and mathematical theories. Recent AI research is goal-oriented, pursuing the best performance and prediction rate, not its modeling's explicability and generalizability. Thus, people believing that ML is the AI are easily misled to conclude that AI is not related to human-like intelligence. However, the reality is much more complex, and AI is not defined in a single term and field of practice because AI is such a massive assemblage of various sub-disciplinary studies, and it has even constantly evolved into various shapes throughout history (Crane, 2003; Konar, 2000).

# CHAPTER 2. RESEARCH FRAMEWORK

## Paradigm

Paradigm is an essential concept in this research, which aims to explore the common area between the two distinct fields of education and machine science, or the so-called cyborg space. Thomas. S. Kuhn first popularized this term and definition. He said that the scientific community is based on joint prior achievement providing foundational knowledge and practice, what he termed *normal science* and more broadly used the term *paradigm* to explain it. He defined a paradigm in terms of "some accepted examples of actual scientific practice - examples which include law, theory, application, and instrumentation together - provide models from which spring particular coherent traditions of scientific research" (Kuhn, 1996, p.10). The paradigm indicates consensus and commitment – constituted of the standardized rules and theory to follow. Hence, the paradigm is maintained by "strong network of commitment" (p. 42) at the conceptual and methodological levels alike. The scholars can comfortably dwell in this consensual research practice because the existing knowledge and theory provide well-defined problems and questions for their research. However, Kuhn himself never clarified research methods to identify and substantiate the existence of a shared paradigm in scientific research (Small, 2003). His research was rather conceptual and philosophical, mostly supported by anecdotal and conceptual evidence, drawing criticism for its subjective interpretation of the paradigm and normal science (Haraway, 1972).

**Paradigm Convergence**

Thus, this study refined an elusive *Kuhn's paradigm* to create an operational definition of it in order to identify, measure, and explore the interdisciplinary area where the research paradigms converge. Two approaches provide a methodological frame to study the scientific paradigm more systematically. First, Small (2003) insisted that the scientific paradigm is quantifiable and measurable through its bibliometric structure. The bibliometric structure is the academic community structure represented as a network among research papers and textbooks using their metadata such as authors, keywords, institutions, and cited references. Small (2003) wanted to make Kuhn's elusive paradigm more testable by analyzing this quantifiable bibliometric network structure. He found that Kuhn alluded that communication and referencing patterns - the object of bibliometric analysis - can be evidence of a shared paradigm. Also, Kuhn emphasized that textbooks and references are crucial in reproducing past research achievements and maintaining a homogenous research paradigm among scholarly community members. Therefore, Small (2003) assumed that each research paper could be a proxy of a scientific concept, theory, methodology, and practice, although he admitted bibliometric interpretation is only an approximation to the original notion of paradigm raised by Kuhn. Then, shared keywords or references among these research papers represent the research paradigm.

Second, shared metaphorical expression in the scientific community serves as evidence of a shared paradigm in the scientific community. Donna Haraway argued that Kuhn defined paradigm as a shared disciplinary belief and models used in exemplar research. Accordingly, she conceptualized the paradigm as a shared symbolic generalization built upon common belief, value, and exemplar research focusing on the language and image use in the scientific community. Then, she insisted that the paradigm is composed of metaphors, a symbolic generalization of their modeling, saying "paradigms and their constituent metaphors are eminently community possessions whose principal value lies in their growing points" (Haraway, 1972, p. 5). Accordingly, her research traced the paradigm shift from mechanistic to organismic biology with a change in metaphor from atomism to animism. She insisted that such a metaphorical frame is pervasive in the real research practice, even influencing observation and lab testing.

Therefore, in aggregate, the interplay between the educational science and AI studies at the paradigmatic level means a close communication between the two distinctive disciplines through specific cross-referencing patterns and shared metaphorical language expressions. This research aims to find evidence of such close communication between the two fields through bibliometric and metaphor analysis. Accordingly, this research will quantify the degree of paradigmatic convergence between the two fields through bibliometric analysis by measuring citation network structure and features. This research will also explore the detailed research content to find any language and symbolic evidence representing a shared metaphorical expression between the two fields.

## Cyborg: A Paradigm Convergence of Human and Machine Studies

The concept and analogy of cyborg, which have been originally used to represent the human's transformed material basis, are irrevocably shifting due to the advanced technology. According to Muri (2006), Manfred Clynes and Nathan Kline were the first to coin the term cyborg to represent "self-regulating human-machine system" (p. 3). The cyborg sometimes signaled humanity's bright future, expanding limited human capability beyond our material constraints by enhancing or transforming our physical body with new technology. Such a technophilic perspective or so-called transhumanism became a dominant paradigm of our modern society. Fukuyama (2004) pointed out the pervasiveness of transhumanistic desire in modern techno-science, occupying a dominant position that frames our socio-cultural fabric. In particular, he pointed out that the recent advancement of bio-medicine, gene therapy, and brain science reflects human desire to enhance itself to approximate the eternal being. Grunwald (2011) defined the modern world as a performance-enhancement society wherein everyone is obsessed with their enhancement to be in a better social position. In this social atmosphere, transhumanistic vision is not a radical idea anymore but everyone’s dream, thus being readily accepted and desired by the general public (Grunwald, 2011). In particular, the recent development of new technologies such as nanotechnology, biotechnology, cognitive science, artificial intelligence, and robotics reshaped the condition of human life. It enhanced transhumanists’ belief that our current definition of humanity is one in the permanent-transitory human evolution process (Jeffrey, 2016).

The general message from the most recent development of the transhumanists’ techno-philia is that the human mind and body are a materialistic, physical, and mathematical structure; thus, we can code and recode our operational functions to emulate its structure and function with reverse engineering technology. This paradigm's intellectual roots are cyberneticians of the early 20th century, including Norbert Wiener and Claude E. Shannon, who created the mathematical theory of machine and animal information processing. They claimed that the mathematical theory of the information process constituted of input, output, feedback, and control is commonly applicable to animate and inanimate beings and humans are not an exception (Wiener, 1985; 1989). On that account, Ray Kurzweil (2012a) emphasized the human brain and computer parallelism, explaining that “statements along these lines (the brain is not a computer) are akin to saying, “Applesauce is not an apple." Technically that statement is true, but you can make applesauce from an apple…a computer can become a brain if it is running brain software. That is what researchers, including myself are attempting to do” (p. 181). Kurzweil (2012a) argued that this brain-machine parallelism is a source to build a better intelligent machine for the next step of evolution. He suggested that biological understanding can inspire the creation of machine intelligence. The ultimate human enhancement through mechanical structure and function is the common desire lurking inside transhumanists’ dreams. Kurzweil's vision of uploading the mind to the computer is one of the outgrowths of technophilic vision to enhance humans beyond the organic frame.

On the other hand, the cyborg has represented a dystopian techno-future where the machinic transformation severely impairs humanity's purity and integrity. Muri (2006) found that films and fiction used cyborg to describe “troubled, dark, corrupt, or post-apocalyptic future” (p. 4). In academia, there has been a prolonged technophobic sentiment. Religious critiques insist that once our mind escapes the organic body, we will lose our soul, the fundamental essence of the human endowment from God (Lilley, 2013). According to Zimmerman (2011), Heidegger warned against the techno-utopianism concerning human instrumentalization to serve the techno-industry. Zimmerman (2011) said that “According to Heidegger, even though humans may think themselves to be in charge of technoscience, in fact, we are servants of the technological juggernaut. Technology is no longer a means to human ends, but rather an end in itself" (p. 104). He also pointed out that Heidegger was expressly against computer technology, calling it "the self-release of being into machination". This release takes a man into unconditional service. In a similar vein, Dupuy (2010b) denounced transhumanists' endeavor, especially the cybernetics movement led by Norbert Wiener, calling it "a decisive step in the rise of antihumanism" (p. 228). Dupuy (2010b) claimed that cognitive science's primary task is demystifying the human mind as a mere machine kind, only threatening humanity. Francis Fukuyama (2004) also disputed transhumanism, calling it a dangerous scheme threatening American egalitarianism and liberalism. He considered that human equality was built upon an assumption that all human beings share some uneducable human essence, which became a bottom line to establish political liberalism.

The posthumanists' cyborgism suggested by Donna Haraway and her colleagues provided a new perspective of neither technophilic nor technophobic point of view. Through her academic essay *Cyborg Manifesto*, Haraway (2016) suggested what we, especially for feminists, should be after this technological and material-based-modernity, where all the traditional boundary surrounding humans has irrevocably shifted. Cyborg was her creation to represent the hybrid and flexible boundari-ness of our being, against the humanistic and human-centered classical Western world view, which emphasized the fixed purity and essence of being human. She defined the cyborg as "a cybernetic organism, a hybrid of machine and organism, a creature of social reality as well as a creature of fiction" (Haraway, 2016, p. 5). This definition indicated that the cyborg is an ambivalent entity that is not only real but also fictitious. Firstly, the cyborg is a real thing in that we have created our material bases on cybernetics and information theory, which included mechanization of our own body and social system. Our lives are conditioned by miniaturization, microelectronics, and mechanical, industrial circuits with an advanced electronic industry. Communication sciences translate our being into programmable codes, and biotechnology writes and rewrites that code. She succinctly summarized these conditions as "Biological organisms have become biotic systems, communications devices like others. There is no fundamental, ontological separation in our formal knowledge of machine and organism, of technical and organic" (Haraway, 2016, p. 60). Secondly, the cyborg is fiction, a metaphorical assemblage of the anti-humanistic ideal of Haraway. She mentioned that "A cyborg body is not innocent; it was not born in a garden; it does not seek unitary identity and so generate antagonistic dualisms without end (or until the world ends); it takes irony for granted. One is too few, and two is only one possibility" (Haraway, 2016, p. 65). Jeffrey (2016) saw this strategy of using cyborg as a metaphorical tool as her "deliberate irony" (p. 25), picking up cyborg, the product of modern industrial-military complex, as a symbol of fighting against the holistic and pure human ideal of humanism.

Haraway herself was once a part of early transhumanists, a cybernetic group. In the interview with Gane (2006), Haraway revealed that her intellectual trace is similarly hybrid like a cyborg, trained in biology and laboratory experiment and then moving into critical and feminist movement studying science history. Given this dual background, there is sympathy toward her former colleagues and intellectual home in the lab on the one side of her mind. She confessed that she enjoyed meeting organisms in her early career in the laboratory as objects of knowledge. She said, "For me, it was always about the materialities of instrumentation of organisms and laboratories, [I was] really interested in the various non-humans on the scene. *The Cyborg Manifesto* came out of all that" (Gane, 2006, p. 136). Haraway  said that she was also in the core group of cybernetic and biological study at her time under the advisement of Evelyn Hutchinson, one of the prominent biology scholars who adopted cybernetic theory for the biological descriptions.

Haraway's background and experience of once being a part of the cybernetic group in biology may have influenced her not to fear  the rapid technological advance but rather recognize it as a condition we live with. She said, "This is about those objects that we non-optionally are. Our systems are probabilistic information entities. It is not that this is the only thing that we or anyone else is. It is not an exhaustive description but it is a non-optional constitution of objects of knowledge in operation" (Gane, 2006, p. 139). Also, she mentioned that "I'm sympathetic to certain kinds of cybernetic efforts to think through autopoiesis" (Gane, 2006, p. 139), a little bit distancing herself from the critical theorist against transhumanism. She added that even though such a material condition has many problems, "we had better inhabit as more than a victim. We had better get it that domination is not the only thing going on here. We had better get it that this is a zone where we had better be the movers and the shakers, or we will be just victims" (Gane, 2006, p. 139). She said that technology is not a thing to fear but rather a thing to play and live with, explaining, "The machine is us, our processes, an aspect of our embodiment. We can be responsible for machines; they do not dominate or threaten us. We are responsible for boundaries; we are they" (Haraway, 2016, p. 65).

What Haraway was most outrightly against was the dominance of any radical single vision and narratives. She criticized radical transhumanistic visions such as mind uploading to machines suggested by Morevec, blaming such insistence as ahistorical and short-sighted vision. Primarily, she emphasized that the mind without embodiment is an impossibility, saying, "I don’t care what you are talking about, but if you think that virtualism is immaterial, I don’t know what planet you are living on!"(Gane, 2006, p. 148). She also said the situation wherein everything was turned into codable information, which is necessarily inseparable from the global capital and commodification of life, is not a desirable condition she imagined; instead, it was close to a nightmare. Besides, Haraway also criticized classic humanists for their human narcissism. Inspired by Derrida's essay on three wounds to human narcissism, constituted by Copernican, Darwinian and Freudian wounds, Haraway suggested the emergence of lively machines as the fourth wound of human-centric narcissism. By recognizing and adding lively machines near us to  the list of “the wounds of human narcissism,” Haraway emphasized our relationality with even non-organic existence. She also tried to provoke humanistic thinkers by reminding them of intelligent machines having a creative ability. On the same ground, she also criticized critical philosophers of technology, saying, "it is crazy to be stuck in that relentless complaint about technology and techno-culture and not getting the extraordinary liveliness that is also about us" (Gane, 2006, p. 142).

Haraway dreamed of a fundamental ontological and epistemological turn, opposing the classic humanists’ ideal of pure soul and quality difference between humans and other beings. Transhumanists and cyberneticians insisted earlier that posthumanists also accept that subjectivity or intelligence can emerge from inorganic beings. Hayles

(2010) captured this point saying "…subjectivity is emergent rather than given…" (p. 291). This point indicates that subjectivity can be emergent in any animate or inanimate beings in their embodiments, as she recognized the intelligent quality of connectionists' deep neural circuit. Similarly, Braidotti (2013) realized that Guittari's recent perspective about the self-organizing matters already recognized qualitative contingency between organic and inorganic matters, thus defining machines as intelligent and generative beings. The posthumanists and Haraway emphasize that recognizing a quality of machinic intelligence is to abandon a human-centric worldview and redefine ourselves as a more relational being. Hayles (2010) also stressed that in the posthumanist world view, "partnership between humans and intelligent machines replaces the liberal humanist subject's manifest destiny to dominate and control nature" (p. 288). She argued that our question is not whether we will become transhuman or not, but instead how we will incorporate intelligent machines.

In that regard, posthumanists pursue coexistence with other intelligence agencies on the future planet. Hayles (1999) wrapped up her famous and controversial book of *How we Became a Posthuman*, saying: "Although some current versions of the posthuman point toward the antihuman (transhumanism) and the apocalyptic (anti-transhumanism), we can craft others that will be conducive to the long-range survival of humans and of the other life-forms, biological and artificial, with whom we share the planet and ourselves" (p. 291). Braidotti (2013, p. 60) also mentioned that posthumanists' relational boundary was almost limitlessly expanded beyond egoistic human self as such a transhumanists' mechanistic and materialistic vision on consciousness convincingly claimed universality of subjectivity codable on the infinitely expandable circuit of information. In this way, the posthumanists' world view on the intelligent being expands beyond transhumanists' narrow and technical definition. In posthumanism, any intelligent entity is already ethically treated as an autonomous being, not the object of intervention or control.

**Exploring a Cyborg Space**

As illustrated so far, the concept of the 'cyborg' gives a glimpse of how to interpret human and machine integration in the time of advanced AI technology. Particularly, Haraway succinctly captured new posthuman conditions defined by fast evolving electronic, communication, and computational technology, using a cyborg metaphor that symbolizes human-machine hybridity. Haraway's cyborg is also our imaginary being, based on a belief that humans and machines are the same kind of information processor, being considered comparable and even compatible. The cyborg has been considered a convincing theoretical model explaining modern revolutionary shifts with rising techno-science (Mirowski, 2002; Pickering, 2009). In the social science field, Mirowski (2002) attempted to frame modern economics as a cyborg science closely interacting with cybernetics and AI development. Pickering (2009) claimed that cyborg science, an inheritance of wartime regime in the 20th century, was instrumental in creating a new paradigm of social science studies. However, in education studies, an attempt to frame education as a cyborg science has been almost missing, although its knowledge is deeply rooted in the cybernetics-related mind sciences such as psychology, cognitive science, and neuroscience. Thus, this research will explain how the hybridity and convergence of human and machine intelligence and the learning process has emerged in the intellectual history of cross-reference and the cross-metaphor of human and machine mechanisms. The next two literature review chapters will then introduce the intellectual history of education science and artificial intelligence wherein the cyborg, hybridity of human and machine, has been nurtured and become a dominant paradigm framing human and machine intelligence.

# CHAPTER3. LITERATURE REVIEW

## Intellectual History Of Education Science

Education study is multidisciplinary in nature, combining and interacting with various other fields of study such as philosophy, history, politics, sociology, psychology, statistics, anthropology, and economics (Lagemann, 2000; Mehta, 2013). At the same time, the multidisciplinary nature of the education study has made it vulnerable to outer influences. Lagemann (2000) pointed out the vulnerability of education study, saying that "…this domain of scholarly work has always been regarded as something of a stepchild, reluctantly tolerated at the margins of academe …" (p. x). In turn, as can be seen from the report commissioned by the National Research Council (Bransford et al., 2000; Shavelson & Towne, 2002), which declared education study as education science, such a vulnerability made education scholars and communities pursue its intellectual trustworthiness by borrowing scientific features from other sciences. Amongst the many, the so-called mind sciences (such as psychology, cognitive science, and neuroscience), which are characterized by rigorous experimental data collection methods, quantification, and objectification that mostly borrowed methodological orthodoxy from hard sciences, have made the most remarkable impact on the tradition of education study shaping its paradigmatic base (Glaser, 1984; 1991; Mayer, 1992; Bransford et al., 2000). On that ground, this chapter aims to illuminate the history of education studies from the perspective of broader intellectual currents which have shaped mind sciences in general. In particular, this paper focuses on how machines, with the advent of advanced science and technology, made people understand that the human mind is a machine, thus reshaping a mechanical view of the human mind that has ruled mind sciences for the last several centuries.

*Defining the mechanical perspective*

The meaning of a mechanical perspective is manifold, but the mechanical perspective will be defined in two ways in this chapter. First and foremost, there is an explicit mechanical perspective that refers to machines to explain non-mechanic things such as organisms and natural phenomena. Such an analogy entails the presumption that machine and non-machinic beings are fundamentally the same, thus being comparable and even compatible with each other. Garber (2002, p. 185) found that the mechanical philosophers around the 18th century believed that "the whole world can be treated as if it were a collection of machines." Robert Boyle, one of the founders of mechanical philosophy, depicted mechanical philosophy as an attempt to explain the natural phenomenon in reference to Strasbourg's clock (Van Lunteren, 2016, p.767). Berryman (2003) said that the mechanistic perspective considers machines a perfect guide in examining natural phenomena but not vice versa. Boden (2006, p. 58) classified Descartes as a perfect example of a mechanical thinker on the ground that "he often drew explicit analogies between living creatures and man-made machines, seeing these as different in their complexity rather than their fundamental nature."

Sometimes a mechanical perspective can be implicit when there is no explicit reference to machines. This includes any attempt to explain the natural phenomenon with mechanistic features such as functional, physical, sensible, and material features. Garber (2002, p. 185) said "According to the mechanical philosophy, everything in nature is to be explained in terms of the size, shape, and motion of the small parts that make up a sensible body." According to Van Lunteren (2016, p. 767), Robert Boyle also defined mechanical philosophy "as the attempt to explain all natural phenomena in terms of those "two grand and most catholick principles of bodies, matter, and motion." Osler (2001) defined mechanical philosophy as thinking that "sought to reduce all causality to the contact and impact between particles and matter" (p. 154). This means that a mechanical perspective is an attempt to borrow the logic of natural sciences such as physics and chemistry to explain organismic and metaphysical entities. Boden (2006) defined a mechanical view on the human mind as a belief that "the same type of scientific theory could explain processes in both minds and mind like artifacts" (p. 168). In a similar vein, one of the reasons Boden (2006) saw Descartes as a mechanical thinker was that "he believed that the principles of physics could explain all the properties of material things, including living bodies" (p. 58).

In addition to the explicit and implicit categorization, Berryman (2003) agreed that there can be an exclusive and inclusive categorization of the mechanical view depending on its level of tolerance for the other ways of explaining the natural phenomenon. Exclusive mechanical perspective enshrines a mechanical perspective as sole way of gaining real knowledge. Berryman (2003) insisted that all mechanical perspective is exclusive as it does not allow any room for other thinking such as vitalism or holism.  However, such a clear-cut distinction does not apply to every case. For instance, dualistic views are using a mechanical perspective in explaining only a part of the natural phenomenon (Hatfield, 1995). Therefore, this paper will recognize an inclusive mechanical perspective as one of the different options for identifying mechanical thinkers' positions in history, what I term a quasi-mechanical view.

*A Machine As A Metaphor*

As the machine has evolved, our social norms and epistemic bases have changed. A machine represented our modern mode of logical thinking, while it became a robust conceptual and linguistic tool to deepen our self-knowledge, thus perpetuating the co-production of technology and social norms (Jasanoff, 2004). Using the most successful and sophisticated machines and artifacts as a reference point to explain natural phenomena, we defined and redefined the world and ourselves (Van Lunteren, 2016, Jasanoff, 2004). The latest development of such recognition position our mind as a machine, which can be called a mechanical view of the human mind.

The mechanical view on the human mind did not come in a day or a week. There was a long historical evolution behind the scene, which incrementally formulated our paradigmatic and epistemological base. This paper assumes that this change was incremental in the following two ways. Firstly, the object of mechanical reasoning was expanded from nature to the human body and body to mind. Secondly, with the advent of more sophisticated and authentic machines, machines' initial analogical expression was gradually replaced by the spread of thinking that a mind is a real machine. In other words, as the machine perspective advanced, our mechanical perspective on the human mind evolved from quasi-mechanical to the explicit and exclusive mechanical view, making it impossible to comprehend the human mind without mechanical logic.

*Nature As A Machine*

After the Renaissance, there was a growing tendency to understand the natural phenomenon as a mechanical process (Merchant, 1980). It was signaled by Henry Power (1623-1668): “I think it is no rhetorician to say that all things are artificial: for nature itself is nothing else but the art of God” (cited in Cook, 2001, p. 133). This statement exemplifies how people started depicting the god as scientists, engineers, or skilled craftsmen creating complicated machines as technology advanced (Bullock, 2008; Cook, 2001). Banfield (2011) assumed that most of the modern hard sciences, such as physical, biological, and social sciences, share some “founding metaphors” (p. 105). He found that the founding metaphors were shifted from animal spirits to electrical currents. In other words, this transition in metaphor from spirit to machine indicated there was a growing tendency to understand nature as materials having a purposive function, not a mystic and metaphysical force. In terms of the machine used to describe the natural phenomenon, its symbolic representation has changed as technology advanced. Van Lunteren (2016) found that the dominant metaphorical symbol of the machine used in depicting natural phenomenon has been changed from clock, balance, steam engine and finally to computer. Besides, Cook (2001) found that Robert Boyle understood the universe as “the operations of a vast machine that ultimately depended on the external agency of the divine artificer who created it” (p. 140). Charles Babbage (1791-1871) tried to prove that geological transformation is a mechanical process moved by a machine developing a calculous machine that could simulate nature-like irregularity out of the systemic mechanical structure (Bullock, 2008).

*A Human Body As A Machine*

As a corollary of nature's mechanization, people started to understand the human body as a machine. Descartes was one of the earliest pioneers who explicitly expressed an understanding of animals and a human body as a machine - materially constructed functionaries - and established the basic methodological frame of an experiment for physiology (Black, 2014; Boden, 2006; Gardner, 1985; Hatfield, 1995). Descartes understood that humans' bodily functions follow physical law; mostly, he described the heart as a heat engine and explained its mechanism based on pure physical law (Boden, 2006). This mechanical understanding of the human body let Descartes envisage bodily organs as machines (Black, 2014). Descartes' idea of the mechanical view on the human body remained a lingering impact on the Western intellectuals. Salisbury (2011) found that renowned scientists of the early 20th century such as Thomas Edison (1847 - 1931) and Hermann von Helmholtz (1821-1894) understood that a human body was largely based on materialistic and mechanistic condition. Most recently, Campenot (2016) even insisted that without electric power in our body, we would be just "less than vegetables" (p.1).

On the other hand, Descartes' version of a human's mechanical view was only a partial endeavor as he left human mind as a territory of a holistic and mystic soul, autonomously controlling the mechanical human body (Black, 2014; Boden, 2006). His description of the human mind was still ambiguous. Boden (2006) found that Descartes acknowledged the mind as having an independent and autonomous free will apart from the mechanical body, while he also recognized that body and mind are physically connected. That is, on the one hand, Descartes continued differentiating the holistic mind from the mechanical body. However, on the other hand, Descartes also sensed that the mind and the physical brain are correlated. In this way, Descartes obscured the line between the materialistic and immaterialistic view of the human mind, putting uncertainty into its study. Hatfield (1995) considered that the Cartesian dualism of mind and body begot contradictory results. Firstly, the differentiation of mind and body justified establishing separate disciplinary study specific to the human mind, psychology. Secondly, but ironically, a Cartesian origin of psychology was made vulnerable to the intervention of physics and quantification of the human mind as Descartes also invented a physical approach to the human mind through physiological experimental study (Hatfield, 1995, p. 194).

*A Human Mind As A Machine*

It was not until there was a significant progress in the human brain's anatomy that the human mind was depicted as a mystic soul given by God. Salisbury (2011) suggested how the advanced anatomical skills, so-called technologies of seeing inside, changed our view of our brain and mind. For this, Salisbury examined the case of Broca's brain anatomy and penetration. Paul Broca (1824-1880) implemented an autopsy of one brain whose language competency was impaired and found that the brain damage at the brain's specific location was correlated to the dead person’s incompetency. Based on this discovery, Broca mapped out a brain structure with the conviction that a specific location in the brain may imply structural bases of certain functions of human minds. Salisbury (2011) considered that this localized brain function theory became the mechanical model of language production and mind. He insisted that people started using the metaphor of a machine to describe brain functions as the technology opening up our skull to see inside of it advanced. For example, Wernicke described a damaged brain causing language problem as ‘the malfunctioning telegraph’ (Salisburg, 2011), and Edison admired the sophistication of the human brain by describing the brain’s subsets as a phonograph cylinder, one of his inventions (Salisburg, 2011). Besides brain anatomy, Boden (2006) and Husbands et al. (2008) found that at the end of the 19th-century, neuropsychologists influenced by Descartes’ mechanical view uncovered the mechanistic base of the nerve system in the human brain. One of the strange figures in this field, Alfred Smee, insisted that he could explain all types of human thought and behavior based on an electrical machine in the nerve system (Boden, 2006). Danziger (1990) saw that the advance of brain anatomy and physiology brought a view that brain function binds to structure; thus, the function of the human mind could only be defined through structural identification.

Since the early 20th century in Western countries, more radical thought has emerged. At that time, some of the radical thinkers proactively conceptualized the human mind as a real machine, diverging from the metaphorical representation of the human mind in its previous time. Black (2014) concluded that thinking about the human mind as a machine entails the assumption  that “(minds) are interchangeable with other machines and whose functioning can be understood by looking at machines rather than our bodies themselves”(p. 9). He called this epistemological stance “backward causation”(p. 9). Boden (2006) explained how radical it was to bring an idea of the mind as a real machine in the early 20th century by saying, “Maybe minds are machines, too! As late as 1930, this shock-horror thought had not even attained the status of a heresy. It was not a heresy, because no one believed it. Indeed, no one had even suggested it" (p. 168).

These thinkers were from the various new fields of engineering, cybernetics, computation, and psychology. Admittedly, the development of computer and computer-related information technology during the Second World War period through cybernetics and computer engineering made a significant contribution to mainstreaming this idea (Gardner, 1985). A father of modern computing and artificial intelligence, Alan Turing (1912-1954) was one of the early contributors of this paradigm shift, envisaging the human mind as a real machine (Husbands et al., 2008). Hodges (2008) found that his wartime experience of code-breaking of German's communication system and following the success of mathematical development of computation made him incrementally convinced that minds are machines. McCulloch (1889-1969), as a psychologist,  said, "Everything we learn of organisms leads us to conclude not merely that they are analogous to machines but that they are machines''(Boden, 2006, p. 182). Kenneth Craik (1914 - 1945) was also the pioneer of such a paradigm shift, who is often considered a founding father of cognitive psychology and cybernetics (Husbands et al., 2008). He raised the seemingly radical but critically influential idea that the human brain is one of the machine-kinds, not the other, based on his human nervous system's findings. Following him, Herbert Simon, one of the founding fathers of cognitive science and artificial intelligence, insisted that the human brain and computer are both adaptive information processors (Kline, 2011).

*Mechanical View In Mind And Education Sciences*

As mentioned earlier, there is a view that the intellectual history of education study can be defined as a rationalization process pursuing more rigorous scientific knowledge and practice borrowed from mind sciences. In a sense, this rationalization process made the human mind's mechanical view more pervasive in education studies and mainstreamed such a view as a dominant epistemic base. In other words, a growing number of scholars in education studies started to explicitly depict the human mind as a machine, having specific quantifiable and material features. Some scholars were immersed into the mechanical view, accepting that the human mind can be understood with its materiality and functionaries in the same way that we measure objects in natural sciences such as physics, chemistry, and biology. Given this account, this section will present how the mind sciences accepted and cultivated the mechanical view of the human mind and then transferred this view to education studies.

**Charles Darwin and Wilhelm Wundt.** America's early modern psychology was nurtured and inspired by Charles Darwin's evolutionary theory (Boakes, 1984; Green, 2009; Greenwood, 2008) and Wilhelm Wundt's physiological psychology (Boakes, 1984; Danziger, 1990). It brought uncertainty to American psychology as Darwin was a naturalist believing the vital and active force of organisms, while Wundt was more geared toward fixating organisms under experimental mechanical conditions. This section will describe how this contradictory dream of naturalism and a mechanical worldview co-created early American psychology.

Firstly, Darwin's evolutionary theory revolutionized the Western intellectual community by breaking the boundary between animal and human. Darwin's naturalism promoted the view that humans and animals have a strong continuity, not fundamentally differentiated but distinctive to a degree (Boakes, 1984; Greenwood, 2008). Both functionalism and behaviorism shared the common paradigmatic ground that animal psychology shares a significant commonality with human psychology (Greenwood, 2008). More than anyone else, Darwin influenced American psychologists to the extent that John Dewey and Edward Lee Thorndike commonly believed that the human condition could be improved by psychological intervention and enhancement (Greenwood, 2008).

Secondly, the German psychologist Wilhelm Wundt (1832-1920) frequently appears as a critical figure who influenced American psychology's foundational figures in psychology history books. For example, Danziger (1990) marked the establishment of Wundt's psychology laboratory at the University of Leipzig as modern psychology's birth date. Influenced by his research advisor Helmholtz who had a rather rigorous materialistic view of the human mind, understanding the human mind as a particular state created by chemical-physical forces (Boakes, 1984), Wundt's research was also driven by the mechanical understanding of the human mind. According to Danziger (1990), Wundt believed psychology was a supplemental form of physics studying psychological causality. His laboratory was run almost the same as chemistry and physics labs in any German university, thus serving as the birthplace of experimental psychology. Wundt intended to apply the experimental method developed in physiology to psychology and psychological issues in speculative philosophies (Boakes, 1984). Such an innovative and pioneering effort made him one of the most influential and prominent figures for the fledgling academic society of American psychology.

**William James: Mind Is Not A Soul**. In the mid 19th century in the US, new generational scholars in philosophy and other natural sciences were influenced by European thinkers such as Charles Darwin, Robert Spencer, and Wilhelm Wundt. William James (1842-1910) was usually considered a founder of American psychology; one of his students was Edward Lee Thorndike. Although James turned away from his initial enthusiasm for experimental psychology in his later career, revisiting metaphysical explanation of the human mind and soul, it is undoubtful that he advocated for the mechanical view of the human mind in his powerful writings, which influenced his students and readers significantly (Boakes, 1984; Evans, 1990; Greenwood, 2008). In other words, James was the first American psychologist who conceptualized the mind not as a soul, but a natural process, establishing an image of "naturalistic and secular representation of mind" (Evans, 1990, p. 443). James' secularization distinctly contrasted with the University of Cambridge's conservatism, which refused to acknowledge psychology as a natural science, denouncing it as a demeaning effort against religion (Greenwood, 2008). Also, inspired by Darwin's evolutionary theory and Wundt's physiology, James tried to explain human consciousness as evolutionarily purposive, and physiologically structured like a machine (Green, 2009). James emphasized the human brain's material and functional bases and its cerebral mechanism as a source of scientific knowledge of psychology. James (1892) argued that the physical state of the brain represents the state of mind

> We do not know precisely what a nerve current is, it is true, but we know a good deal about it. We know that it follows a path, for instance, and consumes a fraction of a second of time in doing so. We know that physically considered, our

brain is only a mass of such paths, which incoming currents must somehow make their way through before they run out. We even know something about the consciousness with which particular paths are especially 'correlated,' those in the occipital lobes, e.g., being connected with the consciousness of visible things. Now the provisional value of such knowledge as this, however inexact it be, is still immense. It sketches an entire programme of investigation and defines already one great kind of law which will be ascertained (p. 152).

**Functionalism And John Dewey: The Mind Is Not A Machine.** Functionalism emerged around the 1890s in the US psychological community as a move toward revamping Wundt's physiological psychology. The functionalists emphasized the adaptation of the human mind to the changing socio-cultural environment while refusing to recognize the impact of physical development at the physiological level (Green, 2009). Functionalists saw that functions are not necessarily bound to their physical structure as they are non-reducibly whole, but independently evolve and adapt to the given environment (Green, 2009).

One of the representative figures of functionalist was John Dewey. Without a doubt, John Dewey has been recognized as one of the most influential scholars of education studies. Dewey's ideas have a lingering effect on the educational discourse and practice broadly in the US. Lagemann (2000) saw that the intellectual legacy of Dewey is still alive, inspiring scholars and practitioners to seek alternative ways of education, despite his earlier defeat to Thorndike. Mehta (2013) insisted that in order to revamp the entire education system of the US, it is necessary to redesign schools as a "center of inquiry" according to the spirit and ideas of Dewey. Bereiter (2002) said that Dewey's educational philosophy revives whenever there is a futuristic reform discussion in education even this far after his life. As an educationist, Dewey is defined as a progressivist believing in human progress through education, and school as a powerhouse of social reform. He was a hegemonic rival of Thorndike with his situated, developmental, and less-technocratic philosophy (Lagemann, 2000). Unlike behaviorists, Dewey emphasized that humans can actively adapt to their ecological or social environment through continuous learning by experiencing (Popp, 2007). He also insisted that education should be designed not for the sake of other vague beliefs or assumptions on the human mind, but for the alignment with the intrinsic developmental pathway as preconfigured in students' minds.

John Dewey took a distinctively different pathway in his intellectual pursuit in both fields of education and psychology, compared to his contemporary scholars such as behaviorists (Lagemann, 2000). Thus, we can expect that he may be an exception from the mechanization trend of the human mind, advocating ideas for imagining humans differently. Such an expectation is half met. Firstly, John Dewey was very explicit in opposing the mechanization of the human mind. In his book titled *How We Think*, Dewey (1933) expressed his anti-mechanical sentiment describing human thinking as something that is not a machine:

> Thinking is specific, not a machinelike, ready-made apparatus to be turned indifferently and at will upon all subjects, as a lantern throws its light as may happen upon horses, streets, gardens, trees, or river…Thinking is not like a sausage machine that reduces all materials indifferently to one stereotyped,

marketable commodity, but is the power of following up and linking together the specific suggestions that specific things arouse. (p. 46)

According to Lagemann (2000), Dewey was also explicitly against an attempt to define education as natural or hard science such as physics and chemistry. Bredo (1998) also insisted that Dewey refused to use metaphors of mechanical logic and warned against the reductionistic physiological process. Instead, he primarily considered education science as social scientists studying "the conditions which secure intellectual and moral progress and power" (Lagemann, 2000, p. 50). He also wanted to establish a comprehensive naturalistic approach to education by borrowing conceptions and notions from biology and history (Bredo, 1998). In the same vein, he designed his laboratory school to counter the dominant movement of experimental psychology, targeting the nurturing of a democratic mind with adaptive social and vocational skills.

On the other hand, Dewey's psychological and philosophical presumptions for child education had the potential to be interpreted as a mechanical perspective in a broader sense. Admittedly, the machines of his time were not more than "sausage machine" and "lantern," "indifferently" repeating simple motions regardless of any change in the given environment. Thus, it is understandable that Dewey abhorred any attempt to equate the human mind with the machine. However, in the computer age, adaptation and flexibility are not considered unique characteristics endowed to humans or other organisms. It is mainly due to the invention of highly sophisticated computer machines, which have been gradually upgraded with highly flexible features to various natural settings. The Information Age understands organism as the information itself, and its morphological formation is reduced to the coding of information; for organisms, it is reducible to the coding of DNA. Popp (2007) was the one who most actively recognized the possibility to read Dewey mechanically. As he indicated, Dewey's philosophy of the human mind is easily read and compatible with the mechanical philosophy of the cognitive neuroscientists. He argued that Dewey's explanation of human thinking presaged a highly sophisticated information processing machine:

> Dewey constructs no evolutionary explanations of how our parallel processing brains can manipulate highly symbolic material, but he does recognize that the mind did emerge by way of natural selection and that the philosopher's task is to help improve our cognitive architecture…If we again think of the mind as a virtual machine built of a set of rules for processing information, a good deal of the computing power of the mind is devoted to trying to use good sense in practical affairs. Dewey is trying to get us to edit or upgrade the rules we are using to manage our affairs. In other words, Dewey can be interpreted as showing us how to improve the rules that constitute our virtual-computing minds. (pp. 88-89)

Also, Herbert Spencer's influence on Dewey indicates that the mechanical perspective is indirectly connected to Dewey's theory. Developmentalism of the child's mind, which advocated timely accurate instruction and teaching for the child according to their natural process of mental development, was a product of mechanists like Erasmus Darwin and Herbert Spencer inspired by phrenology and physiology (Tomlinson, 1996). The developmentalism of a child's mind laid a foundation of the logic of Dewey's educational idea (Bredo, 1998; Popp, 2007).

In the aggregate, John Dewey and his functionalist approach was against the mainstream mechanistic experimental psychology of his time, but this does not necessarily mean that Dewey sidelined himself from the significant historical currents of mechanization of the human mind. That is, he was the one whose understanding of the human mind was one step ahead of his contemporary scholars because of his insistence on a highly flexible and adaptable feature of the human mind, but his description of the human mind was in a sense mechanistic enough to be easily read and understood by the mechanists after his death. It means, Dewey's conceptualization of the human mind resembled the machine of advanced information processing, which was yet to be invented at his time. In that account, Dewey can be considered as an unusual figure of his time. However, he is still not a drastic game-changer in the longer term perspective. He shared the basic assumption on the Neo-Darwinian human mind; the mind is a more advanced, adaptable, and resilient information processing machine that accepts knowledge following specific orders and patterns as it was designed.

**Behaviorism and Thorndike: Mind Is A Machine.** Psychological research's behaviorist tradition assumes that mind or consciousness is not a proper element to study because they are elusive and metaphysical. Instead, the early behaviorists believed that only observable human behavior could be the object of psychology's scientific method. Levin (1987) saw that behaviorism symbolized the American spirit in the early 20th century representing a practical and optimistic atmosphere. Levin quoted Watson's unwittingly but adamantly positive outlook on the use of behavioristic approach to the study of the human mind: "Give me a dozen healthy infants, well-formed, and my specified world to bring them up in and I'll guarantee to take any one at random and train him to become any type of specialist I might select - doctor, lawyer, artist, merchant chief, regardless of his talents, penchants, tendencies, abilities, vocations, and race of his ancestors"(p. 1683). Also, Levin (1987) pointed out that behaviorism was established when psychology was separated from its former affiliation to philosophy. This inevitably pushed psychologists to distance themselves from the metaphysical philosophy and made them reject introspection as a method of inquiry while acknowledging observable motor behavior. Such an approach was dominant in psychology until it was challenged by cognitivism in the 1950s (Lagemann, 1989).

Thorndike was one of the most frequently mentioned figures in the influence of psychology on education in the early modern history of American education (Lagemann, 2000; Mehta, 2013). At the same time, however, he was also one of the most stigmatized and haunted names in history such that "He faced the charge that his psychology was mechanistic and explained adequately only the most rote kinds of learning" (Glaser, 1984). As a student of William James, primarily inspired by his early mechanical works in psychology, Thorndike was an ardent supporter of using measurable, quantifiable, and physical methods through a psychological experiment (Lagemann, 2000). Thorndike came from animal psychology and physiology, focused on the stimulus-response pattern of human behavior. He was also the one who very explicitly recognized the human mind as a machine. He stated: "The mind is, on the contrary, on its dynamic side a machine for making particular reactions to particular situations. It works in great detail, adapting itself to the special data of which it has had experience" (Thorndike & Woodworth, 1901, pp. 249-250).

Like his contemporary scholars, Thorndike was well aware of physiology and actively used it to explain human behavior. Thorndike and Woodworth (1901) assumed that the high intellectual means having more physiological connection and association. They denied that there is a distinction in quality rather the quantity of the connection that decides the quality of intelligence and behavior: "…the person whose intellect is greater or higher or better than that of another person differs from him in the last analysis in having, not a new sort of physiological process, but simply a larger number of connections of the ordinary sort" (Thorndike & Woodworth, 1901, p. 415). Thorndike's quantification of the human mind was closely related to his mechanical understanding of the human mind, which conceptualized the mind as a physically embedded function with features of motion, size, and structure. Thorndike & Woodworth (1901) said:

> Intellect might be exactly proportionate to the activity of the thyroid gland, or to the proportion of the brain weight to body weight, or to the number of associative neurons in the frontal lobes or to the intensity of a certain chemical process, and hence be measurable by observations of the thyroid's action, or estimates of the brain's volume, or by a count or measurement of neurons, or by a chemical analysis. (p.12)

But, Thorndike could not use this sophisticated physiological measurement during his life as there was a limit in anatomical technology. Instead, he chose human behavior as a proxy to measure physiological states, which in turn represented a state of mind (Thorndike & Woodworth, 1901).

**Cognitivism: The Mind Is A Computer.** The mechanical view of the human mind made a breakthrough after a long impasse of behaviorists' simple rote mechanistic vision. It was only possible as the new generational thinkers conceptualized the human mind as a computer, a highly versatile and sophisticated information processing machine (Abrahamsen & Bechtel, 2012; Boden, 2006; Crane, 2003; Gardner, 1985; Thagard, 2005). As early as 1949, Warren McCulloch was already quite straightforward, insisting that a human mind is a computer machine and saying, “Man's brain is much the most complicated of computing machines, and it requires power to keep its relays in the operating range of voltage…The brain is a logical machine. Each of some ten billion relays has only two states: pulse or no pulse” (p. 492). Ross Ashby (Ashby, 1951, p. 1, cited from Asaro, 2008), one of the founding members of the British cybernetics group, described this transformation in paradigm:

> It has become apparent that when we used to doubt whether the brain could be a machine, our doubts were due chiefly to the fact that by “machine" we understood some mechanism of very simple type. Familiar with the bicycle and the typewriter, we were in great danger of taking them as the type of all machines. The last decade, however, has corrected this error. It has taught us how restricted our outlook used to be; for it developed mechanisms that far transcended the utmost that had been thought possible, and taught us that “mechanism" was still far from exhausted in its possibilities. Today we know only that the possibilities extend beyond our farthest vision. (p. 149)

From the 1940s to the 1960s, some changes were brewing under the mantle, initially dominated by behaviorists’ experimental psychology that had thrived for a while to create a new intellectual movement of cognitive science. Firstly, the wartime research efforts to create automated intelligent war machines begot a new understanding of the human mind and its mechanism (Abrahamsen & Bechtel, 2012; Gardner, 1985; Lagemann, 2000). It laid the foundation for the mechanistic interpretation of the human mind with advanced modeling of human thinking as a physical, materialistic, and mathematical entity. The Macy Conferences, the first one held in 1946 until the tenth one in 1953, set the foundations of the modern computing and cognitive science (Dupuy, 2010a). The participants were key figures of their fields of studies. Claude E. Shannon, a founder of information theory, established a mathematical and logical model of information processing (Dupuy, 2010a, Miller, 2003). In neuroanatomy and neurophysiology, Pitts and McCulloch attempted to establish a psychological model entirely through mechanical explanation combining functional and structural causal mechanisms of the human mind and brain (Dupuy, 2010a). They finally proved that a network like physical architecture could simulate logical function, that is, a Turing machine (Abrahamsen & Bechtel, 2012). Norbert Wiener developed an automated feedback machine inspired by the human nervous system's homeostatic system called cybernetics (Gardner, 1985). Cyberneticians led by Wiener assumed that “thinking is a form of computation” and “physical laws can explain why and how nature…appears to us to contain meaning, finality, directionality, and intentionality” (Dupuy, 2010a, pp. 3-4). John Von Neumann, working closely with cyberneticians and information theorists, developed a modern computer concept completing the functional and structural base of the information processing machine. In particular, by claiming that “anything that can be exhaustively and unambiguously described” can be simulated by computer, he declared that whatever we describe as a mechanism or function of the human mind can be simulated and subsumed by computer (Piccinini, 2018). As a consequence of the diffusion of mechanical thinking, people started believing that the human brain should be computable and even the human mind is one of the computers - which is called pancomputationalism (Piccinini, 2018). This idea substantially impacted cognitive science's critical scholars later on, such as Jerome Bruner and George Miller (Boden, 2006).

Secondly, it became clear that behaviorists' endeavor explaining the human mind only with observable behavior did not show any significant progress in handling more sophisticated theories in linguistics, computer science, and neurosciences (Gardner, 1985; Lagemann, 2000). The experimental method limiting objects only in stimulus and response did not show any sign of approximation to the higher level of human cognitive behaviors especially language competence and symbolic representation skills. Miller (2003) remembered how his close friend Noam Chomsky criticized the behaviorism as an erroneous orthodoxy: "As Chomsky remarked, defining psychology as the science of behavior was like defining physics as the science of meter reading" (Miller, 2003, p. 142). Miller (2003) said that he, Chomsky, and Bruner wanted to revive interest in the human mind and cognition, opening up the black-box encapsulated by behaviorists who prevented any attempt to touch that box. In regard to this move, Abrahamsen and Bechtel (2012) stated, "Both metaphors (computer and communication metaphors) gave rise to information processing by offering engineering—based ways to open the "black box" between stimuli and responses and model mental activity" (p. 13). When Miller and his colleagues looked around to find an alternative to behaviorism, rising stars were innovating and restructuring the field of mind sciences, such as Claude Shannon, Norbert Wiener, Von Jon Neumann, Alan Newell, Herbert Simon, Marvin Minsky, and John

Macarthy frequently mentioned as founding fathers of modern computing and artificial intelligence (Glaser, 1984; Miller, 2003).

Thirdly, the effort to topple the bulwark of behaviorism and new technological development in computing coincided with mounting pressure to redefine whose perception of a human. As Cold War politics matured in the US, there was a governmental and political move to re-conceptualize humans as rational beings and American citizens as flexible and open-minded people against communists and conservatives (Cohen-Cole, 2014). The idealistic vision of intelligent humans was defined as those who can think like scientists who can autonomously explore, discover, and utilize scientific knowledge. According to Cohen-Cole (2014), cognitive science was instrumental to this political process as a group of new mind scientists reimagined the human mind as highly flexible and rational, contrary to the dominant behaviorists' vision and Dewey's progressive education.

In this context, the emergence of cognitive science is understood as a product of the convergence of various fields studying the nature of the human mind, such as computer science, linguistics, physiology, psychology, neuroscience, and biology, which initially diverged from the philosophy of mind in the late 19th century (Frankish & Ramsey, 2012; Gardner, 1985; Talkhabi & Nouri, 2012). George Miller (2003) saw that the interdisciplinary connection between computer science, linguistics, and psychology was the core building block of cognitive science. It signaled that the science had matured enough to directly tackle the issue of a higher level of human cognition, which had been intentionally black-boxed and replaced with observable human behavior by the behaviorists (Gardner, 1985).

From the 1950s to 1970s, the institutional and knowledge base of cognitive science was established. Whatever the reason behind explaining the foundation of cognitive science during this period, it seems evident that it was based on the human mind's mechanical view defining mind as a computing machine and borrowing concepts and theories from hard sciences. Glaser (1991) succinctly summarized this dominant trend: "The analogies between human cognitive processes and the mechanisms of mechanical and electronic systems, such as servomechanisms and computers, captured attention. This work helped set the stage for the present day modeling of human performance in information processing systems" (p. 130). Most of the early generational scholars of cognitive science pursued establishing the symbolic architecture of the human mind. Abrahamsen and Bechtel (2012) defined symbolic architecture as follows: "Symbolic architectures share a commitment to (1) representations whose elements are symbols and (2) operations on those representations that typically involve moving, copying, deleting, comparing, or replacing symbols. A rule specifies one or more operations (e.g., S —> NP VP)" (p. 16). The representative figures are mostly AI scholars such as Marvin Minsky, Alan Newell, and Herbert Simon, so in this chapter, I will skip the discussion of them and will only focus on cognitive scientists closely related to education science.

George Miller was one of the most prominent scholars who initiated cognitive science as an independent study. He innovated psychological approach by introducing Shannon and later Noam Chomsky's generative grammar (Boden, 2006). Influenced by the mechanical view of the human mind, Miller conceptualized the human mind's information processing as a computational process codifying sound elements as bits or phonemes and proved his thesis using experimental study (Boden, 2006). He argued that human memory is limited to storing information, insisting on a magical number seven as a universal limit of memory capacity. He implicitly assumed that there must be physiological structure limiting our capacity saying, “There seems to be some limitation built into us either by learning or by the design of our nervous systems, a limit that keeps our channel capacities in this general range” (quoted in Boden, 2006, p.289).

With George Miller, Jerome Bruner made a significant contribution to open a new field of study. Lagemann (2000) estimated that the opening of the Center for Cognitive Studies in 1960 was a decisive swing to formulate the field of cognitive science, wherein Miller and Bruner studied human cognition in various dimensions. Bruner did not explicitly pursue computation of the human mind as he was not familiar with computer programming, but just like his contemporaries, Bruner took advantage of using computation as a metaphorical representation in explaining human thinking. Bruner said, “New metaphors were coming into being in those mid-1950s and one of the most compelling was that of computing. . . “(quoted in Gardner, 1985, p. 29). His personal exchange with von Neumann nurtured his implicit acknowledgment of the human mind's computational perspective (Boden, 2006, p. 308). Boden insisted that Bruner was imbued with a computational perspective saying, “When Bruner spoke of an inferential ‘mechanism’ underlying perception, he was thinking computationally if not programmatically” (2008, p. 308). Accordingly, Bruner understood human thinking as necessarily logical (Cohen-Cole, 2014; Gardner, 1985) and depicted the thinking process as information processing, including acquiring, retaining, transforming, storing, and presenting information (Boden, 2006). Later on, Bruner's students increasingly accepted computation modeling, and Marvin Minsky, a prominent scholar of artificial intelligence, referred to Bruner's contribution in his *Steps Toward Artificial Intelligence* (Boden, 2006).

Lastly, it seems noteworthy that D. E. Broadbent, a close friend of Miller (Boden, 2006) and who frequently interacted with Miller and Bruner, conceptualized the flowchart of information processing of organisms borrowing a mechanical concept widely used in computer engineering. With this flowchart in mind, Boden (2006) considered that "the organism was here being presented as an integrated system" (p. 292). Also, Broadbent clearly stated that "nervous systems are networks of the type shown in Fig. 7 [the flowchart], and of no other type" (Broadbent, 1958, p. 304). It was the declaration that the human mind is a machine-kind constituted of an information processing mechanism and system - it was beyond just a metaphorical use of a computer. Instead, it was reverse causation of the human mind from its imitated artifact, a computer. Even though Broadbent denied categorizing himself as a positivist, his description of the mind system looks mechanistic due to his emphasis on the causal relationship between structure and function. Broadbent said, "It may often be preferable to explain a physiological fact by reference to its role in a well-understood psychological function'' (Broadbent, 1958, p. 305). Also, he was sympathetic to computer modeling of the human mind  "The great merit of models which can be implemented on a computer ... is that they avoid many... ambiguities. I would firmly believe that in the long run any adequate account of human beings will have to be capable of computer implementation" (Boden, 2006, p. 295). The computational modeling of the human mind is one of the dominant research approaches in cognitive science. Thagard (2005) claimed that the computational-representational understanding of the mind (CRUM) is one of the most foundational paradigms of cognitive science although there is opposition to this idea. Thagard (2005) posited that the CRUM became so popular in this field because the metaphor of computer explaining and describing human's cognitive mechanism was powerful. The CRUM interprets a human's thinking process as data and algorithmic structure, adding accuracy and simplicity to the cognitive modeling. This metaphor can even be extended to the case of brain study, modeling the brain's neuronal process as parallel computational processing.

The influence of cognitive science on education was immense and deemed to be positive. Lagemann (2000) estimated that the advent of cognitive science gave educational study a boost to become a real science more than anything else before. Lagemann put that "insights gleaned from cognitive science and applied to classroom instruction, combined with greater understanding of the ways in which cultural differences influence classroom exchanges, have opened a new potential for effective schooling" (2000, p.xiii). Glaser (1984) said that educational scholars and practitioners had more connection with the scientific experiment through cognitive science inquiry, which made the boundary between basic and applied science blurred. Glaser (1984) also added that the major discovery of cognitive science related to the mental process of learning required teachers equipped with domain-specific knowledge and the science of the learning process.

Along with cognitive science, the mechanical view understanding the human mind as a computer became much more pervasive in education studies. Miller, Bruner, and Broadbent, conceptualizing the human mind as a computer, were highly influential in the field of education studies. There was even more direct extrapolation of computation theory to human cognitive theory. For instance, Glaser (1984) remembered that the strategy to increase multiple computer processing power conceived by Minsky and Papert in 1974, was inspiring to many education scholars using cognitive science. This idea was innovative as it valued knowledge organizing strategy over the innate quality of thinking. Accordingly, Glaser, Chi, and Lesgold adopted the power strategy theory to the cognitive experiment of humans to prove how expert knowledge is different from that of a novice when it comes to organizing strategy of knowledge (Glaser, 1984).

*Conclusion*

The history presented so far in this chapter demonstrated that the effort to make education a science coincided with the general acceptance of the mechanical view of the human mind. Science and technology developing new advanced machines created an epistemological force to reframe nature and the human mind as a machine. In that process, the education studies were also subsumed under the psychologists' mechanical logic through the work of such scholars as James, Thorndike, and Dewey – although Dewey's ideas do not explicitly reflect the mechanical logic - as well as the cognitive scientists including Bruner and Miller. In particular, the findings of this section imply that despite the widespread view amongst cognitive psychologists and even held by Lagemann that cognitive revolution brought the paradigm shift in mind sciences, it was instead a continuation of expanding the mechanical view of the human mind, a process initiated since the middle of the 19th century.

## Intellectual History Of Artificial Intelligence

Automated machines are transforming our society with advanced cognitive learning capacity. There is a view that this transformation may bring a fundamental shift to humankind comparable in its scope to the previous industrial revolutions in human history (Bostrom, 2014; Harari, 2014; Kurzweil, 2012a; Schwab 2016). Despite this substantial growth of AI technology, there is no consensus on AI's definition as it is such a massive assemblage of various sub-disciplinary studies, and it has evolved into various shapes throughout history (Crane, 2003; Konar, 2000). At best, we can get a rough glimpse of AI definitions by referring to some other scholars. Nilsson (2010) offered the following definition: “Artificial intelligence is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment” (p. 13) Khakurel et al. (2018) said:

> AI can be described as a cluster of technologies and approaches, that is, statistical and symbolic that aim at mimicking human cognitive functions or exhibiting aspects of human intelligence by performing various tasks, mostly preceding analytical, analytical mostly preceding intuitive and intuitive mostly preceding empathetic intelligence. (p. 2)

The most systematic and thorough definition came from Russell and Norvig (2001). They categorized AI into four different approaches. The first approach pursues *machines that can act like a human*. The representative figure is Turing who defined “intelligent behavior as the ability to achieve human-level performance in all cognitive tasks, sufficient to fool an interrogator" (p. 5). This approach did not heed much on internal logic of human thinking but instead focused on observable behavior. The second one is the *thinking humanly approach*, the so-called cognitive modeling approach. Simon and Newell initiated this movement and pursued a general problem solver (GPS) imitating the human cognitive process. The third one is *thinking rationally approach,* and in this tradition, scholars focus on a logical language system that can program AI as a rational agent. Lastly, there is a *rational agent approach*. In contrast to the cognitive modeling approach that imitates human cognitive features, this approach is more goal-oriented, less focused on the internal logic structure. This approach takes whatever mechanism that can be perceived as a rational act at the surface value. It involves a deep learning approach.

Despite the variations of the definitions, it is evident that the field emerged and grew out of the desire to imitate the human mind. This fundamental desire enabled AI studies to maintain interaction with the human mind and behavior studies, such as cognitive science, biology, psychology, and philosophy. It implies that the quest for understanding the nature of human intelligence and building up the general intelligence machine has been inseparably intertwined with and inspiring to each other. Given this assumption, this chapter will explore the history of AI studies from its modern inception with a highlight to the interaction with human intelligence and brain studies.

### *Bio-Inspired Early Intelligent Machines And Computers*

The Second World War was a critical event triggering interdisciplinary gathering of scholars from various fields. In both the US and the UK, the governments put much effort to enhance human physical and cognitive capacity using various technologies such as biology, psychology, engineering, physics, and chemistry (Heims, 1991). This wartime situation created a new intellectual stream framing the human mind as a machine (Heims, 1991; Husbands & Holland, 2008). It is often called a cybernetics movement, the precedence of an AI study (Kline, 2011). The cybernetics movement was a transatlantic phenomenon based on interaction in specific groups gathered in the conferences and the small seminars in the US and UK. The first generation of cybernetics scholars mostly came from psychiatric studies (Pickering, 2009). This gathering aimed to study brain mechanisms by creating an adaptive behavioral machine that can imitate animal or human behavior.

On the US side, the Macy conference was the primary gathering of the cyberneticians (Heims, 1991). In the first conference in 1943, the presentations were mostly about the parallelism between computer machines and the human brain, setting the following conferences' direction. Warren McCulloch was interested in neurobiological, physiological, and engineering discoveries. John Von Neumann proved that metal tubes in the computer could function like neurons in the human brain. Lorente de No presented that the neuronal electrochemical impulse works with a computational binary zero and one signal system, proving that the human brain is a kind of computing machine. Norbert Wiener presented his idea of developing an autonomous machine with sensory motors and a feedback loop. Further, he suggested framing biology, statistics, psychology, and social science altogether as a common discipline handling an issue of communication and information processing. Heims (1991) esteemed this first conference as a watershed of the western intellectual history, putting that “Characteristically, the new concepts spanned the human and the inanimate, leading to mechanical metaphors for human characteristics and anthropomorphic descriptions of machines” (p. 22).

The British cybernetics started in the basement room in the National Hospital for Nervous Diseases, and this was the first venue for the Ratio club’s meeting (Husbands & Holland, 2008). The Ratio club members were mostly from the brain and neuronal studies. This group of scholars pursued a universal intelligence theory that can explain information processing in both the brain and machine. This desire was well-reflected in one of the themes discussed in the Ratio Club: “Can the members agree on definitions, applicable equally to all systems—biological, physiological, physical, sociological—cf: feedback, stability, servo-mechanism” (Husbands & Holland, 2008, p. 119). The British cyberneticians defined intelligence as adaptive behavior that constantly modifies its behavior based on the input or feedback loop with a contact to the outer environment. The representative artifact reflecting this simple but powerful definition of intelligence was Ashby’s Homeostats. Ashby was a critical figure who initiated an innovative approach to brain study, using rigorous mathematical measures and brain mechanism modeling (Asaro, 2008). His Homeostats maintained the system's stability by controlling output proportional to its given input, which constantly changed. It was the simulation of the organic feedback loop system. Another was Walter’s Tortoises, highly esteemed as the first-ever mobile autonomous robot. This robot could autonomously move, navigating the environment with its multiple sensorimotor systems that imitated the human nervous system. One of the club members was Alan Turing, and his later work on the conceptualization of the first modern computer was deeply influenced by this early cybernetic movement (Husbands & Holland, 2008).

Alan Turing is widely recognized as a father of modern computing, who suggested a universal computation principle, which was widely adopted in modern computer design (Nilsson, 2010). Turing designed a computer as a universal problem solver, later called the Turing machine. He proved that every Turing machine could solve universal problems once it was given an encoded input in the tape indicating how the other Turing machines operate. That means each Turing machine is designed to simulate and mimic whatever the other machines do; thus it is a universal problem solver, in other words, the Universal Turing Machine (UTM) (Crane, 2003, p. 98). Turing assumed that intelligence is a mental capacity to “process representations in a systemic way” (Crane, 2003, p. 85). In this way, Turing indicated that the human brain is one of the machine kinds which is highly sophisticated enough to manipulate symbols. He truly believed that the human body could be built up with mechanical parts such as a camera, motors, and microphone (McCorduck, 2004). He said “The electrical circuits which are used in electronic computing machinery seem to have the essential property of nerves. They are able to transmit information from place to place, and also to store it” (Cited from McCorduck, 2004, p. 68). Given this claim, McCulloch and Pitts insisted that the human brain's neural system is a Turing machine with their experimental evidence. Von Neumann saw this discovery of McCulloch and Pitts as decisive proof that “anything that can be exhaustibly and unambiguously described is realizable by a Turing machine” (Piccinini, 2018, p. 436). The Universal Turing machine was indicative that there is a universally common intelligence identifiable across different embodiments. It means any function and structure logically recognizable by human cognition can be simulated by a Turing machine. One of the partial byproducts of this recognition, while having a prolonged impact on the following generations of scholars, was that a human mind is a Turing machine. Thus, the invention of the Turing machine made the distinction between human and the machine intelligence ever more blurred and ambiguous.

As much as he was dedicated to mechanizing the human mind, Alan Turing sought to create machine intelligence that could learn and evolve like human intelligence. He was eager to simulate the human learning mechanism in the advanced intelligent machines. According to Bostrom (2014), Turing assumed the child brain as the most idealistic model of an intelligent machine, quoting what he said: “Instead of trying to produce a programme to simulate the adult mind, why not rather try to produce one which simulates the child's? If this were then subjected to an appropriate course of education one would obtain the adult brain” (p. 23). Michie (2008) found Turing’s desire to approximate to adaptive human learning capacity from this following quote:

> Let us suppose that we have set up a machine with specific initial instruction tables, so constructed that these tables might on modify these tables… In such a case one could have to admit that the progress of the machine had not been foreseen when its original instructions were put in. It would be like a pupil who had learnt much from his master, but had added much more by his own work. When this happens I feel that one is obliged to regard the machine as showing intelligence. As soon as one can provide a reasonably large memory capacity it should be possible to begin to experiment on these lines. (p. 65)

John Von Neumann is a well-known figure as an architect of the modern computer. He completed the computer's basic structure with a sequential program, machine language program, and modifiable memory (Von Neumann, 2002). However, it is not widely acknowledged that the human brain mechanism inspired his modern computer design. In his book, *Computer and Brain*, he delicately compared the details of the artificial machine and human brain (Kurzweil, 2012b). This book shows how the early computer design and brain studies exchanged inspiration with each other. He is also a critical figure in the intellectual history of AI and mind sciences because of his membership to both first conferences of AI, Dartmouth Conference, and cognitive science, Hixon Symposium. He was also substantially immersed in the cybernetics movement - he was even a core member of the Macy conferences. His connection to Warren McCulloch is well-known, and he was inspired by the basic neural concepts such as conjunction, disjunction, and negation in organic feedback loops (Boden, 2006). In a sense, Von Neumann was the first person who explicitly translated organic neuron systems to design the mechanical digital computer by ignoring discrepancies between human and machine systems and seeking for convergence at the abstract level. His computer left a significant mark in the rest of AI history (Boden, 2006).

*Symbolism vs. Connectionism*

Since the AI study started its official history at the Dartmouth workshop in 1956, the AI scholars have been largely divided into symbolism and connectionism groups (Cardon et al., 2018). The most prominent approach in the early phase was symbolism. The representative figures of symbolism were John McCarthy, Marvin Minsky, Allen Newell, and Herbert A. Simon. The symbolism group identified the source of human-like intelligence as a symbolic logic system. This approach is often called deductive machine design (Cardon et al., 2018). That is because, in the symbolic AI, the machine has a predetermined set of rules and programs enabling it to manipulate the input symbols to create output symbols, so-called heuristics. Advocates of symbolism defined intelligence as a set of rules such as “searching, knowing, recognizing, trying, remembering, choosing, and the like” (Boden, 2006, p. 317). Simon defined learning as “a permanent alteration in the repertoire of heuristics to guide search and actions of an information processing system, involving knowledge acquisition and increasing complexity of perceptual chunks” (Kao & Venkatachalam, 2018, p. 14). For instance, McCarthy designed the AI program by inserting axioms in the system and let the machine manipulate such representation through its program (Russell & Norving, 2001).

The symbolism group referred to psychology and cognitive science and even significantly impacted these two fields. This interdisciplinary character of their research led them to attempt to humanize machines (Boden, 2006). Russell and Norvig (2001) named this AI approach as “thinking humanly approach” (p. 17). Accordingly, they proactively studied discoveries in cognitive science and adapted them to build up new machines. The symbolism group preferred to imitate the problem-solving mechanism of humans (Russell & Norvig, 2001).

In 1980, Simon and Newell developed an AI machine called SOAR, which was a goal-directed machine. For this, they utilized various conceptions of cognitive science such as “integrated perception, attention, memory, association/inference, analogy, and learning” (Boden, 2006, p. 433). As their research matured, Simon claimed that machines have intelligence once they can fulfill goal-oriented tasks and adaptive behavior in various settings, simulating in chess play, solving math problems, and diagnosing diseases (Frantz, 2003). The symbolism remained the most dominant paradigm of AI from the 1960s to the 1980s.

In contrast to symbolism, connectionism is an approach to design the intelligent machine in an inductive way (Cardon et al., 2018). The connectionism defined human thinking as parallel information processing at a massive scale in the neural nets, producing emergent behavior. The initial inventors of this bio-inspired design approach were McCulloch and Pitts. According to Medler (1998), they designed each neuron to take input individually, producing on/off mode accordingly, and then as the sum exceeds a certain arbitrarily set threshold, the bigger neuronal system can be activated, creating an emergent pattern of behavior. This design is called inductive machine design, in which AI is programmed to produce a specific program capturing a pattern in the world when the massive amount of data is inserted into the minimally preconfigured program. This minimally designed program is a bio-inspired neural net. Specifically, this machine should be equipped with the following four properties: "1) The connectivity of units, 2) the activation function of units, 3) the nature of the learning procedure that modifies the connections between units, and 4) how the network is interpreted semantically" (Medler, 1998, p. 22).

The formalism of the connectionists' modeling of the AI was developed further after McCulloch and Pitts' model. Selfridge developed a Pandemonium model, which processed image patterns on a massively parallel scale (Medler, 1998). This model processed information like nerve cells at the lower level while could manipulate the symbols at a higher level. Nilsson (2010) deemed the Pandemonium foreshadowed the most recently developed method of machine learning. The Perceptron developed by

Frank Rosenblatt is considered the first connectionists' model that pioneered the machine learning algorithm (Boden, 2006; Cardon et al., 2018; Medler, 1998; Nilsson, 2010). Inspired by McCulloch and Pitts neural net modeling, Rosenblatt completed the machine learning system's design around the 1960s, about 50 years before it came to be recognized as the most potent AI architecture. He initially designed this model with a conviction that this would be the universal intelligence modeling not only for machines but for humans (Nilsson, 2010). He conceived that the neural net is constituted with input, intermediate, and output neurons. He assumed that only the intermediate or hidden layer of neurons could learn through the training, modifying its weight parameters with a statistical mechanism.

The Perceptron became a model still widely used in machine learning software algorithm with minor modifications. The learning process in it is reduced into the simple mechanical process modifying parameters in the nerve system, deciding whether to turn on or off the specific level of nerve cells at each level. Konar (2000) described this learning process with a metaphor to a child's learning of pronunciation:

> The hearing system of the child receives the pronunciation of the character "A" and the voice system attempts to imitate it. The difference of the mother's and the child's pronunciation, hereafter the error signal, is received by the child's learning system through the auditory nerve, and an actuation signal is generated by the learning system through a motor nerve for adjustment of the pronunciation of the child. The adaptation of the child's voice system is continued until the amplitude of the error signal is insignificantly low. (pp. 35-36)

The connectionists saw that intelligent behavior emerges as a consequence of a massive cascading effect activating nerve cells. This modeling's strength is the high responsiveness to the external stimuli modifying its next behavior to produce an optimal outcome (Raschka & Mirjalili, 2017). Again, this design is inductive as the intermediate layers of nerve cells are considered to operate in correlation with the input and output, which is only minimally designed in a statistical sense.

Overall, the symbolism's long endeavor since its inception from Dartmouth was sidelined from the mainstream AI studies with the rise of parallel distributed processing (PDP) model in the connectionism group (Cardon et al., 2018; Nilsson, 2010). David Rumelhart, Geoffrey E. Hinton, and Ronald J. Williams announced their discovery, adding a mathematical modification to the Perceptron in a book *Parallel distributed processing: Explorations in the microstructures of cognition* in 1986, and reinstated their previous modeling in an article *Learning representations by back-propagating error* (Rumelhart et al., 1988). It became an early version of a deep neural network, the most dominant machine learning model in the 2010s. There was another new technological breakthrough in the connectionists' approach of AI called reinforcement learning. The term reinforcement learning was borrowed from behaviorists' psychology, which emphasized the trial and error process as a learning loop (Nilsson, 2010; Sutton & Barto, 2018) - Nilsson (2010) and Sutton and Barto (2018) explicitly mentioned the name of Thorndike as an intellectual father of reinforcement learning. In reinforcement learning, the intelligent agent learns a shortcut to the outcome through random walks and valuation to each selected pathways. Each pathway's valuation is continuously updated and modified based on its additional experience of trial and error (Nilsson, 2010).

By far, the history of modern AI was illuminated with its connection to the understanding of human intelligence and its mechanism. This specific perspective taught us that the modern approach to developing advanced autonomous machines originated from the abstracted knowledge about human intelligence and the brain. The connectionists borrowed McCulloch and Pitts' discovery of the neural net structure and its function to create a highly homeostatic machine, automatically adapting to the changing environment. Reinforcement learning is based on psychological behaviorism and is closely related to Thorndike's study of stimulus and reaction mechanisms. The symbolism always pursued knowledge from psychology, cognitive science, and biology to simulate almost human-like intelligence mechanisms in machines.

# CHAPTER 4. METHODS

The literature review indicates that intelligence, traditionally considered an animated feature, has been understood as universal even for inanimate beings, including machines. In the same manner, education studies have also strived to gain authority by mechanizing its logic and assumption about the human mind, traversing to hard sciences. This suggests that there has been a co-space where the studies, ideas, and meanings about human and machine intelligence interacting and even converging with each other. This study calls this interdisciplinary domain a *cyborg space*, where the study of universal intelligence looms over. Given that this convergence has never been explicitly explored before through any measure, this research aims to analytically describe how the studies of a human and machine intelligence have interacted with each other over time, thus creating an interdisciplinary space between the two seemingly distanced disciplines that share a common paradigm assuming that human and machine is compatible and interchangeable with each other. To reveal this hybridity, this research used sequential explanatory mixed methods. The mixed-methods included quantitative and qualitative research with a rigorous process of data collection and analysis, ultimately aiming to integrate views obtained from the different methodologies (Cresswell, 2014).

The mixed-method study involves quantitative and qualitative data collection and analysis, ultimately aiming to integrate views from the different methods (Cresswell, 2014). There are various reasons to adopt this relatively new research approach, but mostly they pursue dialectical resolution in the face of conflict and tension between qualitative and quantitative research traditions (Creswell et al., 2011). Occasionally but not exhaustively, the mixed method entails pragmatists' approach given the primary goal of answering the diverse research questions and problems. Also, recently, there is a growing pressure to increase research reliability and validity by triangulating various data sources (Hesse-Biber, 2010). In other words, there is a growing demand for the cross-validation between different research approaches. More broadly, this research approach may include the following types of research:

- focusing on research questions that call for real-life contextual understandings, multi-level perspectives, and cultural influences;
- employing rigorous quantitative research assessing magnitude and frequency of constructs and rigorous qualitative research exploring the meaning and understanding of constructs;
- utilizing multiple methods (e.g., intervention trials and in-depth interviews);
- intentionally integrating or combining these methods to draw on the strengths of each; and
- framing the investigation within philosophical and theoretical positions.

(Creswell et al., 2011, p. 1)

According to Cresswell (2014), the idea of combining qualitative and quantitative research methods came into being almost a half-century ago. Campbell and Fisk, in 1959, attempted to combine multiple methods and data in psychology research, which triggered more widespread usage of the mixed-method approach. The early recognition was that every method has its weaknesses to cover the whole research topic; but, these limitations could be  overcome by combining the methods. This approach brought to the fore the triangulation method in qualitative research. In the 1990s, a systematic effort to establish the research tradition of the mixed method first emerged. In 2003, the *Handbook of*

*Mixed Methods in the Social and Behavior Sciences* by Tashakkori and Teddlie was published, and this book comprehensively summarized the characteristics and procedures of the mixed methods. Now that the mixed methods gained growing attention, the community is expanding rapidly. Research journals dedicated to the mixed methods include *Journal of Mixed Method Research, Quality and Quantity, Field Methods*, and *International Journal of Multiple Research Approaches* (Cresswell, 2014).

Creswell (2014) suggested several different research designs in the mixed-method approach based on the procedural method of merging and analyzing data. Firstly, there is an approach called a *convergent parallel mixed method design*. Creswell said that when the researchers were first motivated to use the mixed methods, they would present both qualitative and quantitative in parallel. The critical feature of it is to use the same variables and constructs to guide research questions and instrument design in both qualitative and quantitative research. Secondly, there is an *explanatory sequential mixed method design*. The researchers with quantitative research background preferred this design. In this design, the researchers initially conduct quantitative research and then design the next qualitative research session according to the quantitative research outcome. The quantitative research guides the next steps to follow in the complementary qualitative research session. The qualitative data analysis is supposed to provide more in-depth information missing in the previous quantitative analysis. Thirdly, an *exploratory sequential mixed methods design* is precisely opposite to the explanatory sequential mixed method's order. For this time, the researchers do qualitative research first and design the next quantitative research based on the results. The intention is to generalize the findings from the narrow population in the qualitative research to the bigger size of a quantitative research population. For instance, the researchers can use an interview with a representative small population sample to develop a valid psychometric questionnaire.

Given the research questions design to emphasize quantitative data analysis and complement more detailed information through qualitative research, this research used the sequential exploratory study design, which allows the quantitative research data to illuminate more details in support of the qualitative research data (Cresswell, 2014). Accordingly, this research first implemented a quantitative study and then conducted a qualitative study to supplement more details.

# CHAPTER 5. QUANTITATIVE ANALYSIS: BIBLIOGRAPHIC NETWORK ANALYSIS

## Methodological Background

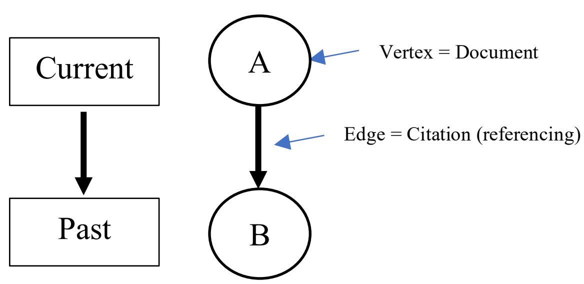
### *Bibliographic Network Analysis*

This research will use *bibliographic network analysis method*, one of the network science methods, to measure the extent to which educational psychology and AI engineering converge towards each other. Newman (2010) defined the network as "a collection of points joined together in pairs by lines" (p. 1). Dots in the network are called vertices, and lines connecting each dot are called edges. In other words, a network is a *pattern of connections* and network analysis seeks to capture a basic pattern of the connections by reducing complicated network features in nature into a rather abstract representation composed of vertices and edges. Bibliographic network (or citation network) means networks existing in the web of documents through a practice of citation. Academic, legal, and patent documents commonly attach a bibliography list at the end of the writing to reveal their source references. In this bibliographic network, each document is a vertex, and once document A refers to B, it is considered a directed edge between them (Newmann, 2010). Although there are various reasons to refer to a particular document, there is no doubt that researchers using citation networks acknowledge the fact that academic writing and knowledge production is not an isolated process, but a collective and cumulative process exchanging influence to each other (Biscaro & Giupponi, 2014). For instance, authors refer to papers by other authors to give them credit or even present radical opposition to the previous papers, forming a group of scholars sharing common knowledge and cultural bases. Neumann (2010) considered "In general, however, if one paper cites another it is usually an indication that the contents of the earlier paper are relevant in some way to those of the later one, and hence citation networks (bibliographic networks) are networks of relatedness of subject matter" (p. 68).

The bibliographic network has some distinctive features compared with the other network structures. Leicht et al. (2007) and Neuman (2010) provided the most accurate and comprehensive summary of the features. First, the bibliographic network is a directed network; the arrow comes from one document to another. For instance, if document A refers to document B, then the arrow comes from A to B but not vice versa. Second, the bibliographic network is relatively static in that once the document was published, the reference cannot be modified. On that account, the network tends to grow as time passes but never diminishes. Leicht et al. (2007) said that the bibliographic network's static feature intrigued many information scientists providing a laboratory-like condition of the network dynamics. Lastly, the bibliographic network is acyclic. Acyclic means that there is no circular loop in the network. It is a directed network; the arrow only comes from relatively current to the past one. For instance, document B published in 1950 cannot refer to document A published in 1980 as B was not published when A was published.

Figure 1.

*Bibliographic Network's Characteristic*



*Measuring Convergence and Interdisciplinarity*

This study aims to measure the level of convergence between the studies of educational psychology and AI engineering. Information science studies have focused on this issue for a while because they could predict a rise and fall of the interdisciplinary research field by measuring convergence and divergence between different group of papers (Karunan et al., 2017; McCain, 1998; Leydesdorff, 2006; Leydesdorff & Rafols, 2011; Leydesdorff et al., 2017; Porter & Rafols, 2009). There are three different ways to measure a convergence between different disciplinary studies. First, the number of common papers categorized in two distinctive fields can represent a degree of convergence. Karunan et al. (2017) measured the degree of interdisciplinarity between two fields by measuring the number of common papers, labeled as multiple categories, thus searched in both topics in the search query. They calculated interdisciplinarity through a ratio of the number of common papers ($A \cap B$) to combined papers ($A \cup B$). Second, the exchanged number of direct citations between two fields can indicate a degree of convergence. Frank et al. (2019) studied citation networks between AI studies and other fields of study, including philosophy, art, sociology, chemistry, economy, etc. Based on the number of exchanged direct references, they concluded that there was less and even decreasing connection between AI and other studies such as psychology and philosophy. Similarly, Núñez et al. (2019) found that cognitive science, including AI studies as its relevant field, had developed less interdisciplinarity, observing the diminishing number of direct cross-reference with the other area of studies. Lastly, there is a measure of *bibliographic coupling*, which calculate the strength of the tie between documents at an aggregated level of journal, institution, author, and nation, based on the number of commonly cited references by two different disciplinary studies (Glanzel & Czerwon,1996; Gazni & Didegah, 2016; Jarneving, 2007; Kessler, 1963; Sen & Gan, 1983; Thijs et al., 2015; Zhao & Strotmann, 2008). Newmann (2010) put that "Two vertices in a network are structurally equivalent if they share many of the same network neighbors" (p. 211). The bibliographic coupling measure indicates the strength of the tie between documents simply measured by counting the number of shared network neighbors between vertex i and j – in the bibliographic network study, each vertex refers to the document, and each edge between the vertex means shared cited references.

This research selected *bibliographic coupling* to measure AI and educational psychology research's similarity instead of using common papers and a direct citation network. There are two reasons for this. First, this research sought to find interdisciplinary interaction at the fundamental paradigmatic level. The paradigmatic interaction can be less visible because it does not exchange information directly with each other but brings a much broader impact on the fields as they influence symbols, concepts, theories, and models. This research aimed to reveal that even the papers not explicitly pursuing interdisciplinary studies could share similarities with the other fields because they are structurally embedded in the larger paradigmatic network where the interdisciplinarity emerges. Like the tip of the iceberg, a vast hidden layer of the network connection between the seemingly distanced two fields may lie beneath the surface. By definition, the *bibliographic coupling* means the network's hidden layer, the latent structure of the network showing co-membership and structural similarity among clusters (see, figure 2). As Klavans and Boyack (2007) said, the coupling tie and its strength represent a research paradigm that a group of researchers is commonly grounded. Thus, this research excluded self-claimed interdisciplinary papers, because they explicitly pursue interdisciplinary studies. For instance, the *International Journal of Artificial Intelligence in Education*, as the name indicative of interdisciplinary field of AI and education studies, was categorized into *Computer Science, Interdisciplinary Applications* in the search database, and this research did not include Interdisciplinary Applications studies from the database's category search. Second, even though some may want to study with the common papers and direct citation, there are very marginal numbers of common papers and direct citations between AI and educational psychology studies. A previous database search query did not create any overlapping papers or journals (common papers or journals)[1]. It means either that the two fields do not interact with each other through explicit and direct ways or that the database puts aside interdisciplinary studies to the separate categories reserved for interdisciplinary or applied studies, making them almost invisible in the category search. Therefore, the bibliographic coupling network was the best known measure of paradigmatic convergence and interdisciplinarity between different disciplinary studies for this research.

Figure 2.

*Comparison of Directed Citation and Undirected Bibliographic Coupling Network*

[1] Once I put the search keyword as "psychology, educational AND artificial intelligence, computer science", it produced zero result, meaning that the database does not have any papers belong to educational psychology and artificial intelligence at the same time.

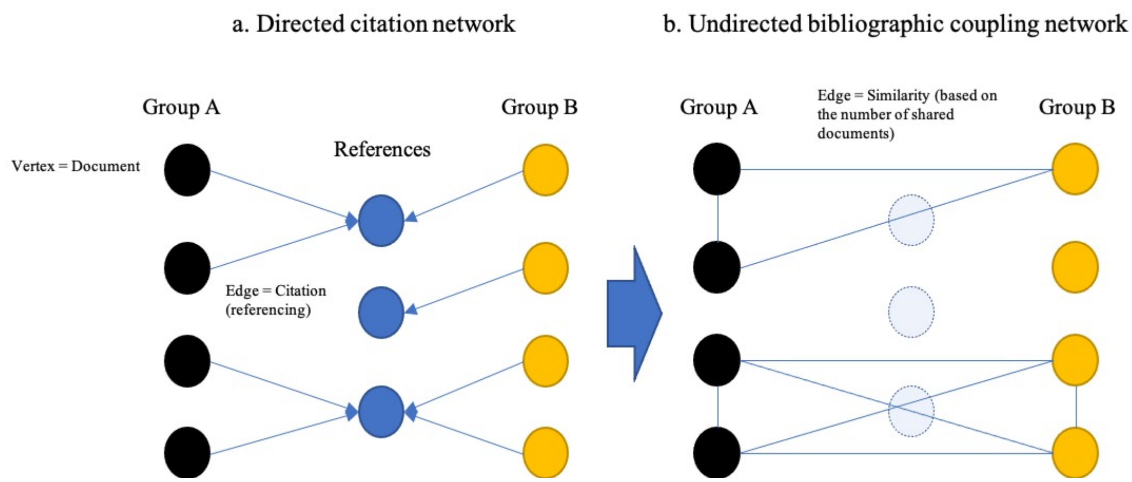



Table 1

*Multiple Layers Of Interdisciplinarity Measures*

| **Methods** | **Size** | **Interdisciplinary Connection** | **How Connection is Made** |
|---|---|---|---|
| Common papers | Marginally small | Explicit, Direct | Self-claimed |
| Direct citation | Relatively small | Direct | Direct cross-referencing |
| Bibliographic coupling | Relatively large | Implicit, Indirect | Indirect connection through the shared references |

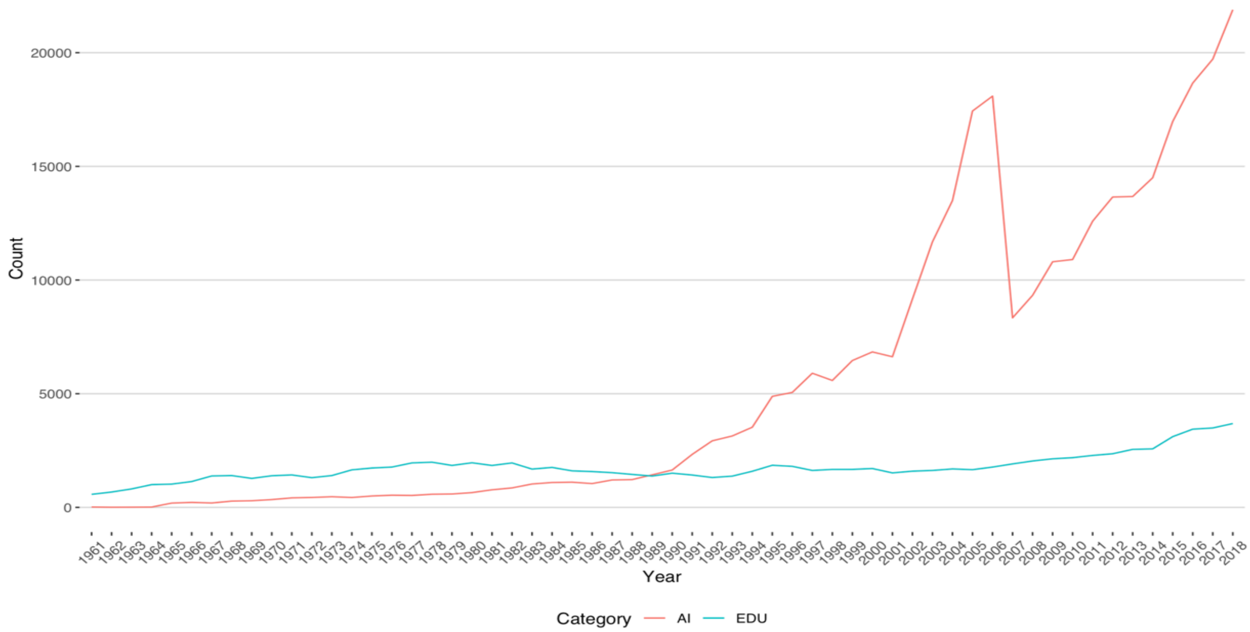
## Method

### *Sampling*

The bibliographic data was retrieved from Web of Science (WoS), one of the largest databases of academic writings operated by Clarivate Analytics, previously run by Thomson Reuters. WoS is a recognized academic authority that started measuring academic impact and performance index, developing Science Citation Index (SCI) in 1964. WoS pioneered online and digital database systems, publishing its index through CD-ROM and the internet. With this vast amount of digital academic data, Eugene Garfield, the founding father of WoS, opened up the field of information science dedicated to academic indexing and citation network analysis (Moed, 2005). In this accumulated history, WoS was a preferred online database to retrieve a large set of bibliographic data (Calero-Medina & Noyons, 2008; Kajikawa et al., 2007; Leydesdorff & Rafols, 2009; Leydesdorff et al., 2012).

Figure 3.

*Annual Publication of Educational Psychology and AI Papers*




For this research, I searched the large set of bibliographical data of AI and education studies using the Web of Science Core Collection database. This Core Collection database provides more detailed and full citation information than general search, including topic, country, institution, Web of Science category, funding source, geographical region, and etc. I searched a broad set of data related to AI and education psychology using the Web of Science category[2]. The Web of Science category (WC) has

[2] This was the search term and specific condition: WC=(Computer Science, Artificial Intelligence OR Computer Science, Cybernetics OR Robotics). Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, ESCI. And WC=(Psychology, Educational). Timespan: All years. Indexes: SCI-EXPANDED, SSCI, A&HCI, ESCI.

been used to map out academic disciplines before in the other studies (Leydesdorff & Rafols, 2009; Moya-Anegón et al., 2007; Rafols et al., 2010). This approach allowed access to the vast amount of papers from theoretical to various applied studies of AI and educational psychology, yielding 364,607 unique papers for the AI studies from 1900 to 2018 and 120,049 unique papers for educational psychology from 1900 to 2018. The search query in the WoS retrieved 460,358 unique papers. Then, the initially searched papers were refined filtering out papers with anonymous authors, omitted citations list, and publication before 1961, the year when the AI paper firstly appeared in the database. In total, 33,003 papers were identified inappropriate after the screening in the R software program. After removing these papers, the count was reduced to 427,355.

Table 1.

*The Number of Papers Per Each Category*

| | Artificial Intelligence | Educational Psychology | Total |
|---|---|---|---|
| Timespan | 1961 - 2018 | | |
| Sources (Journals, Books, etc) | 1,014 | 109 | 1,124 |
| Documents | 312,292 | 101,672 | 427,355 |
| Average years from publication | 14.2 | 27.1 | 16.9 |
| Average citations per documents | 21.58 | 20.99 | 20.78 |
| References | 3,800,242 | 1,205,653 | 5,118,168 |
| Authors | 258,382 | 100,505 | 355,983 |
| Documents per author | 1.21 | 1.01 | 1.2 |
| Annual production growth rate | 13.38 | 3.32 | 5.53 |

*Analysis*

This research used multiple data analysis methods, including bibliographic coupling, Jaccard Index, inter-intra coupling ratio analysis, and community detection analysis methods. For this analysis, I used the Bibliometrix and igraph packages provided by R software, specialized for online bibliographic data analysis and network analysis per se, with a slight modification in their original function codes (Aria & Cuccurullo, 2017, Csardi & Nepusz, 2006).

**Bibliographic Coupling**. The bibliographic coupling has been widely adopted in the bibliometric studies so far. Since Kessler (1963) first introduced this concept to measure the strength of ties between documents and to cluster documents based on this coupling strength, bibliographic coupling has been used to map a the research field (Zhao & Strotmann, 2008; Boyack et al., 2008), to identify the source of research ideas (Biscaro & Giupponi, 2014), to identify new research fronts and emerging topics (Jarneving, 2007), and to measure the structural similarity between journals and databases (Klavans & Boyack, 2007). The shared number of commonly cited reference between documents is calculated as in the formula below:

$$n_{ij}= \sum_k A_{ik}A_{kj}$$

In this formula, $n_{ij}$ is the number of shared cited references between document i and j, and $A_{ik}$ and $A_{kj}$ each indicates an adjacency matrix between document i and cited references k, and between k and j. By multiplying the two adjacency matrices and counting the sum, the total number of shared documents between i and j documents can be simply counted. In particular, the bibliographic coupling converts the directed citation network into an undirected network presenting each vertex as a document and their connected edges as a degree of similarity or tie between vertices.

Then, this individual document level coupling count can be aggregated to the coupling strength of higher cluster levels such as author and journal. There are many different ways of aggregation, but no consensus on the best method yet. Ma (2012) suggested that there are three ways of aggregating document-document coupling to the author-author or journal-journal coupling; simple, minimum, and combined method. The simple method is to simply assign a value of one for at least one shared cited reference, or zero otherwise regardless of how many times the cited reference appeared in each journal or author. Then, the binary score per each cited reference is aggregated at the journal or author level to show the number of shared cited references between journals or authors. For instance, the simple method measures the coupling strength as one, although the shared cited reference A appears three times in two journals of J1 and J2. If the J1 and J2 share three distinctive cited references, their coupling strength is simply three. This simple method has been widely adopted for its convenience of calculation (Boyack et al., 2008; Rousseau, 2010; Thijs et al., 2015). The minimum method is selecting the smaller number of times that the shared cited reference appears in each journal as a coupling strength. For instance, the shared cited reference A appears two times in journal J1 and three times in journal J2, then two is selected as the coupling strength between J1 and J2. The combined method is to multiply the appearance number of the shared cited reference A in journals J1 and J2. If the A appears two and three times for each J1 and J2 journal, then the coupling strength of J1 and J2 is 2 x 3=6. This combined method has also been commonly used in research because of its computing efficiency in processing big data (Aria & Cuccurullo, 2017; Batagelj & Cerinšek, 2013; Yan & Ding, 2012). This research used a combined method to aggregate document-document bibliographic coupling to the journal-journal coupling strength to visualize and do the network analysis. Hence, the edge connecting two journal vertices means an abstracted index representing coupling strength between the journals in this research. The formula to produce aggregated coupling strength is:

$$JC = DJ^T * DC \quad (1) \qquad\qquad JJ = JC * JC^T (2)$$

DJ is the [document x journal] incidence matrix and DC is [document x cited reference] incidence matrix. JC, [journal x cited reference] matrix can be derived from matrix multiplication of transposed DJ and DC. Then, JJ, [journal x journal] adjacency matrix can be calculated with JC and transposed JC matrix multiplication. In the JJ matrix, each matrix element indicates bibliographic coupling strength between two journals. This research presented the bibliographic coupling strength using descriptive statistics, network visualization, and adjacency matrix visualization.

**Jaccard Similarity Index**. The Jaccard similarity index indicates the ratio of inter-cluster to total network coupling strength (Leydesdorff, 2007). The Jaccard similarity index helps to measure the relative size of interdisciplinary coupling against total coupling in the journal bibliographic coupling network. Then, this measure can be calculated by dividing a sum of inter-cluster coupling strength by a total sum of coupling strength of the entire network.

$$\frac{EDU \cap AI}{EDU \cup AI}$$

The Jaccard index is a normalization score ranging from 0 to 1; 0 means no overlap between two clusters or subgraphs, while 1 means perfect overlap. This index entails the relative size of the intersection against that of union between two clusters or subgraphs. Thus, this research used this normalized similarity index as an indicator of interdisciplinary convergence between the studies of AI and educational psychology.

**Inter- To Intra-Disciplinary Coupling Ratio.** The journals in the same disciplines can be connected to each other through a bibliographic coupling tie, and such an intra-disciplinary coupling is generally supposed to be bigger than inter-disciplinary coupling. It is because the journals in the same discipline are more likely to share common references with each other than those in the other disciplines considering their topic similarity and community culture. Thus, this research also compared the size of an intersection to that of a single cluster. That is, the relative size of inter-disciplinary coupling between AI and educational psychology studies was measured against the single intra-disciplinary coupling strength in total. The score also ranges from 0 to 1, and when the score approximates to 1, a perfect score, then it may mean the discipline is submerged into another discipline.

$$\frac{\mathrm{EDU} \cap \mathrm{AI}}{\mathrm{EDU}} \ (1) \qquad \frac{\mathrm{EDU} \cap \mathrm{AI}}{\mathrm{AI}} \ (2)$$

**Community Detection.** Another way of measuring paradigmatic convergence between two different disciplines is the community detection method in the network analysis. While there are many kinds of community detection methods, this research used fast-and-greedy and Louvain clustering methods. The fast-greedy and Louvain clustering methods are commonly based on the modularity score. The modularity is a discrepancy between the observed number of edges in the same clusters and the edges in the random graph, the graph where the vertices are connected totally by a random chance. The higher modularity score means the clustering or modularity of the given graph is less by chance but rather due to systemic reason (Csardi et al., 2016; Newman, 2004). These methods start by assuming that every single vertex belongs to separate clusters. Then, it calculates the change in modularity score in every step of merging two clusters and chooses the best option that can maximize modularity increase. Then, this research compared predetermined binary clustering according to the disciplines of AI and educational psychology to the identified network clusters based on the community detection methods. Using the fast-greedy community detection method, this research identified journals that belong to the clusters having both AI and educational psychology studies, labeling them as bridge or interdisciplinary journals located in the middle of the two disciplines. Using the Louvain community detection method, this research identified intra-journal clusters in each educational psychology and AI studies to see the differential network connection patterns depending on the sub-clusters in each field.

## Findings

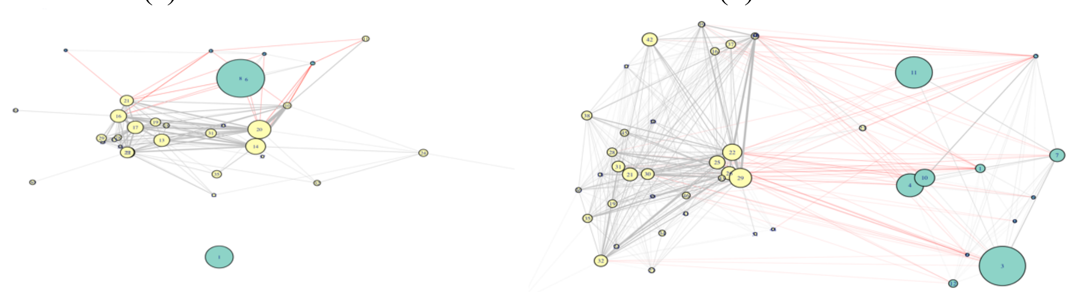
### *Bibliographic Coupling Network*

The journal-journal bibliographic coupling strength was measured to construct a coupling network. The coupling network showed a difference across the time from the 1960s to 2010s. Figure 4 indicates the journal-journal coupling network. The yellow vertices mean educational psychology journals, while the green vertices are AI journals. The grey edges represent intra-disciplinary edges, connecting journals only in the same disciplines, and red edges mean inter-disciplinary edges, linking journals only in the different disciplines. The size of each vertex correlates with Pagerank centrality, one of the centrality measures giving more credit to the vertices having neighbors with higher degree centrality (Newman, 2010, p. 175). The layout function positioning the vertices in the space was the multi-dimensional scaling (MDS). The figure clearly shows that the absolute size of the network has grown for the last half-centuries with an increasing number of journals and bibliographic coupling strength. The number of journals increased from 37 in the 1960s to 1,120 in the 2010s, growing almost 30 times larger. The number of edges between journals increased from 149 to 108,939 from the 1960s to the 2010s, the size about 731 times larger. The summation of bibliographic coupling strength, indicating similarity and interdisciplinarity of the two disciplines, increased from 1,654 to 3,077,102 in the same period. The network plot also shows a clear distance between educational psychology and AI studies as the vertices of the two fields are located at a separate left and right position in the network space. However, the connection between the two fields has been maintained and even dramatically increased over time. The network figures show that the red edges in the middle, indicating interdisciplinary edges, have increased over time with the growing overall size of the network.

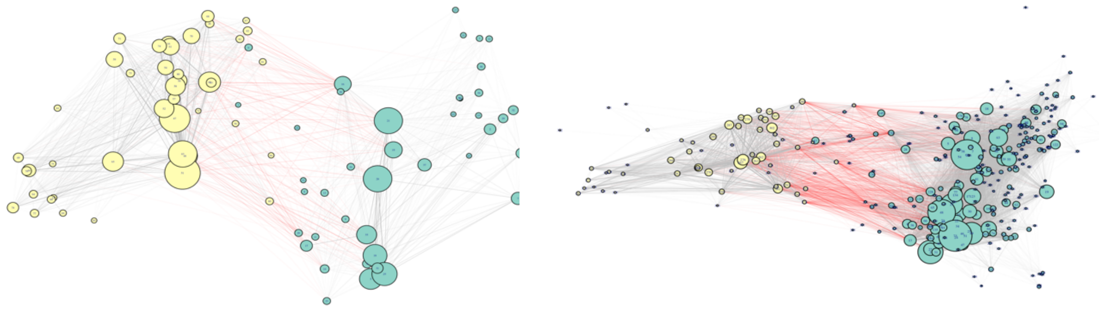
Figure 4.

*Network Visualization: Evolution of Journal-Journal Bibliographic Coupling Network from 1961 to 2018*

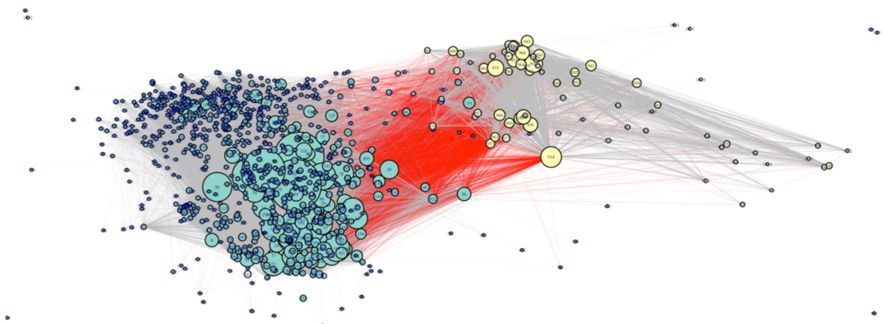
(a) 1961-1970 (b) 1961-1980



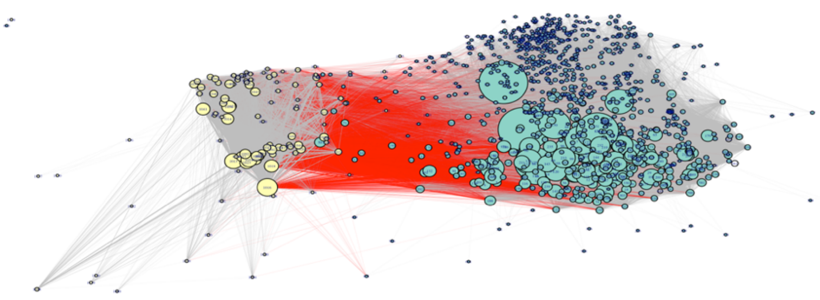
© 1961-1990 (d) 1961-2000



(e) 1961-2010



(f) 1961-2018



The visualization of the journal-journal adjacency matrix in the figure 5 also presents a similar result with the bibliographic coupling network visualization. The x-axis and y-axis of the figure indicate the same list of 1,120 journals from AI and educational psychology studies between 1960 and 2018. The dots on the pane mean a pair of journals have at least one shared common reference. Mostly, the clusters at the bottom-left are common references between AI journals, while those at the top-right are common references between educational psychology journals. The figures show a clear trend of an increasing number of AI studies and their clustering, while the educational psychology studies are pushed to the top-right corner of the pane. The dots in the top-left and bottom-right corner indicate shared documents across AI and educational psychology disciplines, and the figures show they grow over time.

Figure 5.

*Journal-Journal Bibliographic Coupling Adjacency Matrix*

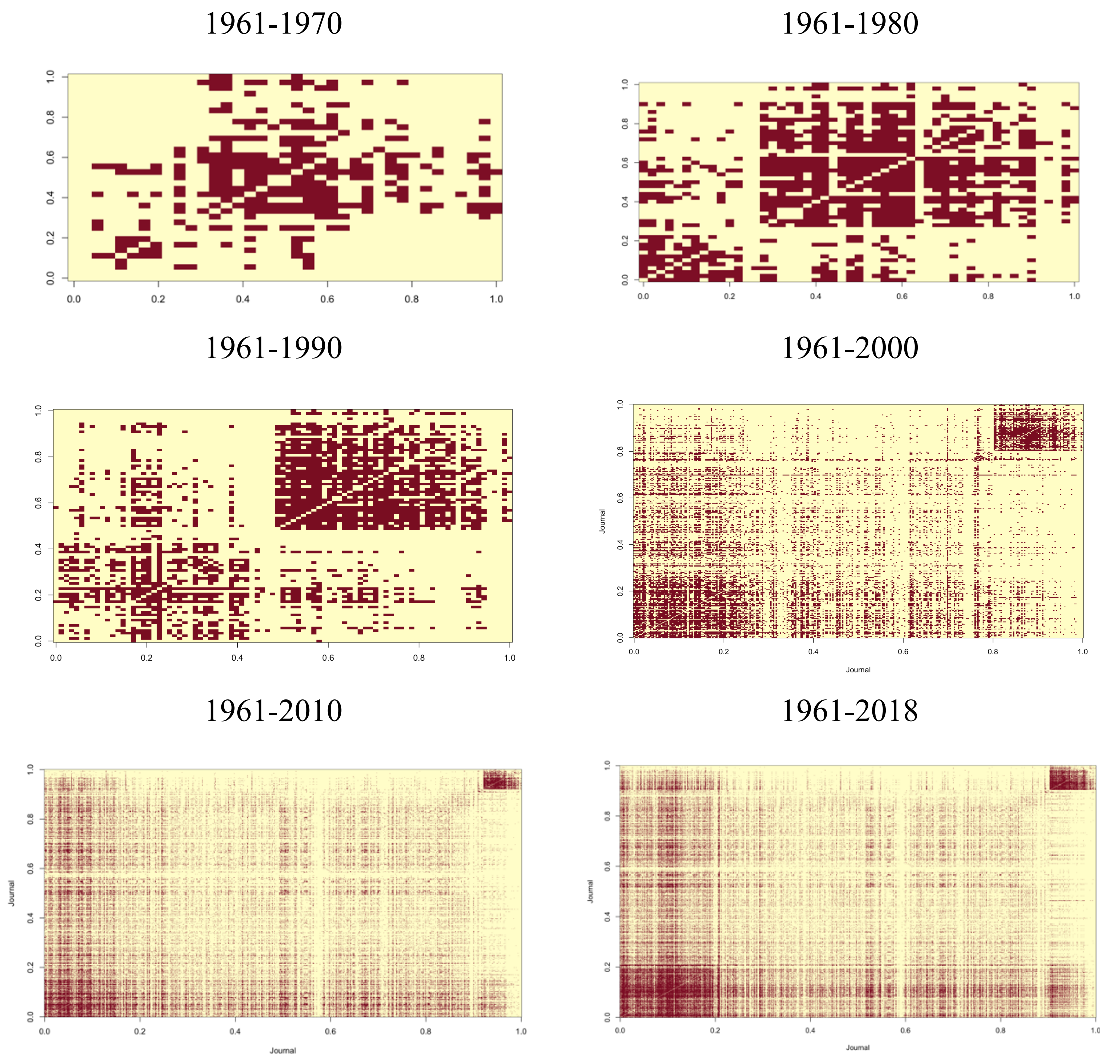



Table 2.

*Descriptive Statistics of Journal-Journal Bibliographic Coupling Network*

| N | 1960~1970 | 1960~1980 | 1960~1990 | 1960~2000 | 1960~2010 | 1960~2018 |
|---|---|---|---|---|---|---|
| **Vertex** | 37 | 51 | 89 | 326 | 1,014 | 1,120 |
| **Edge** | 149 | 472 | 1,100 | 9,841 | 72,551 | 108,939 |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Edge weight** **(Coupling strength)** | 1,654 | 10,787 | 31,597 | 163,062 | 1,000,600 | 3,077,102 |
| **Mean edge weight** | 11.10 | 22.85 | 28.72 | 16.56 | 13.79 | 28.24 |

*Jaccard Similarity Index*

However, the increase of the absolute network size does not guarantee that there has been a growing trend of paradigmatic convergence between educational psychology and AI studies. Hence, this research examined normalized scores measuring structural similarity between different network clusters such as the Jaccard index and Inter-Intra network ratio ranging from 0 to 1. Figure 6 and Table 3 show that the Jaccard index has increased over time from about 1.5% to almost 2.2%, meaning that the proportion of interdisciplinary coupling strength has increased slightly relative to the total bibliography coupling strength in the entire graph. The historical trend shows that the size of interdisciplinary coupling peaked in the 1980s, then continuously decreased over time until the 2000s. Then, only recently, this decreasing trend was reversed to reach the peak again in the 2010s. This result shows a different dimension of interdisciplinary network, compared to the study of Frank et al. (2019), which suggested decreasing strength of direct reference between AI and other disciplines, including psychology studies. This finding rather indicates the strength of reference between AI and educational psychology studies in terms of the bibliographic coupling maintained the same or even strengthened slightly over time.

*Inter-Intra Disciplinary Coupling Ratio*

The inter-intradisciplinary coupling ratio reveals a more detailed trend inside the bibliographic coupling network, making visible another facet of the bibliographic network. The score of the inter-intra coupling ratio usually indicates the relative size of interdisciplinary coupling strength to the intradisciplinary coupling strength. For instance, the inter-intra coupling ratio of educational psychology in the 2010s is about 26%, and it means the coupling strength between AI and educational psychology studies is about 26% of the coupling strength only between educational psychology journals in this network. The 26% is not a trivial score of inter-intra coupling ratio because intra-coupling strength is a significant measure of clustering journals into the same discipline. As this coupling ratio gets close to 1, the disciplinary boundary can become much more blurred while interdisciplinarity emerges, and even dismantle the established disciplinary boundary. The contradictory downward and upward trend of the ratio score in table 3 may entail that the AI studies are becoming multidisciplinary, covering a vast range of social and behavioral science studies as it grows its publication size rapidly. It can also be interpreted that most key references shared among the educational psychology journals are also cited in the AI journals and provide a theoretical background to develop AI studies.

Figure 6.

*Jaccard Similarity Index from 1961 to 2019*

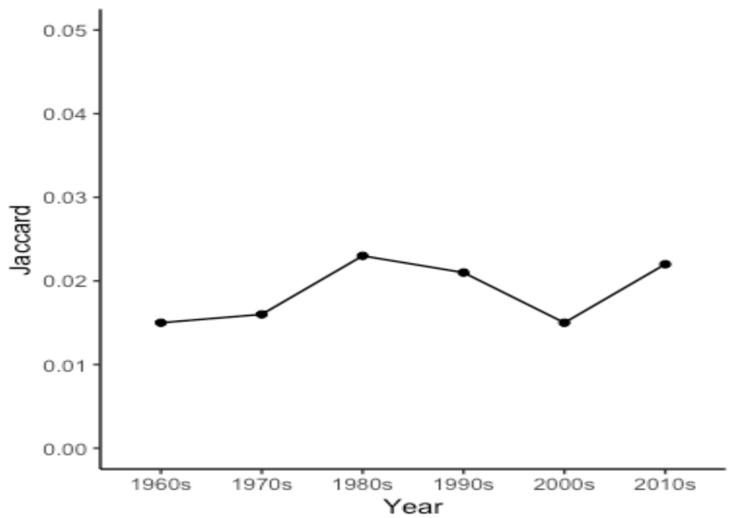



Table 3.

*Jaccard Similarity Index and Inter-Intra Coupling Scores per Discipline*

| N | 1960~1970 | 1960~1980 | 1960~1990 | 1960~2000 | 1960~2010 | 1960~2018 |
|---|---|---|---|---|---|---|
| **Edge count** | **133** | **396** | **898** | **8,952** | **69,621** | **101,082** |
| EDU-EDU | 125 | 350 | 583 | 1,122 | 1,585 | 2,965 |
| AI-AI | 8 | 46 | 315 | 7,830 | 68,036 | 98,117 |
| EDU-AI | 16 | 76 | 202 | 889 | 2,930 | 7,857 |
| **Edge weight (Coupling strength)** | **1,654** | **10,787** | **31,597** | **163,062** | **1,000,600** | **3,077,102** |
| EDU-EDU | 1619 | 10,182 | 23,591 | 42,761 | 85,550 | 187,569 |
| AI-AI | 10 | 430 | 7,290 | 116,852 | 900,303 | 2,822,206 |
| EDU-AI | 25 | 175 | 716 | 3,449 | 14,747 | 67,327 |
| **Jaccard Index** | **0.015** | **0.016** | **0.023** | **0.021** | **0.015** | **0.022** |
| **Inter-intra coupling (Edu)** | **0.015** | **0.017** | **0.029** | **0.075** | **0.147** | **0.264** |
| **Inter-intra coupling (AI)** | **0.714** | **0.289** | **0.089** | **0.029** | **0.016** | **0.023** |

*Community Detection*

**Identifying Interdisciplinary Journals.** Figure 7 and table 4 suggest the result of fast-greedy community detection in the journal-journal bibliographic coupling network. The community detection algorithm identified three clusters in the entire network, cluster one for AI, cluster two for also AI, and cluster three for a mixture of AI and educational psychology journals. The red-colored vertices and edges in the network of figure 7 indicates AI journals clustered together with the educational psychology journals, and their coupling edges. These interdisciplinary journals, all of which are AI studies, were identified to belong to a cluster where most of its member journals are from educational psychology studies. That is, these journals in table 4 were clustered together with the educational psychology studies, although they are AI journals according to the WoS database. There are many journals related to human-computer interaction, such as *Advances In Human-Computer Interaction, International Journal Of Human-Computer Interaction, Human-Computer Interaction, Journal Of Human-Robot Interaction*, and *Brain-Computer Interfaces: Lab Experiments To Real-World Applications*. Also, there are cybernetics studies of *IEEE Systems Man And Cybernetics Magazine* and *Engineering Cybernetics*. In aggregate, these interdisciplinary studies identified in the community detection analysis are mostly about human-machine interaction exploring cognitive compatibility between human and machine in various forms. The studies of human-machine interaction are to establish theoretical and physical continuum between human and machine intelligence.

Figure 7.

Visualization of Interdisciplinary Journals Identified through Community Detection

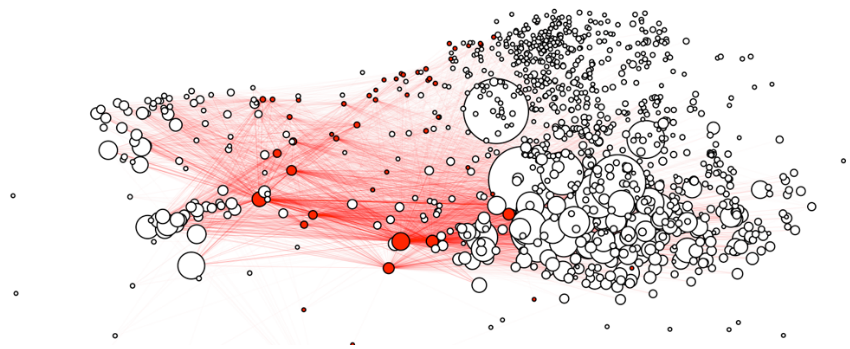


Table 4.

*List of Interdisciplinary Journals in the AI Studies*

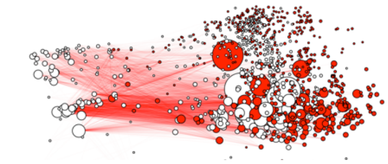
| | Journal | | Journal |
|---|---|---|---|
| 1 | Advances In Human-Computer Interaction | 26 | Human Computer Interaction With Mobile Devices |
| 2 | International Journal Of Human-Computer Interaction | 27 | Visual Interfaces To Digital Libraries |
| 3 | Human-Computer Interaction | 28 | Cognitive Technology: Instruments Of Mind, Proceedings |
| 4 | Behavior & Information Technology | 29 | Engineering For Human-Computer Interaction |
| 5 | International Journal Of System Dynamics Applications | 30 | Haptic Human-Computer Interaction, Proceedings |
| 6 | Journal Of Human-Robot Interaction | 31 | Logic Based Program Synthesis And Transformation |
| 7 | Interacting With Computers | 32 | Human Error And System Design And Management |
| 8 | Entertainment Computing | 33 | Intelligent Tutoring Systems, Proceedings |
| 9 | International Journal Of Mobile Human Computer Interaction | 34 | Cooperative Buildings: Integrating Information, Organization, And Architecture |
| 10 | International Journal Of Social Robotics | 35 | Defense Applications Of Multi-Agent Systems |
| 11 | International Journal Of Human-Computer Studies | 36 | Problem-Solving Methods: Understanding, Description, Development |
| 12 | Machine Translation | 37 | Haptic And Audio Interaction Design, Proceedings |
| 13 | Cognitive Systems Research | 38 | Constraint Solving And Language Processing |
| 14 | AI & Society | 39 | Engineering Human Computer Interaction And Interactive Systems |
| 15 | Universal Access In The Information Society | 40 | Intelligent Media Technology For Communicative Intelligence |
| 16 | Journal Of Usability Studies | 41 | User Modeling 2005, Proceedings |
| 17 | Annual Review Of Cybertherapy And Telemedicine | 42 | Affective Dialogue Systems, Proceedings |
| 18 | Brain-Computer Interfaces: Lab Experiments To Real-World Applications | 43 | Computer Human Interaction: Proceedings |
| 19 | ACM Transactions On Human-Robot Interaction | 44 | Advances In Knowledge Acquisition |
| 20 | IEEE Systems Man And Cybernetics Magazine | 45 | Multimedia, Hypermedia And Virtual Reality: Models, Systems |
| 21 | Smart Graphics, Proceedings | 46 | Lessons From Learning |
| 22 | Spatial Cognition Iv, Reasoning, Action | 47 | Networking: Connecting Workers In And Between Organizations |
| 23 | Cognitive Brain Research | 48 | International Journal Of Man-Machine Studies |
| 24 | Comparative Evaluation Of Multilingual Information Access Systems | 49 | Engineering Cybernetics |

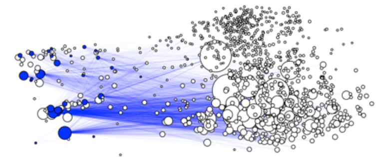
**Identifying Sub-Clusters In Each Field.** The community detection created sub-clusters in each AI and educational psychology studies. The Louvain cluster analysis detected five different sub-clusters in the AI field. The identified clusters in the AI were *symbolic AI, neural network, image processing, robotics*, and *soft computing,* while the clusters in the educational psychology were *educational psychology, educational measurement, child development,* and *learning science studies*. The cluster names reflect the dominant number of journal names in each cluster. For instance, the symbolic AI cluster included *Expert Systems with Application, Kybernetics, Argument and Computation*, and *Cybernetics and Systems*. Also the child development cluster has *Child Development, Journal of Early Intervention*, and *Behavioral Disorder*. The figure 8 shows the interdisciplinary edges from the sub-cluster in each discipline. It indicates that the symbolic AI and neural network studies take most of the interdisciplinary edges in the AI side, while the educational psychology and educational measurement studies are two clusters having most interdisciplinary edges on the educational psychology side. Figure 9 presents the change in interdisciplinary coupling proportion of the sub-clusters in each field. The most visible trend is that of the increasing proportion of symbolic AI and child development studies in the AI and educational psychology studies.

Figure 8.

*Interdisciplinary Coupling Edges of the Sub-Disciplines*

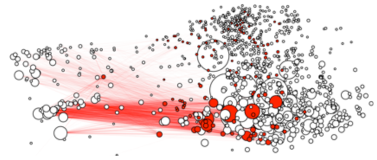
1961-2019 Symbolic AI 1961-2019 Educational Psychology

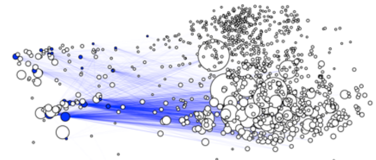




1961-2019 Neural Network

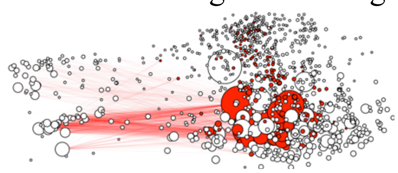
1961-2019 Educational Measurement

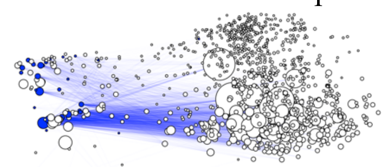




1961-2019 Image Processing

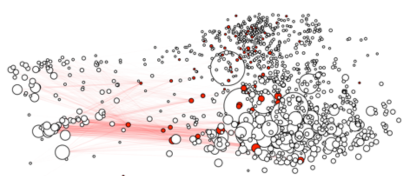
1961-2019 Child Development

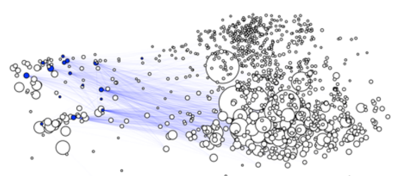




1961-2019 Robotics

1961-2019 Learning Science (Reading and Writing)





1961-2019 Soft Computing

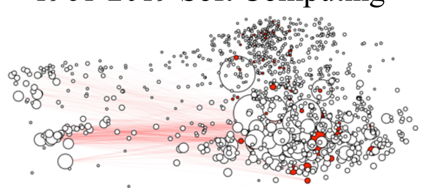
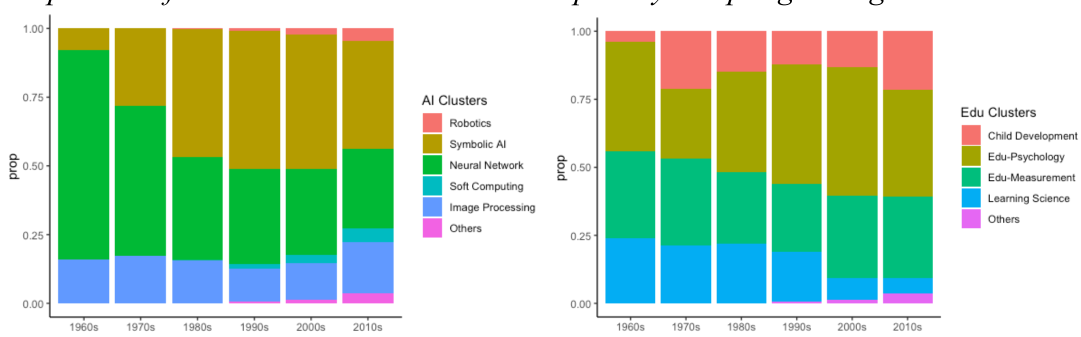


Figure 9.

*Proportion of Each Clusters in the Interdisciplinary Coupling Strength*



The quantitative analysis so far, adopting bibliographic network analysis method, identified the 'cyborg space' in the bibliographic network space between educational psychology and AI studies. There were four major findings. First, the findings confirmed the existence of a common source of knowledge between the two fields, the bibliographic coupling strength, which in turn was based on the number of shared cited references. The network visualization showed a non-trivial amount of interdisciplinary coupling strength between educational psychology and AI studies. Second, the Jaccard similarity index showed that the strength of the tie has slightly increased over time or at least maintained at the same level. Also, the inter-intra coupling ratio indicated that as the AI field expands rapidly, the educational psychology studies tend to be disproportionately affected by the AI studies. Third, the community detection analysis identified the interdisciplinary journals categorized together with the educational psychology journals, while belonging to the AI field according to the WoS category. These are mostly the studies of human-machine interaction and interface. Fourth, the community detection of each discipline identified the sub-disciplinary studies in the educational psychology and AI studies respectively. It showed that the interdisciplinary coupling has been largely led by symbolic AI and neural networks in the AI, and educational psychology and educational measurement studies in the education field.

# CHAPTER 6. QUALITATIVE ANALYSIS: METAPHOR ANALYSIS

## Methodological Background

### *Metaphor In Science Studies*

Over the last several decades, scholars highlighted the ideological and reproductive function of metaphor and found the real force framing our cognitive function and socio-ideology in the metaphorical usage of language (Goatly, 2007; Hasse, 1988; Lakoff & Johnson, 1980). This perspective contrasts with the existing perception of the metaphor considering it as redundant, noisy, and decorative. Hasse (1988) argued that the importance of metaphorical use of language had been largely ignored, given a widespread belief that analytic and scientific description should be based on stable, reliable, and univocal expressions only consisting of literal statements. Against this taken-for-granted assumption, she insisted that "all language is metaphorical" (p. 1). She understood our language is a network system, a semantic web interwoven with its relationally to each other based on similarity and difference.

This cognitive linguistic perspective on the metaphor has significant implications for the scientific studies as scientific discovery is mediated only through the language we use. Concerning this, Van Lunteren (2018) said: "Nature never speaks for itself. Classifications, analogies, technical vocabularies, and other conceptual tools that enable us to make sense of natural phenomena are as much of our own making as the instruments that we use to interrogate nature" (p. 763). Again, Hasse (2000) insisted that the scientific model is a metaphorical expression in its nature. For instance, DNA model built with a colored ball with its connected network is itself a metaphorical expression. By describing the DNA structure with this network system of colored balls, people tend to understand DNA as a connected system of distinctive molecules. Meanwhile, this metaphor tends to hide the features of size, weight, and shape of DNA as a consequence of using the specific kind of metaphorical expression. As hinted in Hasse's point that metaphorical expression has an emancipating force, the metaphor provides momentum to shift the epistemic and ontological base of scientific studies by breaking the existing semantic web in scientific thinking. This emancipatory aspect of metaphor enables scientists to create new theories and models such as string theory and orbital models, which is not achievable through literal expressions (Colburn & Shute, 2008).

There are more studies exploring the relationship between science and technology, and metaphorical expressions. Colburn and Shute (2008) examined how the technological invention of the computer was first conceptualized based on the preexisting similarity with the real world, emerging similarity, and finally enforcing similarity as an effect of its creation. Greenwood and Bonner (2008) criticized the established view that scientific theory should be built upon "literal and precise" language, insisting that: "Metaphorical expressions constitute at least temporarily an irreplaceable element of the linguistic machinery of a scientific theory. These are metaphors which researchers employ to express theoretical claims for which no adequate literal paraphrase is known" (p. 160). They suggested an example of a computer, which does not have any established literal paraphrases that can capture the essential concept of it except for the metaphor of the human brain and mind. Struik et al. (2008) examined how the new field of plant neurobiology emerged by borrowing concepts and terms from animal physiology. They found that the scholars of this new field tried to explain a complex molecular phenomenon in plants using metaphors of an animal mechanism such as synapses, neurons, and the brain system.

Donna Haraway (1972) was one of the earliest thinkers who recognized a significant implication of a metaphor for the science activities. In her doctoral dissertation, she traced the history of how the new paradigm emerged when developmental biology first appeared. To explain the transition in intellectual history, she borrowed the concept of Thomas Kuhn's paradigm shift. In her dissertation research, Haraway built on the idea that Kuhn's indication of paradigm is mainly constituted by two factors: one is a member of communities and their interaction, and two is shared symbols and concepts amongst the members, which can be mediated through metaphorical expressions. She considered that metaphoric systems are "the core of structural coherence" of science, unlike the widely held assumption that objective and literal description is the gist of scientific statements. She found that at the moment of the paradigm shift in biology, "a major reorientation of fundamental metaphor occurred, leading workers in a field to see new problems and accept radically different sorts of explanations" (p. 8). More specifically, she revealed that new generational scholars of biology, who created a new sub-discipline of developmental biology, abandoned mechanical and vitalism metaphors and adopted organic metaphors as a new conceptual anchor.

**Method**

Max Black and Mary Hesse understood scientific modeling and theory building as a metaphorical process (Hesse, 2000). Black said the discourse of scientific modeling is constituted by primary and secondary systems. He understood through the metaphorical discourse the primary system, literal meaning of the words, and associated with the secondary system, metaphorical meaning of the words. For instance, Black analyzed a statement “Life is a journey” as consisting of the primary system of life and secondary system of journey. Black said the primary system of life and secondary system of the journey come together to reframe our conventional view of life with a filter of the journey (Black, 1993). However, the metaphor is not a binary system of the primary and secondary system. Black and Hasse emphasized the interaction effect of the two systems creating a new space of meaning, which Black named as “associated common place.” Hasse pointed out when we just say “Man is a wolf,” not only the meaning of man but also that of wolf transform, resulting in the idea that wolves become more human as much as humans become more wolves. In this interaction, Hesse insisted, neither system can be considered as a genuine truth-bearer (Hesse, 2000).

Building on the metaphor analysis approach undertaken by Black and Hasse, I conducted the metaphor analysis to identify the primary system, secondary system, and associated common place in the identified key shared references across educational psychology and AI studies. Firstly, this study identified specific metaphor expressions related to human and machine at the micro level, such as phrase, sentence, and paragraph. This approach targets only very explicit analogical statements or expressions. Secondly, I identified the metaphors more broadly analyzed at the document level to see how the primary and secondary systems associate with each other to create a new common place where machine and human are interconnected, compatible, and even non-discernable anymore. This method analyzed implicit and abstract indication of metaphorical connection between human and machine, using statistical, mathematical, and mechanical assumptions and expressions. Then, finally, as Hesse stated, this study synthesized the analysis to interpret the metaphorical association between human and machine as an action to create a new world view, set a norm, and corresponding perspective (Hesse, 2000).

*Sampling*

This research identified key shared references between AI and educational psychology studies from 1961 to 2019. There is no established consensus on how to measure the popularity across two or more than two different groups. Giving a combined score of popularity in the network is what degree centrality sometimes means. The degree centrality is a simple count of edges for each node and the nodes with large numbers of edges are ranked at the top of the table. However, in a complex network with many different modules, cliques, and clusters inside, degree centrality does not properly represent global popularity of the nodes. For instance, although one of the references was cited 100 times in AI studies, its global popularity will be low in general once it was referred 0 times in educational psychology. Meanwhile, although one of the references was cited only 50 times in AI studies, its global popularity can be higher in general if it was cited 50 times in educational psychology. Thus, balancing between global popularity, the popularity in more than one cluster, and local popularity, the popularity just in one cluster is a challenging issue to identify the key document in the complex network.

Therefore, this research extracted top-20 most cited references both from educational psychology and AI studies based on the multiplication score of the citation with a certain threshold. The multiplication method is calculating a citation score just simply multiplying the number of citations from educational psychology and AI studies for each reference. For instance, in table 5, Cohen, J. (1988) is at the top of the ranking table because its multiplication score, multiplying citations received from AI (246) and educational psychology (2426), is 596,796. The strength of multiplication compared with the simple addition is that it can filter out the references mostly cited only from one discipline. Additionally, this research considered the citation count of less than 50 from one discipline is not meaningful, setting 50 as a threshold to filter some references having disproportional citation counts. This filtering eliminated references having 5,000 from AI and 30 from educational psychology, to give an example, leaving only the references having a lesser gap. The total number of cited references in the collected bibliographic data from AI and educational psychology studies in the WoS was initially 5,220,428. Then, the reference count cited at least once both from the two studies was 28,066, and after deleting references to less than 50 citations at least from one discipline, the count came down to 122. The top 20 cited references were identified in this list.

This research also identified the top-20 most cited authors both from AI and educational psychology studies to complement the cited reference list. This list showed the authors having published papers that have been consistently cited by both fields. Then, this research identified papers written by the top authors: Cohen (1988), Dempster et al., (1977), Vygotsky (1978), Nunnally (1978), Newell and Simon (1972), Bandura (1986), McClelland et al. (1986), Cronbach (1951), Gibson (1979), and Anderson (1996). From this shortlist, the decision was made to analyze Dempster et al., (1977), and Nunnally (1978) among the statistical papers, and Newell and Simon (1972), Anderson

(1983), Gibson (1979), McClelland et al. (1986), and Vygotsky (1978) among the psychology papers. The papers by Cohen (1988), Cronbach (1951), and Dempster et al. (1977) were excluded because their statistical theory now became a generic grammar of all statistical studies, thus not representing unique characteristics of the interdisciplinary space between educational psychology and AI. In the meantime, for the purposes of this study, I decided to select one statistical papers: Nunnally's *An overview of psychological measurement*. Nunnally's (1978) paper introduced a statistical approach in the general science for the psychological and behavioral science studies. Also, Newell and Simon (1972) and McClelland et al. (1986) were selected given the influence of the authors in the symbolic and connectionist AI studies, respectively. This study picked up Anderson (1983) due to his wide contribution in the AI studies across symbolic and connectionist AI. This research considered the unique position of Gibson's ecological psychology refuting both behaviorism and cognitivism and also its influence on neuroscience and robotics (Lobo et al., 2018).

Table 5.

*List of Top 20 Most Cited Reference both from the Educational Psychology and AI Studies*

| Rank | Cited References | Category | Citation from EDU | Citation from AI | Multiplication |
|---|---|---|---|---|---|
| 1 | Cohen, J. (1988). *Statistical Power Analysis For The Behavioral Sciences*. Academic press. | Stat. | 2,523 | 264 | 666,072 |
| 2 | Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society: Series B (Methodological)*, *39*(1), 1-22. | Stat. | 141 | 2,892 | 407,772 |
| 3 | Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard university press. | Psy. | 1,310 | 185 | 242,350 |
| 4 | Nunnally, J. C. (1978). An overview of psychological measurement. *Clinical diagnosis of mental disorders*, 97-146. | Stat. | 611 | 300 | 183,300 |

| | | | | | |
|---|---|---|---|---|---|
| 5 | Newell, A., & Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, NJ: Prentice-hall. | Psy. | 311 | 539 | 167,629 |
| 6 | Bandura, A. (1986). *Social Foundations of Thought and Action: A Social Cognitive Theory*. Englewood Cliffs, NJ: Prentice Hall. | Psy. | 1,084 | 151 | 163,684 |
| 7 | McClelland, J. L., Rumelhart, D. E., & PDP Research Group. (1986). Parallel distributed processing (Vol. 2). Cambridge, MA: MIT press. | Psy. | 61 | 2,175 | 132,675 |
| 8 | Bandura, A. (1997). *Self-efficacy: The exercise of control.* W H Freeman/Times Books/ Henry Holt & Co. | Psy. | 1,103 | 86 | 94,858 |
| 9 | Schank, R. C., & Abelson, R. P. (1977). Scripts, Plans, Goals, and Understanding. Hillsdale. *U.: Laurence Erlbaum* | Psy. | 243 | 331 | 80,433 |
| 10 | Hebb, D. O. (1949). The first stage of perception: Growth of the assembly. *The Organization of Behavior*, *4*, 60-78. | Psy. | 68 | 1,054 | 71,672 |
| 11 | Glaser, B., & Strauss, A. (1967). The Discovery of Grounded Theory: Strategies for Qualitative Research | | 513 | 132 | 67,716 |
| 12 | Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press. | Psy. | 448 | 125 | 56,000 |
| 13 | Deci, E. L., & Ryan, R. M. (1985). Intrinsic Motivation and Self-Determination in Human Behavior. Berlin: Springer Science & Business Media. | Psy. | 710 | 77 | 54,670 |
| 14 | Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. Psychometrika, 16, 297-334 | Stat. | 470 | 113 | 53,110 |
| 15 | Gibson, J. J. (1979). *The ecological approach to visual perception.* Houghton, Mifflin and Company. | Psy. | 109 | 484 | 52,756 |
| 16 | Bollen, K.A. (1989). *Structural Equations with Latent Variables*. John Wiley and Sons, Inc., New York. | Stat. | 610 | 84 | 51,240 |
| 17 | Johnson-Laird, P. N. (1983). Mental models: Towards a cognitive science of language, inference, and consciousness (No. 6). Harvard University Press. | Psy. | 257 | 193 | 49,601 |
| 18 | Efron, B., & Tibshirani, R. J. (1994). *An introduction to the bootstrap*. CRC press. | Stat. | 96 | 487 | 46,752 |
| 19 | Anderson, J. R. (1996). The architecture of cognition (Vol. 5). Psychology Press. | Psy. | 208 | 212 | 44,096 |
| 20 | American Psychiatric Association, A. P. (1994). Diagnostic and statistical manual of mental disorders (DSM-IV). | Psy. | 752 | 55 | 41,360 |

*Note.* EDU and AI stand for educational psychology and artificial intelligence studies respectively, and the papers whose authors also appear in the list of top 20 cited authors were highlighted.

Table 6.

*List of Top 20 Most Cited Authors both from the Educational Psychology and AI Studies*

| Rank | Authors | EDU | AI | Multiplication |
|---|---|---|---|---|
| 1 | Holland J | 624 | 3,165 | 1,974,960 |
| 2 | Cohen J | 4,345 | 451 | 1,959,595 |
| 3 | Piaget J | 2,593 | 508 | 1,317,244 |
| 4 | Rumelhart D | 545 | 2,349 | 1,280,205 |
| 5 | Bandura A | 3,137 | 358 | 1,123,046 |
| 6 | Anderson J | 821 | 1,035 | 849,735 |
| 7 | Vygotsky L | 2,007 | 298 | 598,086 |
| 8 | Newell A | 457 | 1,261 | 576,277 |
| 9 | Nunnally J | 1,268 | 445 | 564,260 |
| 10 | Dempster A | 144 | 3123 | 449,712 |
| 11 | Carroll J | 930 | 284 | 264,120 |
| 12 | Hair J | 432 | 595 | 257,040 |
| 13 | Bruner J | 1,144 | 219 | 250,536 |
| 14 | Cronbach L | 1,396 | 169 | 235,924 |
| 15 | Gibson J | 228 | 1,031 | 235,068 |
| 16 | Chomsky N | 415 | 510 | 211,650 |
| 17 | Csikszentmihalyi M | 559 | 350 | 195,650 |
| 18 | Sternberg R | 1,103 | 177 | 195,231 |
| 19 | Ekman P | 200 | 927 | 185,400 |
| 20 | Minsky M | 118 | 1,505 | 177,590 |

*Note.* EDU and AI stand for educational psychology and artificial intelligence studies respectively, and the authors whose papers also appear in the list of top 20 cited references were highlighted.

*Analysis*

For the metaphor analysis, different analysis strategies were used depending on the level of explicitness of the metaphorical expressions. First, I searched for key terms indicating human and machine per se in every paper to identify human-machine metaphorical expressions at the sentence and paragraph level. Other than human and machine, I used man, intelligence, brain, cognition, psychology, mind, and mental to indicate *human*, while adopting computer, information processor, information processing machine to represent *machine*. This method was useful to identify straightforward human-machine metaphors. For instance, saying that human is a system constituted with a variety of functions or substructures belongs to this case. Second, I also identified expressions describing humans as a battery of machine-like functionaries with a formalism although there is no direct statement of human-machine metaphors. For this, the study explored sentences and paragraphs including terms such as function, structure, system, operation, classification, probabilistic, and statistical to see how such mechanical expressions describe human, human mind, and human behavior. Lastly, the study also examined abstract concepts and models applicable across the human and machine boundary, thus ensuring comparability, compatibility, generalizability, and universality of the model. In particular, the statistical papers include mathematical expressions with their explanation, making it difficult to do the metaphor analysis in a straightforward way. Hence, this research interpreted abstract mathematics in the statistical papers with a focus on how the papers created cyborg space by claiming statistics and measurement as a universal theory.

Table 7.

*Research Framework for the Metaphor Analysis*

| Explicitness | Primary-Secondary | Search Keyword |
|---|---|---|
| 1. Explicit metaphor | Primary | human, man, men, intelligence, brain, cognition, psychology, mind, mental |

| | | |
|---|---|---|
| | Secondary | machine, computer, information processor, information processing machine |
| 2. Implicit metaphor | Primary | human, man, men, intelligence, brain, cognition, psychology, mind, mental, machine, computer, information processor, information processing machine |
| | Secondary | system, function |
| 3. Abstract metaphor: Modeling and theorization | Primary | human, man, men, intelligence, brain, cognition, psychology, mind, mental, machine, computer, information processor, information processing machine |
| | Secondary | Statistical and probabilistic modeling of both human and machine process |

**Findings**

This research examined the explicit metaphor, using direct expressions: *a human is a machine*, or *a human is like a machine*. Such metaphorical expressions created a cyborg space by alluding to either human as a mechanic being and vice versa. The associated common place created by the metaphor between humans and machines is a cyborg space where human and machine features are considered to be a continuum of each other, and one of the branches shares the same root. Such straightforward expressions were mostly present in the psychological study papers of Newell and Simon's (1972) *Human problem solving* and Anderson's (1996) *The architecture of cognition*.

*Human as an Information Processor*

Newell and Simon (1972) explicitly mentioned that the human could be framed as an information processor and such a metaphorical understanding well explains the real cognitive process: "This study is concerned with thinking-or that subspecies of it called problem solving- but it approaches the subject in a definite way. It asserts specifically that thinking can be explained by means of an information processing theory…The present theory views a human as a processor of information" (p. 5). They insisted that this metaphor provides an opportunity to understand human thinking in a more sophisticated way. They even claimed that this is not a metaphor anymore at all, but a "precise symbolic modeling" (p. 5), reflecting an exact mechanism of human's thinking process.

The metaphor describing humans as a continuum of machines has triggered thinking that humans and machines share the same problem space. The problem space is an inventory of tasks that the intelligent agent should navigate through problem-solving competence. One of the reasons why these authors could assert that human thinking resembles information processing was that they focused on the similarity between human and machine problem space, assuming that humans and machines recognize and solve the real-world problems using the same logic. Newell and Simon (1972) wrote:

> One enters a department store: "Where do I find men's suits?" "Third floor, down the center aisle and to your right"; "Thank you"; and off one goes, following directions. Several phenomena here are closely allied to the interests of this book: language production and reception; deciding to ask for information to solve a problem; following directions, once assimilated; perhaps (if the directions were imperfect) solving some smallish subproblems along the way. Certainly this is the behavior of an information processing system. (p. 7)

They assumed the problem space to be contingent and hierarchical, where the tasks are an assembly of smaller problem-solving tasks. Although there is a difference in intelligence level between human and machine or human and animal, they saw that the essence of the thinking process can be abstracted to the universal information processing theory. Thus, they insisted that the task environment of artificial intelligence provides a meaningful insight into understanding human thinking process, assuming compatibility between human and machine thinking process:

> As will become clear, a theory of the psychology of problem solving requires not only good task analyses but also an inventory of possible problem solving mechanisms from which one can surmise what actual mechanisms are being used by humans. Thus, one must work with task environment in which artificial intelligence has provided the requisite array of plausible mechanisms. (Newell & Simon, 1972, p. 6)

Similar thinking appeared in Anderson's work (1996). He insisted on the new unitary cognitive system theory model, arguing for universal law governing seemingly fractured cognitive subsystems. He tried to validate his model because the computer system runs on a single set of principles but has unlimited functionality. He considered the computer simulation of the human cognitive model as an essential step to validate the cognitive theory, and he performed several simulation studies to prove his mental theory for himself. He did not use the term information processing machine. However, he claimed that human cognition is for the computational function, saying: "More generally, I claim that the cognitive system has evolved many different representational types, each intended to facilitate a certain type of computation"(Anderson, 1996, p. 26). He also continued to state that the evolution process of computer and human cognition follows a similar trace because they both follow the same universal logic of computation: "Thus we see that computers have developed distinctive data types in response to some of the same pressures to produce intelligent behavior that humans have felt during their evolution. This is evidence, or a sort, that humans would have evolved the capacity to process different data types" (Anderson, 1996, p.26).

In this unitary problem space common for humans and machines, performing problem-solving tasks, a sign of having human-like intelligence, is not the privileged human ability anymore. Now, whether being called an information processor or computer can solve the problem, thus having an intelligence. It means a universal theory of intelligence under the frame of information processing and computation as Newell and Simon (1972) said:

> How can a problem be solved? What makes a problem difficult? The answers to these questions constitute a theory of problem solving, one that should be applicable to man, beast, and machine alike, insofar as they can be represented as information processing systems of the type posited in Chapter 2. (p.87)

McClelland et al. (1986) also clearly explained that their parallel distributed processing (PDP) model, which was supposed to challenge the Newell and Simon's symbolic model, commonly works for human cognition and computer. Although the PDP model was devised to compete with the symbolic model, which insisted on a sequential process of information through logical symbol manipulation, both shared the information processing theory as a common paradigm. McClelland et al. (1986) said:

> They (PDP models) hold out the hope of offering computationally sufficient and psychologically accurate mechanistic accounts of the phenomena of human cognition which have eluded successful explication in conventional computational

formalisms; and they have radically altered the way we think about the time-course of processing, the nature of representation, and the mechanisms of learning. (p. 11)

However, unlike the other authors in the selected papers, Gibson (1979) opposed describing a human as a machine. He was explicitly against this idea, claiming that humans are structured differently compared to the machinic assemblage. He also opposed using the information processing theory in modeling the human perceptual system, insisting that Claude Shannon's information processing theory failed to explain the human perceptual process. He argued that the information processing theory is only for machines but not for humans. Nevertheless, such an opposition to the information processing theory should not be interpreted as an opposition to the metaphor connecting humans and machines in general. In the following part, this research explained why Gibson was not an exception to the general paradigm describing humans as machines and vice versa.

*Human as a System*

Sometimes, the authors used the terms system, function, and structure to describe human and human features instead of the machine. These terms made the difference between the meaning of human and machine relatively unnoticed. These terms reduce friction between the two that can make the human-machine metaphor look unnatural and unfitted. For instance, behaviorism drew criticism for its direct analogy to a machine in explaining human features. The system was an alternative metaphor or modeling of human's psychological process to avoid such confrontation. Gibson (1979) said: "Neither mentalism on the one hand nor conditioned response behaviorism on the other is good enough. What psychology needs is the kind of thinking that is beginning to be attempted in what is loosely called systems theory" (p. xv). It shows that he tried to introduce the term *system*, a loosely defined notion, to imply human mental process is neither a mystic mentalist's nor a mechanistic behaviorist's model, but somewhere between them.

Similarly, the selected papers' authors posit that system, function, and structure are neutral, comprehensive, and ambivalent concepts, which represent overarching human and machine features under the more universal and scientific framework. The following quotation is a typical example of how these terms, system, and function were used to articulate a theory:

> An information processing theory is dynamic, not in the sense in which that term is used in depth psychology, but in the sense of describing the change in a system through time. Such a theory describes the time course of behavior, characterizing each new act as a function of the immediately preceding state of the organism and of its environment. (Newell & Simon, 1972, p. 11)

The authors of the selected papers often described humans as a system to imply that humans are constituted by substructure and act as sub-functionaries like a system. Newell and Simon (1972) mentioned: "He (human) is a system consisting of parts sensory subsystems, memory, effectors, arousal subsystems, and so on" (p. 3). Also, they assumed the information processing system to be an assemblage of receptors, effectors, processors, and memory. Each subcomponent has a designated function such as accept input, produce output, and store information. Then, the information processing, the system-level function, combines these small components and their function. Anderson

(1996) similarly posited the unitary cognitive system as a hierarchical composition consisting of lower-level declarative, production, and working memory systems. Each memory system performed a function of storage, retrieval, match, execution, application, encoding, and performances to work as a higher level unitary cognitive system. Even Vygotsky (1978) was not an exception to using a human-system metaphor. He framed child as a system consisting of the sub-functionaries, saying:

> The linkage between tool use and speech affects several psychological functions, in particular perception, sensory-motor operations, and attention, each of which is part of a dynamic system of behavior. Experimental-developmental research indicates that the connections and relations among functions constitute systems that change as radically in the course of a child's development as do the individual functions themselves. (Vygotsky, 1978, p. 31)

McClelland et al. (1986) also alluded to hierarchical structure of the cognitive system: "What PDP models do is describe the internal structure of the larger units, just as subatomic physics describes the internal structure of the atoms that form the constituents of larger units of chemical structure" (p. 12). Gibson (1979) also defined a perceptual system as "an organ and its adjustments at a given level of functioning, subordinate or superordinate" (p.234). He also mentioned: "(perceptual capacity of organism) lie in systems with nested functions" (Gibson, 1979, p. 195). He used the term *system* with a slightly different nuance to represent relational and combinatorial function of perceptual organs, rejecting a model that human perceptual capacity is located in "discrete anatomical parts of the body" (Gibson, 1979, p. 195). His view was truly system-oriented as he insisted that all human organs belong to a certain system, do not exist being separated from the system. In this regard, Gibson (1979) often connected organ names with a hyphen such as "head-eye system" (p. 195), "eye-head-brain-body system" (p. 55), and "mustle-joint-skin system" (p. 177) to emphasize the relational character of the human perceptual system.

The replacement of humans or machines with a more abstracted notion of systems had an immediate impact, generalizing uniquely human or machine features onto the other beings without much disturbance. The authors frequently used the term by creating many phrases such as artificial intelligence system, information processing system, recognition system, production system, linguistic system, problem-solving system, and adaptive system to indicate a blurring boundary between humans and machines alike and with the elusive hybrid entities. The following quotation is one of the examples showing how the boundary was crossed and finally erased.

> …the emerging system is remarkably content-free, and without the powers of integrated action shown by the normal adult. Many constraints on the nature of the fully developed system arise from the requirement of self-organization – help from the external environment can only be used after the system has developed itself to a point where it is capable of such assimilation. (p. 7)

Newell and Simon (1972) described a certain system's feature, and we are soon to recognize that the system indicated humans. *The emerging system* means children or adolescents, *while the fully developed system* implies adults, and this is not the special case where the authors used the term *system* in the place for a human. Newell and Simon (1972) called human as an information processing system in the initial part of their book, but in the latter part, most of the information processing system appeared to mean a computer. Similarly, Anderson (1996) declared that human cognition is a unitary mental system performing adaptive control of thought (ACT) and used this term throughout his book instead of human and human cognition. Gibson (1979) stated: “A system can orient, explore, investigate, adjust, optimize, resonate, extract, and come to an equilibrium, whereas a sense cannot” (p. 234). For Gibson, the system now became an agent of action, movement, and behavior, not an object affected by external stimuli. That is, the system began to be described as an animate being.

In a similar vein, Newell and Simon (1972) espoused the thinking that learning is a general process applicable beyond the human, expanding this uniquely human process to the non-human beings saying: “Learning is a second-order effect. That is, it transforms a system capable of certain performances into a system capable of additional ones” (p. 7). Placing the term *system* in the slot initially assigned to humans could transform human learning as a universal feature applicable to whatever they define as a system. For instance, now there is not much difficulty understanding the following statement: it (learning) transforms an information processing system capable of certain performances into an information processing system capable of additional ones. McClelland, Rumelhart, Hinton, and the connectionist group also suggested a mechanistic definition of learning in their PDP model. They posited that the learning is a general phenomenon reduced to a micro-mechanism. For them, learning is a matter of connection, its strength, activation, and pattern. It seems more like tuning of information flow in the network system. They said:

> For if the knowledge is the strengths of the connections, learning must be a matter of finding the right connection strengths so that the right patterns of activation

> will be produced under the right circumstances…it opens up the possibility that an information processing mechanism could learn, as a result of tuning its connections, to capture the interdependencies between activations that it is exposed to in the course of processing. (McClelland et al., 1986, p. 32)

The authors could transfer the meaning and function of the primary system, in this case, human, to the secondary system, machines - or vice versa - by setting compatibility and equivalence between human and machine through the term system or its mechanistic depiction. Thus, in this hybrid space, humans and machines commonly represent systems only in a different kind. Now that they are derived from the same root, they are compatible with and applicable to each other.

Vygotsky (1978) also described humans as a system. Unlike the other authors, Vygotsky (1978) described the human as a system that interacted with the other systems. Vygotsky (1978) saw that human is nested in the social and behavioral system: “From the very first days of the child’s development his activities acquire a meaning of their own in a system of social behavior and, being directed towards a definite purpose, are refracted through the prism of the child’s environment” (p. 30). Also, he insisted that humans can build up their own psychological and behavioral system through the interaction with the external environment: “…the child’s system of activity is determined at each specific stage both by the child’s degree of organic development and by his or her degree of master in the use of tools” (p. 21).

*Human as a Statistical and Probabilistic Being*

The human-machine-system metaphor barely appeared in the statistical papers. The mechanistic metaphor of human and machine was hardly visible and deeply permeated into the general statistical and mathematical modeling. Nevertheless, the statistical and mathematical models are the most powerful metaphors opening up a cyborg space by setting equivalence between heterogeneous beings with quantification. As Haraway (2016) put it: “Human beings, like any other component or subsystem, must be localized in a system architecture whose basic modes of operation are probabilistic, statistical” (p. 32). She also accurately pointed out that under this probabilistic mode of operation, “any component can be interfaced with any other if the proper standard, the proper code, can be constructed for processing signals in a common language” (Haraway, 2016, p. 32). Interestingly enough, there are similar statements by the authors of the selected papers. Some authors mentioned that statistical and mathematical modeling is generally applicable to various cases regardless of their internal and environmental features once they are set to be equivalent.

Nunnally (1978) claimed that quantitative measurement is the universal representation of the phenomenon. He suggested “generality of measurement problems” (Nunnally, 1978, p. 97), arguing that psychometric modeling and measurement is applicable to the other fields of studies regardless of their specificity and particularity of topics and objects of analysis. Based on such generality, he also predicted that the psychometric measurement would be largely adaptable to the physiological studies studying the human brain and biological organs, and also clinical diagnostic studies of mental disorder:

> I have found similar principles to apply in an extremely wide variety of scientific issues in psychology, psychiatry, numerous fields of medicine, and law, and in special issues in the physical sciences and engineering, particularly biomedical engineering. Indeed, I have been surprised at the commonality of issues regarding psychological measurement that runs through these various disciplines. (Nunnally, 1978, p. 98)

Nunnally also insisted that whatever can be reduced into unit elements is quantifiable, thus being an object of quantitative measurement and analysis. Once such assumption is accepted as stylized fact, it can provide a basic condition to eliminate the boundary between the human and non-human object as Newell and Simon (1972) did in their model, converting human into mechanical information processing system:

> Measurement consists of rules for assigning numbers to objects in such a way as to represent quantities of attributes. No matter how one "cuts it," eventually measurement results in attaching a number to people (or material objects) in such a way as to describe the extent to which a particular characteristic is present. (p. 101)

Although this is just an assumption posited to create a model, this hypothetical thesis putting that psychological phenomenon is statistically measurable and analyzable can influence people's belief in the real world. Nunally (1978) pointed out: "...(for mathematicians) the normal distribution of intelligence was an empirical fact and psychologists thought that it was a mathematical fact" (Nunnally, 1978, p. 103). It means that psychologists believe that the statistical representation of human intelligence is a mathematical validation of their psychological model, while statisticians believe it as an empirically proven fact - cross-validating each other, but less grounded on hard fact. This cross-validation creates a kind of *associated common space* between two different models that Hasse and Black insisted in their metaphor theory. This space co-created by statistical and empirical psychology studies is a metaphorical space where the two different semantic elements interact with each other to create new meaning.

The psychological papers posited that the information processing system is based on statistical and mathematical mechanisms, although they did not adopt a statistical model to build up their theory. Newell and Simon (1972) did not explicitly suggest statistical or mathematical models of the information processing theory while concentrating on symbolic logic. However, they mentioned that the human and machine functionaries are mathematical representations or mathematical calculation objects. For instance, Newell and Simon (1972) described human learning as a "formalized systems of mathematics" (p. 105), reducing the whole process to mathematical operations among axioms. They also emphasized that human learning and even organism itself can be translated into probabilistic event, saying:

> The mathematization of learning theory in the last decade shows this very well (Atkinson, Bower, and Crothers, 1965). In the prototype version of mathematical learning theories, the organism is represented by a set of probabilities of occurrence of a fixed set of responses: learning involves changes in these probabilities under the impact of experience. (Newell & Simon, 1972, p. 8)

Anderson (1996) knew that information processing theory is about statistics and probabilities. He was between symbolists Newell and Simon (1972) and connectionist groups, embracing both symbolic and connectionist concepts. Anderson (1996) framed the information processing system probabilistically, and in particular, he posited the activation and decay of certain functions to depend on the probability equations. For instance, he described the mechanism of the memory system in his ACT system as follows: ”When a temporary link is created and there is not a permanent copy of it, there is probability p that a permanent copy will be created. If there is a permanent copy, its strength will be increased by one unit” (Anderson, 1996, p. 24). It means that creating a permanent copy of the temporary link in the memory system is a probabilistic event. Anderson (1996) also structured the production selection system of the ACT being regulated by the statistical and mathematical formula of *1-es/bn*, which means the probability of completing given test in time is a function of s, “strength of the production” (p. 24), and n, “the number of productions simultaneously being tested” (p. 24). He also posited that the machine's performance could be predicted probabilistically, and even the ACT system’s object of analysis depends on the statistical properties.

The metaphorical understanding framing human psychology as a probabilistic process culminates in the connectionists’ idea. McClelland, Rumelhart, Hinton, and their connectionist group assumed human cognition, decision-making, and its micro-neural system largely follows the statistical mechanism (McClelland et al., 1986). This idea was well reflected in the following description of the language learning process:

> Each active input unit contributes to the net input of each output unit, by an amount and direction (positive or negative) determined by the weight on the connection between the input unit and the output unit . The output units are then turned on or off probabilistically, with the probability increasing with the

> difference between the net input and the threshold, according to the logistic activation function. (p.239)

They also saw that this statistical system, distributing decision-making to the multiple parallel units, produces a more consistent and reliable outcome than the single logic circuit model. They sometimes called their PDP model a competitive learning model because, in their learning model, multiple neurons compete against each other to be selected as an output signal by strengthening their connection weight. In such a system, the intervention should be a reconfiguration or finetuning of the connection weights, for instance, by changing input patterns and a threshold for activation.

# CHAPTER 7. CONCLUSION

## Summary

This research explored the cyborg space, where the educational psychology and AI studies overlap each other in two different ways. First, this research examined the journal-journal bibliographic coupling network to identify cross-referencing and interdisciplinary journals. It illustrated that the journal-journal bibliographic coupling strength between the two fields has increased over the last 60 years with the increased size of each field. The interdisciplinary coupling count, in other words number of shared cited references, increased from 16 in the 1960s to 7,857 in the 2010s. Along with it, the interdisciplinary coupling strength, an abstracted index indicating journal-journal connection strength, grew from 25 to 67,327 during the same period. Then, this study calculated the Jaccard similarity index, the normalized score from 0 to1 indicating relative size of interdisciplinary coupling strength against the total coupling strength of the given network. The Jaccard similarity index slightly increased from 0.015 (1.5%) to 0.022 (2.2%), but the educational psychology and AI studies received differential impact. The inter-intra coupling strength index measures coupling strength with the other disciplinary journals relative to that with the same disciplinary journals. This inter-intra coupling strength of educational psychology and AI studies have evolved in opposite directions. The score increased from 0.015 (1.5%) to 0.264 (26.4%) for the educational psychology studies, while it decreased from 0.714 (71.4%) to 0.023 (2.3%) for the AI studies. This means that the references cited by educational psychology studies have been increasingly also cited by the AI studies growing the interdisciplinary space as much as the 26% of the total journal-journal coupling strength in the educational psychology field.

Then, this research explored this cyborg space using community detection methods. At first, this study detected the community at the entire network level to identify interdisciplinary journals belonging to the communities where the majority of journals are from the different fields. The journals in this heterogeneous community can be considered as interdisciplinary journals because they are part of a different group from their initial membership to a certain field assigned by the WoS platform. As a result of the community detection, this research identified a group of interdisciplinary journals consisting of journals specialized in human-computer interaction. It means that these human-computer interaction studies tend to be a hub in the network bridging educational psychology and AI studies. Semantically, the result of this analysis is acceptable in that the human-computer interaction studies may need knowledge both from human and AI science. Then, as a next step, this study broke down each field of studies into sub-communities again using the community detection methods at each field level. The result showed that the educational psychology studies were comprised of educational psychology, educational measurement, child development, and learning science, while the AI studies were divided into symbolic AI, neural network, image processing, robotics, and soft-computing. The educational psychology and measurement studies took the largest proportion of the interdisciplinary coupling strength in the educational psychology field, and the symbolic AI and neural network were most dominant in the AI field across the time.

Second, this study examined active metaphor with the cyborg space composed of the top 20 most commonly cited references both from educational psychology and AI studies. This study posited that the content summary of these common references represented the characteristics of the cyborg space, where the boundary between studies of human and machine intelligence was blurred. Seven commonly cited publications were selected for the metaphor analysis. The metaphor analysis found that the selected papers described humans as an information processing machine, a system with substructure and probabilistic being. The papers demolished the boundary between human and machine process by depicting human and its psychological process on following a machine-like process composed of systemic and functional parts. Also, the papers built the universal theory applicable to both humans and machines, insisting to follow both the universal law of statistics and mathematics. The papers started redefining learning beyond humans. Newell and Simon (1972) said that the goal of learning is to transform a system doing one task to do another. McClelland et al. (1986) defined learning in terms of strengthening a connection in the neural network in a certain direction. In such mechanical definition, the learning became the universal information process that can happen in whatever the system believed to follow the information processing mechanism.

In summary, this research proved the existence of a cyborg space where the human and machine boundary is ambiguous, with the AI study rapidly expanding its intellectual territory into the field of education science. It alludes to the fact that knowledge specific to human intelligence studies, including education science, has been provincialized and subsumed under a much broader interdisciplinary space of general and system intelligence. Besides, it implies that our educational psychology studies, the biggest field providing scientific rationality for the education studies, are not only the theoretical continuum of the machine studies but even susceptible to the influence from the machine studies, because their knowledge base is not unique anymore. The machine science studies proved that the theories to explain, predict, control, and enhance human intelligence are also applicable to building human-like machine intelligence, becoming a source of cross-validation for the human intelligence theories. That is, machines reflect, explain, and even validate our understanding of humans. This also means that human intelligence, believed to be a privilege of human species differentiating it from the other beings, is provincialized - now it is one of the kinds among multiple intelligence models. There is more to expect in the near future in regard to the integration of human-machine intelligence and growing interdisciplinary space between human and machine studies, which this research findings could not yet identify, but which loom large as the future educational science discourse and paradigm.

Even now, the education field is transforming in the direction of mechanization. Amongst many visionary ideas enlightening education policy and practice, cognitive neuroscience seems to be the potential candidate for the new regime of education science. With the advancement of brain scanning technology such as MRI and fMRI, we can now look inside our brain and capture the image of a decimal point of a moment. It means that we can compare our behavioral patterns with those of electrical impulses in our brain almost synchronously. Also, due to the recent hype and success of the deep neural network in the AI studies, there is emergent scholarly movement to explain human cognitive process with the advanced AI mechanism, reverse-engineering of AI to develop human cognitive models (Griffiths et al., 2015; Lieder & Griffiths, 2020; Wu et al., 2013; Van Gerven, 2017; Zednik & Jakel, 2014). This approach is also called computational rationality, and Gershman et al. (2015) insisted parallelism between the intellectual trajectory of AI and psychology studies, suggesting that Bayesian inferential modeling is the universal learning theory across computers and the brain.

There are already movements toward using brain scans as scientific evidence of educational science to evaluate children's cognitive states or the cognitive impact of education (see McCandliss, 2016; The World Bank, 2018). Martin-Loeches (2015) declared that the technological level already crossed the tipping point, which means that, regardless of whether the opposition might be posed against it, there is no doubt that neuroscience will guide the future educational practice. Ansari and Coch (2006) identified mind, brain, and education (MBE) as newly rising fields in education science. They set the analysis level for the future of education science as follows: test scores, behaviors, systems, networks, neurons, synapses, molecules, and genes. Schneider and Graham (1992) introduced the connectionist approach's pedagogic implication based on the neuronal network mechanism in the human brain. They designed the learning that could ultimately enhance neuronal connection inside the brain to sophisticatedly control the amount of input, repetition, level of difficulty, and enhancement. Therefore, we can expect a more mechanical turn for our future education lying ahead of us.

In the AI field, the most recent boom of machine learning seems to be aloof from the intellectual history interrelating with human intelligence (Kao & Venkatachalam, 2018). However, there some leaders of this field still emphasize cross-referencing and cross-modeling human intelligence. Reinforcement learning, one of the most up-to-date machine learning methods widely adopted, is the behaviorists' machine. In the book *Reinforcement Learning: An Introduction*, Sutton and Barto (2018) addressed Thorndike's S-R theory and his animal experiments in a great detail. They stated that the reinforcement learning algorithm has a clear connection to the behaviorists' psychology (Sutton & Barto, 2018, p.342):

> The algorithms we describe in this book fall into two broad categories: algorithms for prediction and algorithms for control. These categories arise naturally in solution methods for the reinforcement learning problem presented in Chapter 3. In many ways these categories respectively correspond to categories of learning extensively studied by psychologists: classical, or Pavlovian, conditioning and instrumental, or operant, conditioning. These correspondences are not completely accidental because of psychology's influence on reinforcement learning, but they are nevertheless striking because they connect ideas arising from different objectives.

Also, there are growing body of literature advocating consistency and similarity in the human and machine intelligence process (see, Garnelo et al., 2016; Helmstaedter, 2015; Kriegeskorte, 2015; Lake et al., 2016; Somers, 2013). The connectionists led by Demis Hassabis, the Google engineer who developed AlphaGo, published the article titled *Neuroscience-inspired artificial intelligence* (Hassabis et al., 2017). In this article, he said that the connectionism did not abandon the high hopes of exploring the actual human brain function mechanism. He mentioned that there is still immense opportunity to study human intelligence and brain function mainly in two aspects. Firstly, he saw that brain functioning is a source of inspiration to develop new software algorithms as the history of AI vindicated in the case of Turing and McCulloch and Pitts. He claimed that the AI algorithm would be a hybrid between statistical and biological inspiration in the future. Second, he recognized that neuroscientific knowledge is still essential as it can validate the new design of machine learning as a general intelligence system. As Hassabis et al. (2017) predicted, if neuroscience-inspired AI in the future is a dominant architecture, its implication for our society will be non-trivial, reconfiguring our self-knowledge and sustaining our basic assumption constituting various social activities, including education.

**Policy Implications**

Then, what are the policy implications of these research findings? That is, how should we educate our future generation in the world where human intelligence is provincialized, while human-machine intelligence arises? What can be the knowledge and paradigm designing our education while our scientific knowledge of the human mind is subsumed under the AI studies? What are the alternative educational pathways when human intelligence is not the sole choice for the employers and the AI presents its robustness and resilience surpassing that of human. This research will provide some possible answers to these questions. Because it is difficult to connect the findings of this study to new policy design, this research will broadly discuss multiple pathways that we can choose in the future depending on different perspectives.

First, we may choose to compete against the machines, vying for the portion of the problem space. I would call it *compete against the machine* approach. Here, the problem space means the narrowly but neatly defined set of problems and tasks in a mechanical and logical way. Defining the problem space really sets the tone of educational objectives. Once we define specific problem space to entail clear tasks and logical sequence, it is easy to deliver knowledge to the students, but at the expense of that, it also increases accessibility of machines to this problem. The OECD's Program for International Student Assessment (PISA) is a representative case adhering to the competition paradigm, presuming the competitive relationship between humans and machines. The OECD PISA makes the most advanced AI machines take the PISA tests every year and announces the score. Their report warns that students scoring lower than the AI algorithm are at risk of peril in the future, urging educators to push them forward to the safe zone above. The OECD PISA did not change any rules of the game and problem space, but tried to generalize their mechanistic problem space to machine intelligence. We may compete and win against the machines in such circumstances, but the probability is getting low. There are numerous examples where the machines surpass human performance in the narrow problem spaces such as chess, Go, video games, sports games, DNA sequencing, and quiz games. Therefore, it is neither viable nor desirable to sustain education targeting to raise students to become the best performers in the narrowly defined problem space such as tests and artificial subtasks.

Second, we may choose to dominate and govern machines by taking the meta-problem space, where machines cannot enter. The meta-problem space is where we develop AI machines and other systems as a designer of machine intelligence. It is called *build and control the machine approach.* There are substantive policy moves related to this approach. Some governments announced investments in the AI industry to raise AI developers and maintain AI competitiveness. In such a policy paradigm, the governments want to retrain their workforce by directing them into the AI related industry and teach them to be able to design or manipulate machine intelligence. This policy approach is the easiest option to take and even be effective somehow in the short-term period, as many economists predicted the digital transition just started and this boom would continue in the following decades. However, there is a hidden cost here. Firstly, there will be a great divide causing inequality in the job market between those who can and those who cannot manipulate the machines because education systems cannot guarantee all students to be AI scientists. It is probable that the majority of students will remain vulnerable to the automation triggered job crisis. That is, this policy option is not for education *for all,* although it is a realistic educational objective for some engineering and computation majors at the universities. Secondly, enhancing AI workforce may precipitate the automation of industry in general, reducing the labor portion reserved for humans in the longer perspective. Once the AI infrastructure is completed and ready for deployment at a massive scale, there will be even worse job crises.

Third, we can choose to optimize and augment human intelligence by using AI technologies in the conventional problem space. The above two options presumed the competition and hierarchical human-machine relationship, while this approach emphasizes instrumental value of the machine and proactively tries human-machine integration. Thus, I would name it *use the machine* approach. It includes efforts to integrate machines in the individual human process, collective decision-making, and knowledge building process. Such human-machine integration in many different social and economic processes will bring innovation to the society in various dimensions improving productivity, efficiency, and human well-being. In the educational settings, the machine can be more seamlessly integrated into students' learning process to augment students' learning outcome or enhance students' efficacy and agency. The machine will provide an opportunity for humans to maximize performance in the conventional and mechanical problem space compared with the past. With this enhancement, we will be able to reduce students' learning achievement gap, lower the educational operation cost, and promote lifelong learning with the customized personalized AI-mediated learning tools. The AI technologies will augment general efficiency and effectiveness of the educational system, and transform our educational setting dramatically. However, there will be unintended consequences related to this massive AI transition in education, and this new phenomenon will require explanation beyond the instrumental and technological dimensions of human-machine relationships.

Fourth, the previous three approaches all maintained human-machine dichotomy as their basic tenet. Also, these approaches are based on the assumption that automation can be strictly separated from autonomy, thinking that we can exploit machine's automation without yielding autonomy. In contrast, there is another way possible, embracing machines as part of human and societal processes over the dichotomy, coexisting in the conventional problem space. This is what I call *become with the machine approach*, and this approach recognizes that automation of machines inevitably accompanies the growing autonomy of machines. It also requires more humanized machines and more mechanized humans. In the society where this approach is adopted, every part of the social system is reconfigured and redefined to be ready for human-machine integration in various aspects. For instance, there are many social debates in regard to new ethical and legal problems requiring revision to embrace machine intelligence in some decision-making processes. There are data privacy and ownership related issues, as machine learning services are based on big human data. Also, as machines approximate human-like behavior, there are many services launched to provide mental wellness and emotional support through interactive machines. People may feel the mechanization is not only a technical matter, but a more comprehensive socio-ontological issue restructuring social fabric. Education policy heavily reliant on the build the machines and use the machines approach will not be enough to address this restructuring issue. Rather, there is a demand to teach our students how to live *with* these machines. As Sundar (2020) said, humans are situated to negotiate its agency with the machines:

> It may seem ironic that we are turning to machines for limiting machine agency and reclaiming human agency, but it signals an emergent collaboration between humans and machines in negotiating the type and degree of agency. This collaboration rests on a nuanced understanding of the various ways by which machine agency can enhance human agency and the ways by which it may threaten it. (Sundar, 2020)

Such human-machine integration cannot happen with simple pre-defined knowledge or meta-knowledge manipulating machines. The schools will be the places where the students can practice human-machine interaction before entering into the job market or machine-heavy society. The machine subject needs to be addressed not only as machine science, delivering math heavy knowledge, but also as machine literacy with common sense knowledge and more exposure to hands-on user experience.

Lastly, we can redefine the conventional problem space with machines by using this new technological development. It means that we need to consider the machine not just as an instrument, but also as an ontology itself, shifting our worldview through the technology. Such a move may include liberating our worldview from the fixed and single paradigm, while opening up toward pluversality. Thus, it can be named as *beyond the machines approach*. The reason for redefining problem space *with machines* is that, as Haraway once put it, the machine is our non-optional condition framing our mode of life and knowledge production. Totally ignoring machines in redefining a problem space or even redefining the space against the machines is not viable in the state when machines are seamlessly integrated into our mode of existence. Rather, we can use this mechanic transformation as a source of new inspiration. In academic perspective, this approach can create innovative educational theories and ideas because interdisciplinarity with the machine studies can be a source of new ideas and an emergent paradigm. For individuals, outperforming machines as lively as humans lets us reconsider the purpose, meaning, and value of life and intelligence at the foundational level. For the society, the machines give us a chance to reconsider our social, cultural, economic, and political fabric from a different angle. It is another *window of opportunity* to shift our worldview, although it comes with a crisis and chaos. The machines will not just decenter humans; rather, they will challenge people to cross their comfort zone and conventional boundaries, sparking new ideas and perspectives. Humans can expand their functions into the uncharted problem space where new logic, paradigm, and relationality is required.

In such a scenario, we may need to rebuild or radically restructure our educational taxonomy, operational definition, system, and formalism to educate our children to realize our relationality and dependency upon other beings. Our modern education was built upon human exceptionalism, positing that intelligence is a uniquely human possession so that educational purpose should be to achieve high performance in the well-defined intellectual domains. Such an educational paradigm is not tenable anymore as a global model, and does not guarantee job and life satisfaction in our future life. The former World Go champion and the first Go player who played against the machine, Lee Sae Dol, provides an important example that holds implications for us. He retired in

2019, three years after the surprising defeat by the AI machine. He at least tried many different things since his defeat, exploring various alternative options in the post-AlphaGo world. He participated in the Go contest as usual, and even tried to start studying AI technology to make a breakthrough, but to no avail. Nothing could stop him from retirement, which saved him from deep dismay and disappointment. He stayed at the same conventional problem space of Go, while trying to compete against the machine, thus always arriving at the same conclusion: AI is unbeatable. It tells us that it is our educational mandate to raise our children to find new meaning of life and learning in the world where humans are provincialized. For this, we may need to teach our children to recognize that the world cannot be defined by the single order, nor do we need to excel in a single problem space. We need to understand and embrace the reality that machines are not separate from us, and we may need to learn to be with them forever.

## REFERENCES


Abrahamsen & Bechtel (2012), History and core themes, In K. Frankish & W. M. Ramsey (Eds.). *The Cambridge handbook of cognitive science* (pp. 9-28). New York: Cambridge University Press.

Acemoglu. D., & Restrepo. P. (2019). Artificial intelligence, automation, and work. In A. Agrawal, J. Gans, & A. Goldfarb. (Eds.). *The Economics of Artificial Intelligence: An Agenda.* (pp. 197-236). Chicago: The University of Chicago Press.

Agrawal. A, McHale. J., & Oettl. A. (2019). Finding needles in haystacks: Artificial intelligence and recombinant growth. In A. Agrawal, J. Gans, & A. Goldfarb. (Eds.). *The Economics of Artificial Intelligence: An Agenda.* (pp. 149-174). Chicago: The University of Chicago Press.

Anderman, E. M. (2011). Educational psychology in the twenty-first century: Challenges for our community. *Educational Psychologist*, 46(3), 185–196. http://doi.org/10.1080/00461520.2011.587724

Anderson, J. R. (1996). *The architecture of cognition* (Vol. 5). Psychology Press.

Ansari, D., & Coch, D. (2006). Bridges over troubled waters: education and cognitive neuroscience. *Trends in Cognitive Sciences*, 10(4), 146–151. http://doi.org/10.1016/j.tics.2006.02.007

Aoun, J. E. (2017). *Robot-proof: Higher education in the age of artificial intelligence*. MIT press.

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, 11(4), 959–975. http://doi.org/10.1016/j.joi.2017.08.007

Asaro, P. M. (2008). From mechanisms of adaptation to intelligence amplifiers: The Philosophy of W. Ross Ashby. In P. Husbands, O. Holland, & M. Wheeler (Ed.), *The mechanical mind in history*. (pp. 149-184). Cambridge, MA: The MIT Press

Asia Pacific Foundation of Canada (2019). *Artificial Intelligence Policies in East Asia: An Overview from the Canadian Perspective.* https://www.asiapacific.ca/sites/default/files/filefield/ai_report_2019.pdf

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Banfield, M. (2011). Metaphors and analogies of mind and body in nineteenth- century science and fiction: George Eliot, Henry James and George Meredith. In D.

Coleman & H. Fraser (Ed.), *Minds, bodies, machines 1770-1930*. (pp. 105-123). New York: Palgrave.

Batagelj, V., & Cerinšek, M. (2013). On bibliographic networks. *Scientometrics*, 96(3), 845–864. http://doi.org/10.1007/s11192-012-0940-1

Bereiter, C. (2002). *Education and mind in the knowledge age*. Lawrence Erlbaum Associates, Inc., Publishers.

Berryman, S. (2003). Ancient automata and mechanical explanation. *Phronesis*, 48(4), 344–369.

Bessen. J. (2019). Artificial intelligence and jobs: The role of demand. In A. Agrawal, J. Gans, & A. Goldfarb. (Eds.). *The Economics of Artificial Intelligence: An Agenda*. (pp. 291-308). Chicago: The University of Chicago Press.

Biscaro, C., & Giupponi, C. (2014). Co-authorship and bibliographic coupling network effects on citations. *PLoS ONE*, 9(6), e99502–12. http://doi.org/10.1371/journal.pone.0099502

Black, D. (2014). *Embodiment and mechanisation*. Surrey: Ashgate Publishing Limited.

Black, M. (1993). More about metaphor. In A. Ortony (Ed.), *Metaphor and Thought* (pp. 19-41). Cambridge: Cambridge University Press. doi:10.1017/CBO9781139173865.004

Boakes, R. (1984). *From Darwin to behaviourism*. New York: Cambridge University Press.

Boden, M. A. (2006). *Mind as machine: A history of cognitive science Volume1&2*. New York: Oxford University Press.

Bostrom, N. (2014). *Superintelligence: Paths, Dangers, Strategies*. Oxford, UK: Oxford University Press.

Boyack, K. W., Börner, K., & Klavans, R. (2008). Mapping the structure and evolution of chemistry research. *Scientometrics*, 79(1), 45–60. http://doi.org/10.1007/s11192-009-0403-5

Braidotti, R. (2013). *The posthuman*, Polity Press.

Bransford, J. D., Brown, A. N., and Cocking, R. R. (Eds.) (2000). *How people learn: Brain, mind, experience, and school*. Washington, DC.: National Academy Press.

Bredo, E. (1998). Evolution, psychology, and John Dewey's critique of the reflex arc concept. *The Elementary School Journal,* 98(5), 447–466.

Broadbent, D. E. (1958). *Perception and communication*. New York: Program Press.

Brynjolfsson, E., & McAfee, A. (2014). *The second machine age: Work, progress, and prosperity in a time of brilliant technologies*. WW Norton & Company.

Brynjolfsson, E., Rock, D., & Syverson. C. (2019). Artificial intelligence and the modern productivity paradox: A clash of expectations and statistics. In A. Agrawal, J. Gans, & A. Goldfarb. (Eds.). *The Economics of Artificial Intelligence: An Agenda*. (pp.23-60). Chicago: The University of Chicago Press.

Bullock, S. (2008). Charles Babbage and the emergence of automated reason. In P. Husbands, O. Holland, & M. Wheeler (Ed.), *The mechanical mind in history*. (pp. 19-40). Cambridge, MA: The MIT Press.

Calero-Medina, C., & Noyons, E. C. M. (2008). Combining mapping and citation network analysis for a better understanding of the scientific development: The case of the absorptive capacity field. *Journal of Informetrics*, 2(4), 272–279. http://doi.org/10.1016/j.joi.2008.09.005

Campenot, R. B. (2016). *Animal electricity: How we learned that the body and brain are electric machines*. Cambridge, MA: Harvard University Press.

Cardon, D., Cointet, J., & Mazières, A. (2018). Neurons spike back: The invention of inductive machines and the artificial intelligence controversy, *Réseaux* 2018/5 (n° 211), 173-220. DOI 10.3917/res.211.0173

Cockburn. I. M., Henderson. R., & Stern. S. (2019). The impact of artificial intelligence on innovation: An exploratory analysis. In A. Agrawal, J. Gans, & A. Goldfarb. (Eds.). *The Economics of Artificial Intelligence: An Agenda*. (pp. 115-148). Chicago: The University of Chicago Press.

Cohen-Cole, J. (2014). *The open mind: Cold war politics and the sciences of human nature*. Chicago: The University of Chicago Press.

Colburn, T. R., & Shute, G. M. (2008). Metaphor in computer science. *Journal of Applied Logic*, 6(4), 526–533. http://doi.org/10.1016/j.jal.2008.09.005

Cook, M. G. (2001). Divine artifice and natural mechanism: Robert Boyle's mechanical philosophy of nature. *Osiris*, 16, 133–150.

Corbett, A. T., Koedinger, K. R., & Anderson, J. R. (1997). Intelligent tutoring systems. In M. Helander, T. K. Landauer, & P. Prabhh (Eds.), *Handbook of Human-Computer Interaction* (pp. 1–34). New York: North-Holland.

Crane, T. (2003). *The mechanical mind: A philosophical introduction to minds, machines and mental representation*. New York: Routledge.

Creswell, J. W. (2014). *Research design: Qualitative, quantitative, and mixed methods approaches.* Los Angeles, CA: SAGE.

Creswell, J. W., Klassen, A. C., Clark, V. L. P., & Smith, K. C. (2011). *Best practices for mixed methods research in the health sciences.* Bethesda, Maryland: National Institute of Health.

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *Inter Journal*, Complex Systems, 1695. https://igraph.org.

Csardi, G., Nepusz, T., & Airoldi, E. M. (2016). Statistical network analysis with igraph. Springer. https://sites.fas.harvard.edu/~airoldi/pub/books/BookDraft-CsardiNepuszAiroldi2016.pdf

Danziger, K. (1990). *Constructing the subject: Historical origins of psychological research.* Cambridge, UK: Cambridge University Press.

Decuypere, M. (2019). Researching educational apps: Ecologies, technologies, subjectivities and learning regimes. *Learning, Media and Technology*, 44(4), 414–429. http://doi.org/10.1080/17439884.2019.1667824

Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the education process.* Lexington, MA: D.C. Heath and Company.

Dupuy, J. P. (2010a). *On the origins of cognitive science*. (M. B. DeBevoise, Trans.). Cambridge, Massachusetts: A Bradford Book.

Dupuy, J. P. (2010b). Cybernetics is antihumanism: Advanced technologies and the rebellion against the human condition. In G. R. Hansell & W. Grassie (Eds.). *H+/-: Transhumanism and its critics*. (pp. 227-248). Philadelphia, PA: Metanexus.

Evans, R. B. (1990). William James, "The Principles of Psychology". *The American Journal of Psychology*, 103(4), 433–447.

Executive Office of the President. (2016). *Artificial intelligence, automation, and the economy*, Washington. DC: The U.S. Government.

Frank, M. R., Wang, D., Cebrian, M., & Rahwan, I. (2019). The evolution of citation graphs in artificial intelligence research. *Nature Machine Intelligence*, 1(2), 79–85. http://doi.org/10.1038/s42256-019-0024-5

Frankish & Ramsey (2012), Introduction, In K. Frankish & W. M. Ramsey (Eds.). *The Cambridge handbook of cognitive science* (pp. 1-8). New York: Cambridge University Press.

Frantz, R. (2003). Herbert Simon. Artificial intelligence as a framework for understanding intuition, *Journal of Economic Psychology*, 24(2), 265–277. http://doi.org/10.1016/S0167-4870(02)00207-6

Fukuyama, F. (2004). Transhumanism. *Foreign Policy*, 144, 42–43.

Gane, N. (2006). When we have never been human, what is to be done? *Theory, Culture & Society*, 23(7-8), 135–158. http://doi.org/10.1177/0263276406069228

Garber, D. (2002). Descartes, mechanics, and the mechanical philosophy. *Midwest Studies in Philosophy*, 26(1), 185-204.

Gardner, H. (1985). *The mind's new science: A history of the cognitive revolution*. New York: Basic Books, Inc.

Garnelo, M., Arulkumaran, K., & Shanahan, M. (2016). Towards deep symbolic reinforcement learning. *arXiv.org*, 1–13.

Gazni, A., & Didegah, F. (2016). The relationship between authors' bibliographic coupling and citation exchange: analyzing disciplinary differences. *Scientometrics*, 107(2), 609–626. http://doi.org/10.1007/s11192-016-1856-y

Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton, Mifflin and Company.

Glanzel, W., & Czerwon, H. J. (1996). New methodological approach to bibliographic coupling and its application to the national, regional and institutional level. *Scientometrics*, 37(2), 195–221.

Glaser, R. (1984). Education and thinking: The role of knowledge. *American Psychologist*, 39(2), 93–104.

Glaser, R. (1991). The maturing of the relationship between the science of learning and cognition and educational practice. *Learning and Instruction*, 1(2), 129–144. http://doi.org/10.1016/0959-4752(91)90023-2

Goatly, A. (2007). *Washing the brain - Metaphor and hidden ideology*. Philadelphia: John Benjamins Publishing Company.

Graham, E. L. (2002). *Representations of the post/human: Monsters, aliens and others in popular culture*. New Brunswick, NJ: Rutgers University Press.

Green, C. D. (2009). Darwinian theory, functionalism, and the first American psychological revolution. *American Psychologist*, 64(2), 75–83. http://doi.org/10.1037/a0013338

Greenwood, J. D. (2008). Mechanism, purpose and progress: Darwin and early American psychology. *History of the Human Sciences*, 21(1), 103–126. http://doi.org/10.1177/0952695107086189

Greenwood, J., & Bonner, A. (2008). The role of theory-constitutive metaphor in nursing science. *Nursing Philosophy*, 9, 154–168.

Grunwald (2011). What does the debate on (post)human futures tell us? Methodology of hermeneutical analysis and vision assessment. In H. Tiroch-Samuelson & J.B. Hurlbut (Ed.), *Perfecting human futures: Transhuman visions and technological imaginations*. (pp. 35-50). Wiesbaden, Germany: Springer VS.

Gulson, K. N., & Webb, P. T. (2017a). Mapping an emergent field of "computational education policy": Policy rationalities, prediction and data in the age of Artificial Intelligence. *Research in Education*, 98(1), 14–26. http://doi.org/10.1177/0034523717723385

Gulson, K. N., & Webb, P. T. (2017b). "Life" and education policy: intervention, augmentation and computation. *Discourse: Studies in the Cultural Politics of Education*, 39(2), 276–291. http://doi.org/10.1080/01596306.2017.1396729

Hagstrom, R. P., Fry, M. K., Cramblet, L. D., & Tanner, K. (2016). Educational psychologists as scientist-practitioners: An expansion of the meaning of a scientist-practitioner. *American Behavioral Scientist*, 50(6), 797–807. http://doi.org/10.1177/0002764206296458

Hao, K. (2019, December 26). Baidu has a new trick for teaching AI the meaning of language. *MIT Technology Review*. Retrieved from https://www.technologyreview.com

Harari. Y. N. (2014). *Sapiens* [Kindle DX version]. Retrieved from Amazon.com

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron*, 95(2), 245–258. http://doi.org/10.1016/j.neuron.2017.06.011

Haraway, D. J. (1972). The search for organizing relations: An organismic paradigm in twentieth century developmental biology [Doctoral dissertation, Yale University]. ProQuest Dissertations Publishing.

Haraway, D. J. (2016). Cyborg manifesto. In D. J. Haraway (Ed.). *Manifestly Haraway*. (pp. 3–90). Minneapolis: University of Minnesota Press.

Hesse-Biber, S. N. (2010). *Mixed methods research: Merging theory with practice*. New York: The Guilford Press.

Hesse, M. (2000). Models and analogies. In W. Newton-Smith (Ed.) *A companion to the philosophy of science*. (pp.299-307) Malden, MA, Blackwell Publication.

Hatfield, G. (1995) Remaking the science of mind psychology as natural science in C. Fox, R. Porter, and R. Wokler (Ed.), *Inventing human science: Eighteenth-century domains*. (pp. 184-218). Berkeley, CA: University of California Press.

Hayles, N. K. (1999). *How we became posthuman*. Chicago: The University of Chicago Press.

Hayles, N. K. (2010). Wrestling with Transhumanism. In G. R. Hansell, & W. Grassie (Eds.). *H+/-: Transhumanism and its critics* (pp. 215-226). Philadelphia, PA: Metanexus.

Heims, S. J. (1991). *The cybernetics group*. Cambridge, MA: The MIT Press.

Helmstaedter, M. (2015). The mutual inspirations of machine learning and neuroscience. *Neuron*, 86(1), 25–28. http://doi.org/10.1016/j.neuron.2015.03.031

Hodges, A. (2008). What did Alan Turing mean by "Machine"?. In P. Husbands, O. Holland, & M. Wheeler (Ed.), *The Mechanical mind in history*. (pp. 74-90). Cambridge, MA: The MIT Press.

Husbands, P., Holland, O., & Wheeler, M. (2008). Introduction: The mechanical mind. In P. Husbands, O. Holland, & M. Wheeler (Ed.), *The mechanical mind in history*. (pp. 1-18). Cambridge, MA: The MIT Press.

Husbands, P. & Holland, O. (2008). The Ratio Club: A hub of British cybernetics. In P. Husbands, O. Holland, & M. Wheeler (Ed.), *The mechanical mind in history*. (pp. 91-148). Cambridge, MA: The MIT Press.

James, W. (1892). A plea for psychology as a "Natural science." *The Philosophical Review*, 1(2), 146–153.

Jarneving, B. (2007). Bibliographic coupling and its application to research-front and other core documents. *Journal of Informetrics*, 1(4), 287–307. http://doi.org/10.1016/j.joi.2007.07.004

Jasanoff, S. (Ed.). (2004). *States of knowledge: the co-production of science and the social order.* New York: Routledge.

Jee, C. (2020, April 26) Amazon's system for tracking its warehouse workers can automatically fire them. *MIT Technology Review*. Retrieved from https://www.technologyreview.com

Jeffrey, S. (2016). *The posthuman body in superhero comics*. New York: Palgrave Macmilan.

Kajikawa, Y., Ohno, J., Takeda, Y., Matsushima, K., & Komiyama, H. (2007). Creating an academic landscape of sustainability science: an analysis of the citation network. *Sustainability Science*, 2(2), 221–231. http://doi.org/10.1007/s11625-007-0027-8

Kao, Y.-F., & Venkatachalam, R. (2018). Human and machine learning. *Computational Economics*, 1–21. http://doi.org/10.1007/s10614-018-9803-z

Karunan, K., Lathabai, H. H., & Prabhakaran, T. (2017). Discovering interdisciplinary interactions between two research fields using citation networks. *Scientometrics*, 113(1), 335–367. http://doi.org/10.1007/s11192-017-2481-0

Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentataion*, 14(1), 10–25.

Khakurel, J., Penzenstadler, B., Porras, J., Knutas, A., & Zhang, W. (2018). The rise of artificial intelligence under the lens of sustainability. *Technologies*, 6(4), 100–18. http://doi.org/10.3390/technologies6040100

Klavans, R., & Boyack, K. W. (2007). Is there a Convergent Structure of Science? A Comparison of Maps using the ISI and Scopus Databases. *Presented at the 11th International Conference of the international Society for Scientometrics and Informetrics*, Madrid, Spain.

Kline, R. R. (2011). Cybernetics, automata studies, and the Dartmouth Conference on artificial intelligence. *IEEE Annals of the History of Computing*, 33(4), 5–16.

Konar, A. (2000). *Artificial intelligence and soft computing: Behavioral and cognitive modeling of the human brain*. New York: CRC Press.

Korinek. A., & Stiglitz. J. E. (2019). Artificial intelligence and its implications for income distribution and unemployment. In A. Agrawal, J. Gans, & A. Goldfarb. (Eds.). *The Economics of Artificial Intelligence: An Agenda*. (pp. 349-390). Chicago: The University of Chicago Press.

Kriegeskorte, N. (2015). Deep neural networks: A new framework for modelling biological vision and brain information processing. *bioRxiv, 1–22. http://doi.org/10.1101/029876*

Kulik, J. A. (2016). Effectiveness of intelligent tutoring systems: A meta-analytic review. *Review of Educational Research*, (1), pp. 42–78.

Kuhn, T. S. (1996). *The structure of scientific revolutions* (3rd ed.). The University of Chicago Press.

Kurzweil, R. (2012a). *How to create a mind*. London, Penguin Group.

Kurzweil, R. (2012b). Foreword to the third edition, In J. V. Neumann, *The Computer & The Brain*, New Haven: Yale University Press

Muri, A. (2006). *The enlightenment cyborg: A history of communications and control in the human machine 1660-1830.* University of Toronto Press.

Lagemann, E. C. (1989). *The politics of knowledge: The Carnegie Corporation, philantropy, and public policy*. Wesleyn University Press.

Lagemann, E. C. (2000). *An elusive science*. The University of Chicago Press.

Lake, B. M., Ullman, T. D., Tenenbaum, J. B., & Gershman, S. J. (2016). Building machines that learn and think like people. *Behavioral and Brain Sciences*, 1–58.

Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. The University of Chicago Press.

Leicht, E. A., Clarkson, G., Shedden, K., & Newman, M. E. J. (2007). Large-scale structure of time evolving citation networks. *The European Physical Journal B*, 59(1), 75–83. http://doi.org/10.1140/epjb/e2007-00271-7

Levin, H. (1987). Successions in psychology: The cognitive revolution in psychology. *Science*, 236(4809), 1683–1684. http://doi.org/10.1126/science.236.4809.1683

Leydesdorff, L. (2006). Betweenness centrality as an indicator of the interdisciplinarity of scientific journals. *Journal of the American Society for Information Science and Technology*, 58(9), 1303-1319.

Leydesdorff, L. (2007). On the normalization and visualization of author co-citation data: Salton's Cosine versus the Jaccard index. *Journal of the American Society for Information Science and Technology*, 59(1), 77–85. http://doi.org/10.1002/asi.20732

Leydesdorff, L., & Rafols, I. (2009). A global map of science based on the isi subject categories. *Journal of the American Society for Information Science and Technology*, 60(2), 348–362.

Leydesdorff, L., & Rafols, I. (2011). Indicators of the interdisciplinarity of journals: Diversity, centrality, and citations. *Journal of Informetrics*, 5(1), 87–100. http://doi.org/10.1016/j.joi.2010.09.002

Leydesdorff, L., Carley, S., & Rafols, I. (2012). Global maps of science based on the new Web-of-Science categories. *Scientometrics*, 94(2), 589–593. http://doi.org/10.1007/s11192-012-0784-8

Leydesdorff, L., Wagner, C. S., & Bornmann, L. (2017). Betweenness and diversity in journal citation networks as measures of interdisciplinarity—A tribute to Eugene Garfield. *Scientometrics*, 114(2), 567–592. http://doi.org/10.1007/s11192-017-2528-2

Lilley, S. (2013). *Transhumanism and society*. New York: Springer.

Lobo, L., Heras-Escribano, M., & Travieso, D. (2018). The history and philosophy of ecological psychology. Frontiers in Psychology, 9, 2228.

Ma, R. (2012). Author bibliographic coupling analysis: A test based on a Chinese academic database. *Journal of Informetrics*, 6(4), 532–542. http://doi.org/10.1016/j.joi.2012.04.006

Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. Journal of Educational Psychology, 106(4), 901–918. http://doi.org/10.1037/a0037123

Marcinkowski, F., Kieslich, K., Starke, C., & Lünich, M. (2020). Implications of AI (un-)fairness in higher education admissions (pp. 122–130). Presented at the FAT* '20: Conference on Fairness, Accountability, and Transparency, New York, NY, USA: ACM. http://doi.org/10.1145/3351095.3372867

Martin-Loeches, M. (2015). Neuroscience and education: We already reached the tipping point. *Psicología Educativa*, 21(2), 67–70. http://doi.org/10.1016/j.pse.2015.09.001

Mayer, R. E. (1992). Cognition and instruction: Their historic meeting within educational psychology. *Journal of Educational Psychology*, 84(4), 405–412.

Mazlish, B. (1993). *The fourth discontinuity*. New Haven: Yale University Press.

McCandliss, B. (2016, December 6th). *Early education and the brain: Making novel connections* [Video file]. AERA Centennial Lecture Series and Discussion Forum. Retrieved from https://www.youtube.com/watch?v=nGmbcNz9JQ8&feature=youtu.be

McCain, K. W. (1998). Neural networks research in context: A longitudinal journal co-citation analysis of an emerging interdisciplinary field. *Scientometrics*, 41(3), 389–410.

McCorduck, P. (2004). *Machines who think*. Natick: A K Peters, Ltd.

McClelland, J. L., Rumelhart, D. E., & PDP Research Group. (1986). *Parallel distributed processing* (Vol. 2). Cambridge, MA: MIT press.

McKinney, S. M., Sieniek, M., Godbole, V., Godwin, J., Antropova, N., Ashrafian, H., Back, T., Chesus, M., Corrado, G. S., Darzi, A., Etemadi, M., Garcia-Vicente, F., Gilbert, F. J., Halling-Brown, M., Hassabis, D., Jansen, S., Karthikesalingam, A., Kelly, C. J., King, D., Ledsam, J. R., … Shetty, S. (2020). International evaluation of an AI system for breast cancer screening. *Nature*, 577(7788), 89-94.

Means, A. J. (2019). Hypermodernity, automated uncertainty, and education policy trajectories. *Critical Studies in Education*, 00(00), 1–16. http://doi.org/10.1080/17508487.2019.1632912

Medler, D. A. (1998). A brief history of connectionism. *Neural Computing Surveys*, 1(2), 18–72.

Mehan, H. (2014). The prevalence and use of the psychological–medical discourse in special education. *International Journal of Educational Research*, 63, 59–62. http://doi.org/10.1016/j.ijer.2012.10.003

Mehta, J. (2013). *The allure of order*. New York: Oxford University Press.

Merchant, C. (1980). *The death of nature: Women, ecology, and the scientific revolution*. New York: HarperSanFrancisco

Michie, D. (2008). Alan Turing's mind machines . In P. Husbands, O. Holland, & M. Wheeler (Ed.), *The mechanical mind in history*. (pp. 61-74). Cambridge, MA: The MIT Press.

Miller, G. A. (2003). The cognitive revolution: a historical perspective. *Trends in Cognitive Sciences*, 7(3), 141–144. http://doi.org/10.1016/S1364-6613(03)00029-9

Mirowski, P. (2002). *Machine dreams: Economics becomes a cyborg science*. New York, NY: Cambridge University Press.

Moed, H. F. (2005). *Citation Analysis in Research Evaluation*. Dordrecht: Springer.

Moya-Anegón, S. G. F. de, Vargas-Quesada, B., Chinchilla-Rodríguez, Z., Corera-Álvarez, E., Munoz-Fernández, F. J., & Herrero-Solana, V. (2007). Visualizing the marrow of science. *Journal of the American Society for Information Science and Technology*, 58(14), 2167–2179. http://doi.org/10.1002/asi.20683

National Research Council. (2000). *How people learn: Brain, mind, experience, and school*. Washington, D.C.: National Academy Press.

Nasr, K., Viswanathan, P., & Nieder, A. (2019). Number detectors spontaneously emerge in a deep neural network designed for visual object recognition. *Science Advances*, 5(5), eaav7903.

Newman, M. (2010). *Networks: An introduction*. [Kindle Edition]. Oxford, UK: Oxford University Press.

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

Nilsson, N. (2010). *The quest for artificial intelligence: A history of ideas and achievements*. Web Version: Nils J. Nilsson.

Nunnally, J. C. (1978). An overview of psychological measurement. *Clinical Diagnosis of Mental Disorders*, 97-146.

Núñez, R., Allen, M., Gao, R., Rigoli, C. M., Relaford-Doyle, J., & Semenuks, A. (2019). What happened to cognitive science? *Nature Human Behaviour*, 1–10. http://doi.org/10.1038/s41562-019-0626-2

OECD. (2020, March 21). *OECD Principles on AI*. https://www.oecd.org/going-digital/ai/principles/

Oleinik, A. (2019). What are neural networks not good at? On artificial creativity. *Big Data & Society*, 6(1), 205395171983943. http://doi.org/10.1177/2053951719839433

Osler, M. J. (2001). Whose ends? Teleology in early modern natural philosophy. *Osiris*, 16. *Science in Theistic Contexts: Cognitive Dimensions*, 151–168.

Perez. C. E. (2017). *The deep learning AI playbook: Strategy for disruptive artificial intelligence.* [Kindle DX version]. Retrieved from Amazon.com.

Piccinini, G. (2008). Computational mechanisms. In S. Glennan & P. Illari (Ed.), *The Routledge handbook of mechanisms and mechanical philosophy*. (pp. 435-446). New York: Routledge.

Pickering, A. (2009). *The cybernetic brain: Sketches of another future*. Chicago, IL: The University of Chicago Press.

Popp, J. A. (2007). *Evolution's first philosopher*. New York: State University of New York Press.

Porter, A. L., & Rafols, I. (2009). Is science becoming more interdisciplinary? Measuring and mapping six research fields over time. *Scientometrics*, 81(3), 719–745. http://doi.org/10.1007/s11192-008-2197-2

Rafols, I., Porter, A., & Leydesdorff, L. (2010). Science overlay maps: A new tool for research policy and library management. *Journal of the American Society for Information Science and Technology*, 61(9), 1871–1887.

Raschka, S., & Mirjalili, V. (2017). *Python machine learning: Machine learning and deep learning with Python. Scikit-Learn, and TensorFlow* (2nd ed.). Packt Publishing.

Rousseau, R. (2010). Bibliographic coupling and co-citation as dual notions. In L. Birger (Ed.), *A Festschrift in honour of Peter Ingwersen*, special volume of the e-zine of the ISSI, June 2010, 173-183.

Rumelhart, D. E., Hinton, G. E., & Williams, R. J. (1988). Learning representations by back-propagating errors. *Cognitive modeling*, 5(3), 1.

Russell, S. J., & Norvig, P. (2001). Artificial intelligence: A modern approach. New Jersey: Prentice-Hall.

Sachs. J. D. (2019). R&D, Structural transformation, and the distribution of income. In A. Agrawal, J. Gans, & A. Goldfarb. (Eds.). *The Economics of Artificial Intelligence: An Agenda*. (pp. 349-390). Chicago: The University of Chicago Press.

Salisbury, L. (2011). Linguistic trepanation: Brain damage, penetrative seeing and a revolution of the word. In D. Coleman & H. Fraser (Ed.), *Minds, bodies, machines 1770-1930*. (pp. 179-208). New York: Palgrave.

Schneider, W., & Graham, D. J. (1992). Introduction to connectionist modeling in wducation. *Educational Psychologist*, 27(4), 513–530. http://doi.org/10.1207/s15326985ep2704_7

Schwab, C. (2016). *The fourth industrial revolution* [Kindle DX version]. Retrieved from Amazon. com

Sellar, S., & Gulson, K. N. (2019). Becoming information centric: The emergence of new cognitive infrastructures in education policy. *Journal of Education Policy*, 53, 1–18. http://doi.org/10.1080/02680939.2019.1678766

Sen, S. K., & Gan, S. K. (1983). A mathematical extension of the idea of bibliographic coupling and its applications. *Annals of Library Science and Documentation*, 30(2), 78–82.

Shavelson, R. J., & Towne, L. (Eds.) (2002). *Scientific research in education*. Washington, DC: National Academy Press.

Small, H. (2003). Paradigms, citations, and maps of science: A personal history. *Journal of the American Society for Information Science and Technology*, 54(5), 394-399.

Somers, J. (2013). The man who would teach machines to think. *The Atlantic*, 312, 90–100.

Struik, P. C., Yin, X., & Meinke, H. (2008). Plant neurobiology and green plant intelligence: science, metaphors and nonsense. *Journal of the Science of Food and Agriculture*, 88(3), 363–370. http://doi.org/10.1002/jsfa.3131

Sundar, S. S. (2020). Rise of Machine Agency: A Framework for Studying the Psychology of Human–AI Interaction (HAII). *Journal of Computer-Mediated Communication*, 25(1), 74–88. http://doi.org/10.1093/jcmc/zmz026

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed). London, UK: The MIT Press.

Taddy. M. (2019). The technological elements of artificial intelligence. In A. Agrawal, J. Gans, & A. Goldfarb. (Eds.). *The Economics of Artificial Intelligence: An Agenda*. (pp. 61-88). Chicago: The University of Chicago Press.

Talkhabi, M., & Nouri, A. (2012). Foundations of cognitive education: Issues and opportunities. *Procedia - Social and Behavioral Sciences*, 32, 385–390. http://doi.org/10.1016/j.sbspro.2012.01.058

Thagard, P. (2005). *Mind: Introduction to cognitive science*. Cambridge, MA: A Bradford Book.

The World Bank. (2018). *Learning to realize education's promise*. Washington, DC: The World Bank.

The General Language Understanding Evaluation. (2020, March 21). *GLUE diagnostic dataset*. GLUE. https://gluebenchmark.com/diagnostics

Thijs, B., Zhang, L., & Glänzel, W. (2015). Bibliographic coupling and hierarchical clustering for the validation and improvement of subject-classification schemes. *Scientometrics*, 105(3), 1453–1467. http://doi.org/10.1007/s11192-015-1641-3

Thorndike, E. L. and Woodworth, R.S. (1901) . The influence of improvement in one mental function upon the efficiency of other functions.(I). *Psychological review* 8.3 (1901): 247.

Tomlinson, S. (1996). From Rousseau to evolutionism: Herbert Spencer on the science of education. *History of Education*, 25(3), 235–254. http://doi.org/10.1080/0046760960250303

Trajtenberg. M. (2019). Artificial intelligence as the next GPT: A political-economy perspective. In A. Agrawal, J. Gans, & A. Goldfarb. (Eds.). *The Economics of Artificial Intelligence: An Agenda*. (pp. 175-188). Chicago: The University of Chicago Press.

Wiener, N. (1985). *Cybernetics: or control and communication in the animal and the machine*. Cambridge, Massachusetts: The MIT Press.

Wiener, N. (1989). *The human use of human beings*. London: Free Association Books.

Williamson. B. (2016). Digital education governance: data visualization, predictive analytics, and “real-time” policy instruments. *Journal of Education Policy*, 31(2), 1–19. http://doi.org/10.1080/02680939.2015.1035758

Van Lunteren, F. (2016). Clocks to computers: A machine-based “Big Picture” of the history of modern science. *The History of Science Society*, 107(4), 762–776.

Von Neumann, J. (2012). *The computer & the brain* (3rd ed.). New Haven: Yale University Press.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard university press.

Yan, E., & Ding, Y. (2012). Scholarly network similarities: How bibliographic coupling networks, citation networks, cocitation networks, topical networks, coauthorship networks, and coword networks relate to each other. *Journal of the American Society for Information Science and Technology*, 63(7), 1313–1326. http://doi.org/10.1002/asi.22680

Zhao, D., & Strotmann, A. (2008). Evolution of research activities and intellectual influences in information science 1996-2005: Introducing author bibliographic-coupling analysis. *Journal of the American Society for Information Science and Technology*, 59(13), 2070–2086. http://doi.org/10.1002/asi.20910