

The Pioneer Research Journal

*An International Collection of Undergraduate-
Level Research*

Volume 2

2015

Pioneer
academics

The Pioneer Research Journal

*An International Collection of Undergraduate-Level
Research*

Contents

Contributing Readers.	iii
Foreword.	ix
Selection Process.	xi
Effects of Aerobic Exercise on the Hippocampus (Neuroscience)	1
Author: Hugo Taro Blázquez Yamagishi	
School: IES San Mateo – Madrid, Spain	
Pioneer Seminar: Understanding the Sense of Touch	
The Dynamics of Movie Selection: A Study of Social Contagion Using the Revised SI Model Based on Word of Mouth (Computer Science)	37
Author: Xiangting Chen	
School: The Affiliated High School of South China Normal University – Guangzhou, China	
Pioneer Seminar: Networks: from Searching the World Wide Web to Slowing the Spread of Infectious Diseases	
Adolescent Decision Making: Neuroscientific Basis of Risk-Taking Behavior in the Developing Brain (Neuroscience)	71
Author: You Jin Jung	
School: Korea International School – Seoul, South Korea	
Pioneer Seminar: The Socioemotional Brain and Decision Making	
The Effect of Diffraction Aperture Structure on Reciprocal Space (Physics).	103
Author: Eren Kafadar	
School: Robert College – Istanbul, Turkey	
Pioneer Seminar: Optical Diffraction Patterns	

The Hidden Key to Success: How Saudi Arabia’s Economic Needs are Expanding Women's Rights (Culture Studies). 111

Author: Mae Kirkpatrick

School: Hockerill Anglo-European College – Hertfordshire, United Kingdom

Pioneer Seminar: Women’s Rights in the Middle East and South Asia

Which Model is Feasible for Hong Kong? A Comparison among the Democracies in Britain, Mainland China, and Hong Kong (Political Science). 137

Author: Zhimin Lin

School: The Affiliated High School of South China Normal University – Guangzhou, China

Pioneer Seminar: Comparative Political Institutions of the World

The Difference between Oversampling and Undersampling: The Proving of the Nyquist Sampling Theorem (Physics). 179

Author: Yige Liu

School: Beijing No. 4 High School, International Campus – Beijing, China

Pioneer Seminar: Optical Diffraction Patterns of Live Species

Accelerated or Gifted Education——Which One is a Better Fit for Chinese Students? (Education). 189

Author: Sicheng Luo

School: Hangzhou Foreign Languages School – Hangzhou, China

Pioneer Seminar: On Education: Fundamentals, Philosophies, Realities

Transformation of the Meanings of Arts and Crafts in China and Australia: A Comparison (Anthropology) 211

Author: Ziqi Ma

School: The High School Affiliated to Renmin University of China – Beijing, China

Pioneer Seminar: Communication and Culture

Pesticides and Endocrine Disruptors: Understanding the Molecular Mechanism by which Some Pesticides (Aldrin, DDT, Mecoprop) Act as Endocrine Disruptors (Chemistry). 237

Author: Jonathan Mak

School: Chinese International School – Hong Kong

Pioneer Seminar: Crystals, Molecular Structure and Medicinal Chemistry

Sense and Sensibility of Artificial Intelligence: When Machines Can Think
(Literature). 265

Author: Yu Que

School: The Affiliated High School of South China Normal University – Guangzhou, China

Pioneer Seminar: Science, Technology and Society

Brain-Controlled Prosthetics (Engineering). 281

Author: Amanda Rabin

School: Pine Crest School – Fort Lauderdale, FL, United States

Pioneer Seminar: Mechanical Engineering

Hukou, Development, and Fairness (Economics). 297

Author: Darren Tong

School: Chinese International School – Hong Kong

Pioneer Seminar: The Political Economy of International Development

The Pursuit of Truth in Relation to the Faculties of Sensation, Perception,
and Judgment -- An Essay Written in the Context of the Meditations by
Rene Descartes (Philosophy) 323

Author: Yuan Wang

School: The Affiliated High School of South China Normal University – Guangzhou, China

Pioneer Seminar: Descartes' Meditations

Biology (Macromolecular Machines): A Summary of Bacterial Secretion
Systems (with Special Attention to the Type III Secretion System)
(Biology). 343

Author: Timothy Wee

School: Shanghai Community International School – Shanghai, China

Pioneer Seminar: Introduction to Cells and Intracellular Machines

The Motif of Culture: The Costume Culture of the Yi Ethnic Group of
Liangshan (Anthropology). 379

Author: Yilin Wen

School: Chengdu Foreign Language School – Chengdu, China

Pioneer Seminar: Non-Verbal Communication

Is Greece’s Crisis in the Eurozone a “Tipping Point” in the Development of the European Union? The Challenges and Opportunities in the Greek Crisis and the Fate of Europe’s Integration (International Relations) 403

Author: Sibe Wu

School: The High School Affiliated to Renmin University of China – Beijing, China

Pioneer Seminar: European Integration in Comparative Perspective

Women Uniting on the Web: Social Media Contributions Against Gender-Based Violence in Turkey (Culture Studies) 427

Author: Ashley Zhang

School: Crean Lutheran High School – Irvine, CA, United States

Pioneer Seminar: Women’s Rights in the Middle East and South Asia

**Hongkang Yang’s paper, “A Good Intention Resulted in a Global Environmental Threat: Imidacloprid and Its Unanticipated Effects,” was also selected for publication in The Pioneer Research Journal, Volume 2. However, at the time of publication, his paper was under consideration for publication in another research journal. As such, he elected not to publish his paper here. Hongkang Yang attends Shenzhen High School in Shenzhen, China.*

Effects of Aerobic Exercise on the Hippocampus

Hugo Taro Blázquez Yamagishi

Author background: Hugo grew up in Spain and currently attends IES San Mateo in Madrid. His Pioneer seminar program was titled “Understanding the Sense of Touch.”

Abstract

The purpose of this research proposal is to study how the practice of aerobic exercise affects hippocampal formation and functions of the hippocampus, in particular the memory and its complex processes. The relation of the development of the hippocampus to depression and the mood state and to aerobic exercise will also be studied.

Much research has been done on the endless benefits of exercise in averting atrophy of the brain in aged individuals, not specifically about memory processes but about other essential aspects. Therefore, this proposal is focused on the effects in young subjects, whose brain structure and functions are still in development, with the objective of studying how much exercise can strengthen, magnify and prolong this hippocampal growth to reach the best cognitive capacity and maturity conditions.

To research this mainly unknown but important topic, we will develop a within-subjects design to help control for the wide variability in ability between individuals.¹We will compare the change in test scores and scan results between various groups. The differences between pre- and post-

scores of all groups will be measured and compared. The following tests planned to study our hypothesis are only proposals, as we do not have either the required means or time to perform the whole experiment.

A lot is known about the effects of exercise on improving hippocampal function in rodents, but this work has mainly focused on spatial memory or memory encoding, not consolidation per se. There is much less evidence in humans that exercise changes the hippocampus and even less information on its consolidation function.²

Hypothesis

We predict that aerobic exercise training will produce changes in the hippocampus both structurally and functionally. Its enhancement in terms of the size will be correlated with improvements in the functions the hippocampal region is involved in. In particular, exercise will develop young individuals' memory consolidation, a process through which memories are stabilized in the long-term, but also the long-term potentiation phenomenon of it. This means a better capacity for converting short-term memories and information into long-term ones. It is also predicted that depression scores and risks of falling into a depressive mental disease will be lower in trained subjects.

Introduction

Physical activity, specifically aerobic exercise, increases hippocampal and medial temporal lobe volumes, so these parts of the brain in more fit adults are significantly larger. This could imply that exercise is a stimulus to

hippocampal plasticity.^{3,4} Plasticity refers to modification of neural tissue of the living brain and nervous system in response to experience and injury.

These changes in the brains of aerobically trained subjects are bigger, and they demonstrate considerably greater improvement on a neuropsychological test battery including depression scores, sensory thresholds and visual acuity.^{3,4}

The degree to which this exercise practice can alter late adulthood volume of the hippocampus is still unknown.⁵

This small but indispensable region of our brain is located in the medial temporal lobe beneath the cortical surface.

Regarding its structure, it is divided into two halves, which lie in the left and right hemispheres of the brain. Additionally, the hippocampus forms part of the limbic system.

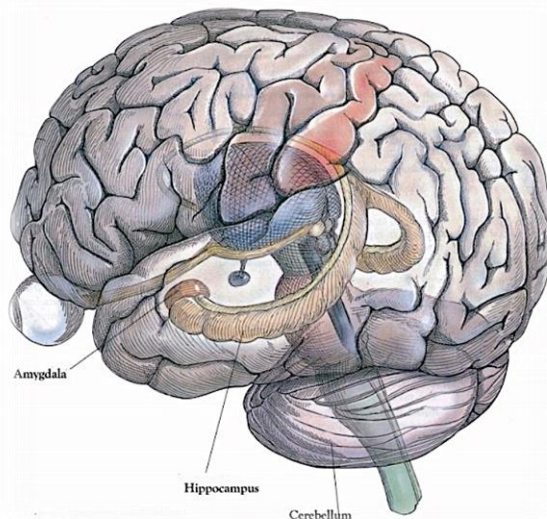


Figure 1. Illustration of the Hippocampus

It can be distinguished externally as a layer of densely packed neurons. In human adults, the volume of the hippocampus on each side of the brain is

about 3-3.5 cm³ as compared to 320-420 cm³ for the volume of the neocortex.⁶

It contains two parts: the cornu ammonis (hippocampus proper) and the dentate gyrus. Both of these parts are separated by the hippocampal sulcus and curve into each other. Below the sulcus lies the subiculum. Since the hippocampus is a part of the allocortex (archicortex), there is a zone that separates it from the neocortex. Some anatomists divide it into the hippocampus proper (Cornu ammonis; CA), the dentate gyrus (DG), the subiculum, and the entorhinal area (EC). This entire set is called the hippocampal formation.

The word "hippocampus" comes from the Greek (πόκαμπος, "seahorse"; from ἵππος hippos, which means "horse" and κάμπος kampos) and it turns out that this organ has a shape that resembles a seahorse.⁷

This region is primarily associated with spatial navigation and it is known that it is also involved in long-term memory, including the storage of all past-acquired knowledge and experiences throughout our entire lifetime. It is thought that the hippocampus works with the amygdala, which is another important part of the limbic system responsible for fear and for storing memories of events for future recognition. Thus both the hippocampus and the limbic system play essential roles in human memory.

Experts believe the hippocampus may work as a gateway through which new memories pass to enter permanent long-term memory storage.⁵ It is generally agreed that the hippocampus also has part of the role of detecting new surroundings, occurrences and stimuli. Some scientists believe it is

specifically involved in declarative memory (explicit memory), which is one of the two main types of memory into which long-term memory is divided, and consists of facts, events and performance skills that can be consciously recalled or declared. It has to do with information that can be explicitly stored and retrieved.⁶

Therefore, this region of the brain when diseased has several potential and devastating clinical consequences. It's the first and most affected structure in most of the neuropsychiatric disorders such as Alzheimer's disease and epilepsy and its damage may lead to anterograde amnesia. It is also involved in severe mental illnesses, for instance schizophrenia, grave depression and bipolar disorder. Patients suffering from hippocampus damage that results in amnesia may remember every detail of their childhood but not things that happened recently after the injury.⁸ Despite this, the results of recent studies indicate that damage to the hippocampus may not affect a person's ability to learn a new skill or to acquire new habits.^{8,9}

For many years, the neuropsychologist Brenda Milner studied a patient known as HM (Henry Gustav Molaison) who suffered from profound amnesia. These studies resulted in one of the most notable discoveries of the brain and meant a huge step forward in neuroscience. At the young age of seven, Henry suffered from epilepsy, and from that time his seizures became increasingly severe. He accepted an experimental and complicated surgery, as he was no longer able to live independently. In this procedure part of his temporal lobe, with a portion of both the hippocampus and the amygdala, was removed. After the surgery, he was able to remember much of his childhood, family, and historical facts but he struggled to remember some

things that happened up to 11 years before. He had severe anterograde amnesia and he lost the ability to form new memories. Milner started to work with him with the purpose of understanding this topic area which wasn't really known yet. She asked patient Henry to draw a line between two outlines of a five-pointed star while watching his hand and the page in a mirror. He was told to repeat this task many times on different occasions. Each time he did it he didn't remember doing this task before, yet his performance improved each time. This demonstrated that despite the fact that he was not really conscious of it his brain was learning new motor skills by practice. Milner concluded that this form of memory must be located in a part of Henry's brain unaffected by his operation.⁹

Various factors can harm the hippocampal region even though the main reason why it can be injured is the global reduction of oxygen in the body. This is owing to its high sensitivity to oxygen deprivation or hypoxia, as during a heart attack, for instance, or respiratory failure, apnea, carbon monoxide poisoning and so on.

Regarding hippocampal pharmacology, the hippocampus receives projections that contain serotonin (5-HT), norepinephrine (NE) and dopaminergic neurons. It contains rich numbers of glucocorticoids, estrogen, and progesterone receptors.⁵

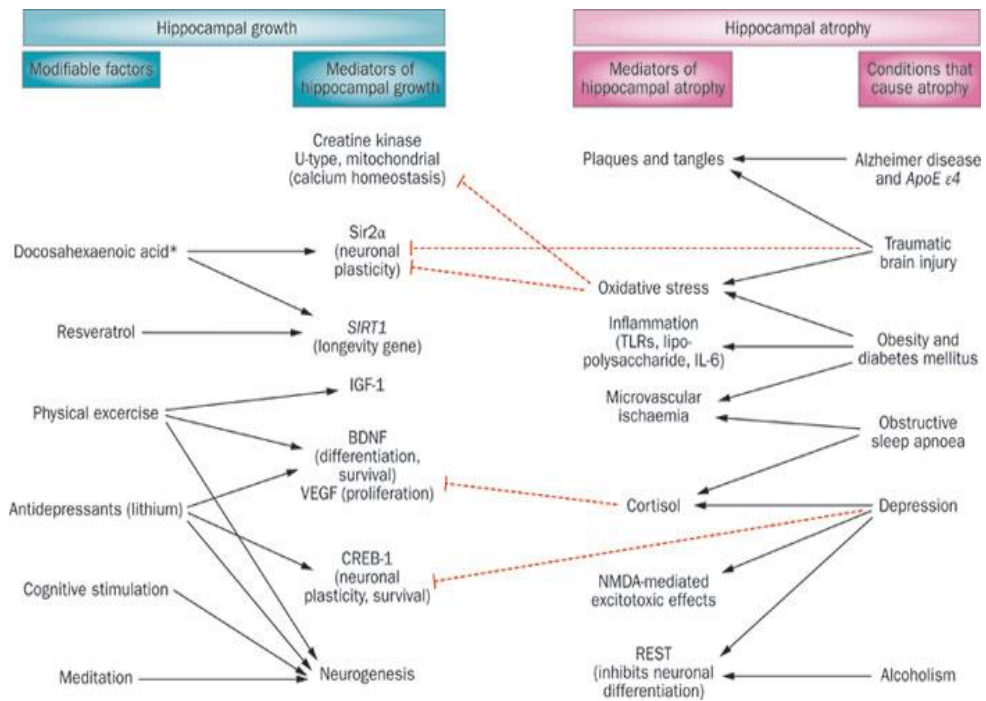


Figure 2. Pathways leading to hippocampal growth or atrophy³⁷

From: Majid Fotuhi, David Do and Clifford Jack (2012)

As indicated in the diagram, physical exercise is one of the main factors that have the capacity to improve hippocampal development basically through stimulating its neurogenesis. It also increases levels of a chemical called BDNF, which means an improvement in hippocampal plasticity.

In this experiment, we will focus on the improvement of the hippocampus in the memory consolidation process and in depression and mood state regulators through aerobic exercise.

The hippocampus has an important role in a process known as memory consolidation. It consists in the transfer of knowledge from short-term

memory to long-term memory where these memories will be stored for a long period of time. It describes the process wherein a memory is refined, stored and stabilized for long-term use.¹⁰ Memory consolidation is usually divided into two phases or processes, synaptic consolidation and system consolidation. Synaptic consolidation takes place in the first hours after learning or encoding information whereas in system consolidation, the memories that depend on the hippocampus become independent over a period of time, from a couple of weeks to some years.

Neurologically speaking, the consolidation process consists of a phenomenon called long-term potentiation (LTP), which through a synapse is able to increase strength and number of signals transmitted between the neurons. This process called potentiation fires neurons synchronously, so it makes those neurons more inclined to fire together in the future. Therefore, long-term potentiation occurs when some neurons of the same group fire together so often that they become permanently sensitized to each other. The brain creates more connections and pathways as new memories and experiences need to be accumulated; it also may rewire itself by rerouting connections and reorganizing them. For example, if dance choreography is repeated several times, this repeated firing of specific synapses in a certain order in the brain makes it easier to repeat this firing later on, which means a better performance of the whole choreography, fewer mistakes with the steps and moving with more ease. This process of neurological changes the brain makes itself is called neural plasticity or synaptic plasticity.

The opposite of long-term potentiation, called long-term depression or depotentiation (LTD), can also occur. In this case by the silencing of their

synaptic connections neural networks involved in erroneous movements are inhibited. It is produced by nerve impulses reaching the synapses at very low frequencies and instead of becoming more efficient these connections are weakened. Contrary to long-term depression, long-term potentiation is triggered by high-frequency stimulation of synapses.¹¹

Methods and Measures

For the experimental design, there will be 3 groups of 10 young people, 5 females and 5 males aged between 15 to 25 years old, which is believed to be the average age at which the brain is fully developed and mature.¹² Groups 1 and 2 will consist of people who have never had a brain injury or a mental disease. Group 3 will be made up of subjects diagnosed with severe depression.

As mentioned before, we will focus on subjects whose brains are not fully developed yet because there has been some research done on how doing sports avoids adults' brain degeneration but not on whether it is possible to improve young individuals' hippocampus function with aerobic exercise while their brains are still developing. How much can aerobic exercise strengthen the development of the hippocampal region, thus improving moods and cognitive function?

Group 1 will be asked to live their normal lives without changing their routines and without doing any significant exercise. Groups 2 and 3 will practice swimming for 60 minutes, 4 days per week for 3 months. Each member that has to do this exercise will be asked to swim one hour at the end of the day 4 days a week at a comfortable and regular pace. Except for

this, they will be following their everyday routine.

The short-term memory capacity of these subjects before the aerobic exercise will be tested with the most common technique, which is a test called Memory Span. It consists of showing the candidate a large list of random words, numbers or pictures one by one, one thing per second. Recognizable and easy patterns mustn't be used.

In this case they will be shown a sequence of 7 pictures, 7 numbers and 7 words in a random order, making a total of 21 items. Having finished the sequence, the candidate will be asked to recall all the items in the right order immediately. Only the items recalled in the right order will be taken into account as a correct answer. The number of right responses of each member of both groups will be written down and saved.

In addition, a long-term memory test will be done. The individuals, just as in the short-term memory test, will have to memorize a list of random words. They will have two minutes to memorize a list of 20 words. Once the time is over, they will have 5 minutes without the list and then try to recite them all in the correct order. In this free 5 minutes they have, they will be asked to undergo an fMRI scan so when the time is over, they will have to recall the words while their hippocampal region is scanned. With this fMRI imaging we will be able to see how their hippocampus, which is working at long-term memory information recall, is initially activated and compare it later.

Every individual will also be asked to practice playing a simple but long piano piece. They shouldn't have ever played a piano before, which means

they are doing a new task for the first time in their life. They will have to practice for 45 minutes. This will allow us to study their long-term potentiation and long-term cognitive functions as well as their capacity to acquire a new skill. Their hippocampus and brain after practicing will be scanned with an fMRI and MRI to see their activated areas. Once the scan is taken, every group will have to practice the same piano composition for two weeks, to be able to play it smoothly and without the score, which means they will have memorized and learned every movement. Having finished the two-week piano training period, they will only have one more week in the middle of the three-month period to practice it again. After that week, they won't touch the piano or see the same score until the next test that will be done after the aerobic exercise training period (for groups 2 and 3) (three empty months for group 1).

A magnetic resonance imaging (MRI) scan of their brain will be taken after having done the Memory Span test and the piano practice session, but before starting the exercise training period.

In this non-invasive test the hydrogen nucleus (a single proton) is used because of its abundance in water and fat.

The hydrogen proton can be likened to the planet Earth, spinning on its axis, with a north-south pole. In this respect it behaves like a small bar magnet. Under normal circumstances, these hydrogen proton "bar magnets" spin in the body with their axes randomly aligned. When the body is placed in a strong magnetic field, such as an MRI scanner, the protons' axes all line up. This uniform alignment creates a magnetic vector oriented along the axis of

the MRI scanner. MRI scanners come in different field strengths, usually between 0.5 and 1.5 teslas. When additional energy (in the form of a radio wave) is added to the magnetic field, the magnetic vector is deflected. The radio wave frequency (RF) that causes the hydrogen nuclei to resonate is dependent on the element sought (hydrogen in this case) and the strength of the magnetic field. The strength of the magnetic field can be altered electronically from head to toe using a series of gradient electric coils, and by altering the local magnetic field by these small increments, different slices of the body will resonate as different frequencies are applied. When the radiofrequency source is switched off the magnetic vector returns to its resting state, and this causes a signal (also a radio wave) to be emitted. It is this signal which is used to create the MR images. Receiver coils are used around the body part in question to act as aerials to improve the detection of the emitted signal. The intensity of the received signal is then plotted on a grey scale and cross-sectional images are built up.¹³

This will allow us to have a clear image of the hippocampal and limbic system volume for later comparison.

We will measure the subjects' hippocampal and limbic system activity as a result of their brainwork during the test and the piano playing with an fMRI scan. Functional magnetic resonance imaging works by detecting the changes in blood oxygenation and flow that occur in response to neural activity. Oxygen is delivered to neurons by hemoglobin in capillary red blood cells. When neuronal activity increases, the demand for oxygen increases and the local response is an increase in blood flow to regions of increased neural activity. Hemoglobin is diamagnetic when oxygenated but

paramagnetic when deoxygenated. This difference in magnetic properties leads to small differences in the MR signal of blood depending on the degree of oxygenation. Since blood oxygenation varies according to the levels of neural activity these differences can be used to detect brain activity. This form of MRI is known as blood oxygenation level dependent (BOLD) imaging.¹⁴

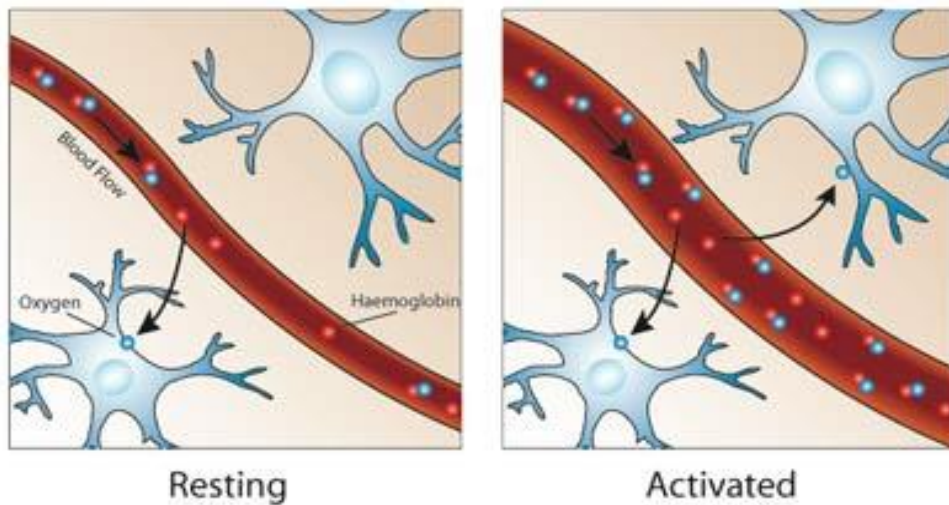


Figure 3. Diagram of the BOLD effect - Courtesy of Stuart Clare, FMRIB

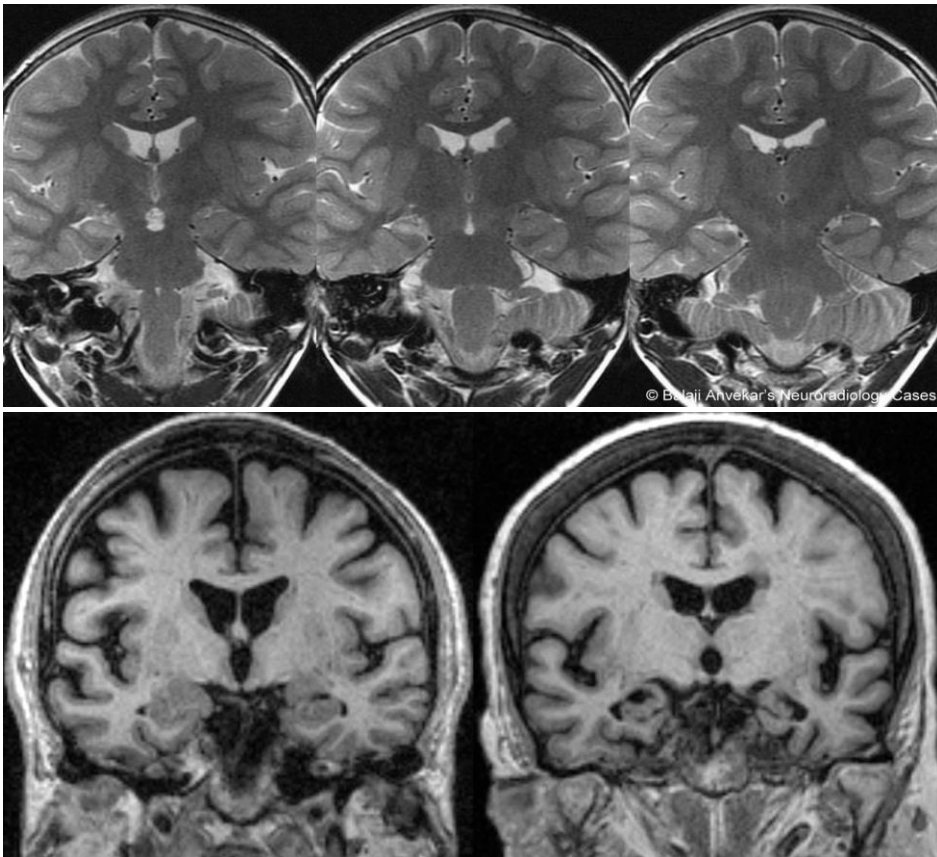


Figure 4. MRI scan of the hippocampus

The magnetic resonance imaging scans of the brain clearly show that in both images the right brain's hippocampal regions are bigger in volume and size than the left ones.

Research suggests that depression doesn't spring only from simply having too much or too few of certain brain chemicals. Rather, depression has many possible causes, including faulty mood regulation by the brain, genetic vulnerability, stressful life events, medications, and medical problems. It's believed that several of these forces interact to bring on depression.¹⁵ It is possible to identify three dimensions of depression: there are biological,

psychological and social factors that may cause it.

Knowing this, it may be practically impossible to prevent it entirely, but it may be possible to improve some conditions of our brain to reduce the risk of becoming depressed.

Biologically speaking, it is known that three main chemicals in the brain called (monoamine) neurotransmitters through which neurons interact in the synapse are involved in depression: they are serotonin (5-HT) (5-hydroxytryptamine), norepinephrine (NE) (noradrenaline) and dopamine (DA).^{15,16}

A neurotransmitter is a chemical substance which is released at the end of a nerve fibre by the arrival of a nerve impulse and, by diffusing across the synapse or junction, effects the transfer of the impulse to another nerve fibre, a muscle fibre, or some other structure. A monoamine is a type of neurotransmitter that is believed to have a crucial role in arousal, emotion and cognition.

Serotonin (5-HT) enhances peace and joy; if its level becomes low the subject may feel anxious, nervous, sad or angry. It also may cause an individual to have obsessive thinking, compulsive repetitive actions or an uncontrolled appetite.

Norepinephrine (NE) plays an important role in emotional stability and mental focus. If this neurotransmitter is in short supply, low mood accompanied by sleep difficulties and inability to have focused

concentration occurs. On the contrary, too much of this chemical can lead to anxiety, restlessness or hyperactivity.

Dopamine (DA) is essential to the learning and memory of pleasure and physical movements; low levels of DA could result in addictions and cravings, disordered physical movement and spatial navigation, and muscle twitching.¹⁷

Although low levels of these neurotransmitters do not always imply depression, together with other factors they can cause depression. Therefore, we will study if through aerobic exercise training it is possible to maintain these neurotransmitters in an adequate quantity, which would mean a lower probability of becoming depressed.

These chemicals' levels will be measured in every individual of each group. Firstly, to measure the amount of serotonin a blood sample test will be done. Blood is most often drawn from a vein. The vein usually used is on the inside of the elbow or the back of the hand.

To check the norepinephrine level, it is necessary to analyze the subject's urine because this neurotransmitter is excreted in the urine. Because levels vary over the day, 12-hour or 24-hour urine collections are used. Norepinephrine can be normalized by creatinine excretion to adjust for body size. To adjust for body size, results for norepinephrine are reported as micrograms of norepinephrine per gram creatinine of urine excretion. Determinations can be made with high-pressure liquid chromatography.

For dopamine, the last of the three main neurotransmitters believed to be involved in depressive illnesses, a PET scan will be used. A brain Positron Emission Tomography scan uses a radioactive drug called a tracer to show brain activity. The tracer may be injected, swallowed or inhaled. The tracer collects in areas that have higher levels of chemical activity, which show up as bright spots.¹⁸

The candidates will be asked not to eat anything for 8 hours before the procedure. They will be able to drink water. Before the scan, the tracers will be administered through an IV in the arm. As it is necessary to absorb them, a one hour waiting time is required. After this time, the subjects will be ready to undergo the scan.

Researchers are exploring possible links between sluggish production of new neurons in the hippocampus and low mood. Evidence from recent studies suggests that if neuronal plasticity in the hippocampus and in the limbic system is strengthened, the risk of developing a depressive disorder becomes lower.¹⁹

Antidepressants regulate levels of some chemicals such as NE and 5-HT, but it is believed that their real value may be in generating new neurons, a process called neurogenesis that consists of strengthening nerve cell connections and improving the exchange of information between nerve circuits.²⁰

These antidepressant drugs increase the amount of neurotrophic BDNF and its receptor TrkB in our brain; this brain-derived neurotrophic factor

neurotransmitter stimulates neurogenesis and increases synaptic numbers.²¹

Glucocorticoids (GCs) are steroid hormones secreted by the adrenal glands when signals from the hypothalamus are received. They are powerful anti-inflammatory compounds because of their ability to inhibit all stages of the inflammatory response. Natural GCs are produced by the cortex of the adrenal gland. The adrenal glands are organs located immediately above our kidneys.²²

On the contrary, stress stimulates the cortex of the adrenal gland and makes it release these natural glucocorticoids.

They are steroid hormones that by reducing BDNF level, reduce hippocampal neurogenesis and impair synaptic plasticity.^{23, 24}

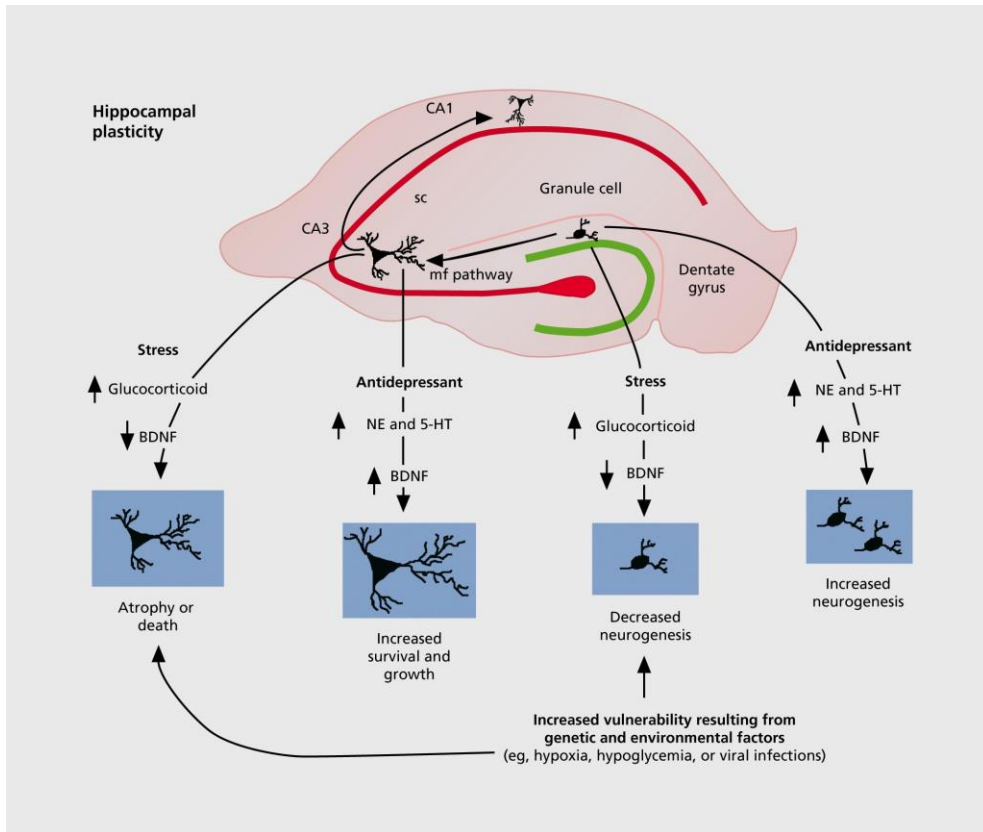


Figure 5. Hippocampal plasticity³⁸

From: Ronald S. Duman (2004)

Furthermore, a psychological test will be done on every subject to study his or her current mood. There are many types of tests and scales used in the clinical field, but this study will use the Hamilton Depression Scale (HDS or HAMD). It is used to assess the grade of depressive symptoms in both children and adults. The main advantage of this method, and concretely this questionnaire, is that it also permits us to measure if there is a possibility for the individual to become depressed. In other words, it is really useful for determining the level of depression before, during, and after a treatment. This will permit knowing not only their current mood but also their risk of

falling into depression in the near future.

Discussion

The present experimental design sought to study and demonstrate whether aerobic exercise training enhances certain cognitive functions vital for the realization of quotidian tasks and also for our well being. This investigation also aimed to show not only the importance of exercise in terms of improving our mental capacity and intellectual health, but also its ability to condition our mood and even our state of mind.

After the aerobic exercise-training period specified before, we expect the following incidents to happen.

The main physical change that we predict is the significant increase in volume and size of the hippocampal formation as a result of stimulation of neurogenesis.

This physical change will have a considerable repercussion on the functions that are regulated by the hippocampus.

It has been proved in several studies that aerobic physical exercise influences the central dopaminergic, noradrenergic and serotonergic systems.²⁵

Hence, we do expect to have these neurotransmitter (serotonin, norepinephrine and dopamine) levels increased in group 2 and 3 subjects, the

ones asked to exercise aerobically for 3 months. Group 1 won't undergo any changes in their monoamine levels.

The amount of these chemicals in every individual after the exercise-training period will be measured with the techniques specified before and compared with the first tests.

We expect the amount of norepinephrine to increase in groups 2 and 3. In the sedentary group 1 this neurotransmitter level will stay invariable.

It is believed that NE, also called noradrenaline, is involved in memory retrieval and in hippocampus plasticity.^{26,27}

Retrieval or recall in memory refers to the subsequent re-accessing of events or information from the past, which have been previously encoded and stored in the brain. It is what is commonly known as "remembering." This retrieval consists in replays of a pattern of neural activity that was originally generated in response to a particular event or information.

It has been demonstrated, not in humans but in mutant mice, that adrenergic signaling is critical for the retrieval of intermediate-term contextual and spatial memories, but not necessarily for the recall or consolidation of emotional memories in general. The role of noradrenaline in retrieval requires signaling through the β 1-adrenergic receptor in the hippocampus.²⁷

Regarding its relation to the hippocampus, it is believed that NE is an essential modulator of memory through its ability to regulate synaptic mechanisms in the hippocampal region. This catecholamine activates both

pre- and post- synaptic adrenergic receptors, which means an increase of neurogenesis in the hippocampus. It enhances its synaptic plasticity and consequently this neurotransmitter positively affects memory.²⁸

Dopamine level is also expected to increase with exercise. It is thought that DA significantly benefits memory consolidation²⁹ and it has been shown that dopamine plays a central role in the hippocampus and that its lack can cause several psychiatric diseases.³⁰

Aerobic exercise will also increase serotonin levels in the body. Recent evidence suggests that serotonin can enhance the production of new neurons in the hippocampus via activation of the 5HT1A receptor. Studies have predominantly focused on the role of serotonin in mediating the structure of the adult brain, but actually more recent research has found that the 5-HT molecule might have structural effects (stimulate neurogenesis) on the brain both during development and in adulthood.^{31,32} We expect this to happen particularly in the hippocampal region of our subjects of groups 2 and 3, due to their aerobic exercise training.

In conclusion, these three main neurotransmitter levels that are released during aerobic exercise will increase and, thus, the risk of becoming depressed or having some psychiatric diseases will decrease. What's more, all of them are involved in hippocampus functions, and particularly in cognitive functions so, through the enhancement of neurogenesis in the hippocampal area, we expect long-term memory and memory consolidation to improve in group 2 and 3 subjects.

It is known that stress is significantly reduced with the practice of exercise

and that it reduces common stress-induced deteriorations in health.³³

Another way to increase hippocampal neurogenesis, apart from the effects of the three neurotransmitters mentioned before, is through the reduction of stress. Stress is responsible for the release of glucocorticoids: the more stress that is accumulated, the more glucocorticoids that will be released. These hormones inhibit the production of BDNF (Brain-derived neurotrophic factor), which is a protein encoded by the BDNF gene.³⁴ This neurotrophin increases neurogenesis in the hippocampus, thus enhancing structural and functional plasticity and, therefore, the growth and development of this organ.³⁵ It is believed that, because of the relation between stress, glucocorticoids, BDNF, neuroplasticity and the hippocampal formation, GCs are considerably involved in memory consolidation and memory retrieval.³⁶

Thus, we expect the practice of physical aerobic exercise to enhance certain memory processes like memory consolidation, retrieval and long-term potentiation.

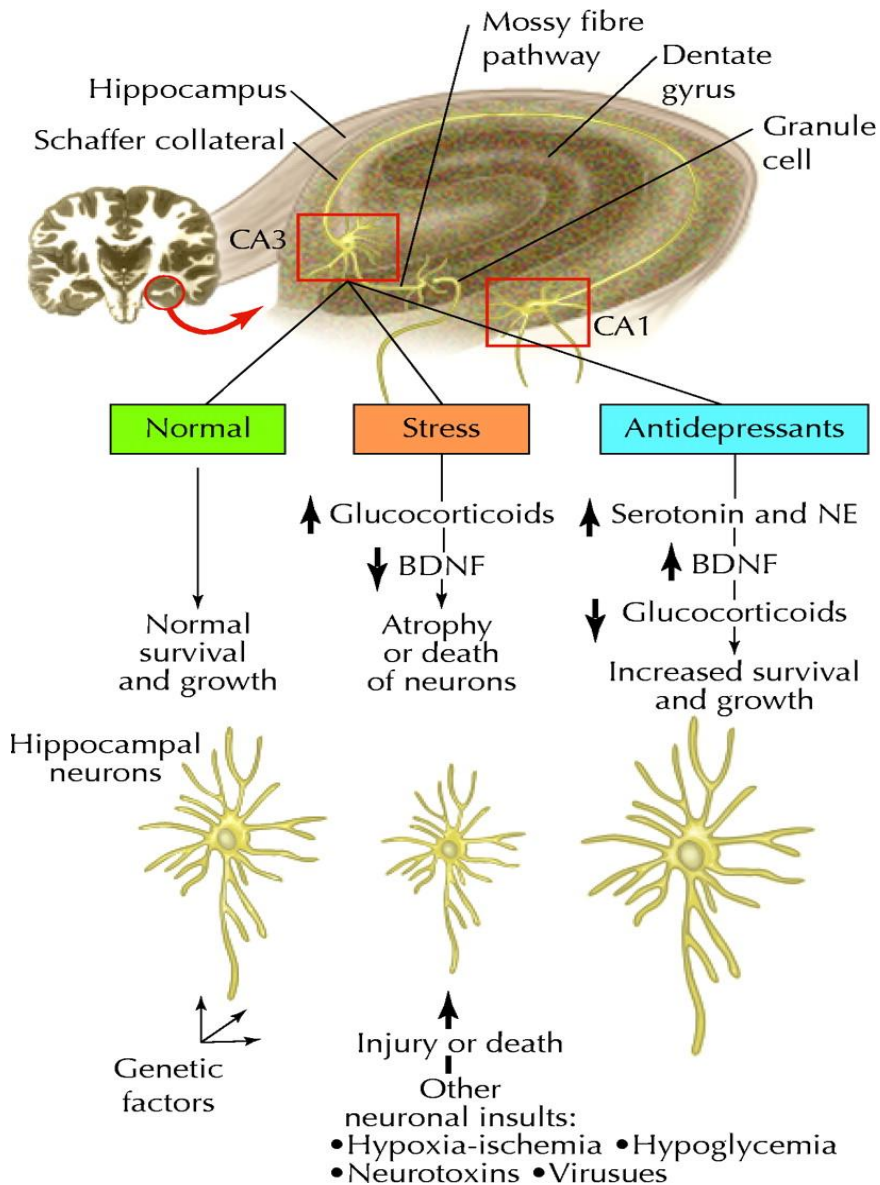


Figure 6. Hippocampus neurogenesis factors

Every group will be tested again with the Memory Span test, which was a method that allowed us to know their current short-term memory capacity.

We are predicting that, even though most neuroscientists don't believe

hippocampus volume and neurogenesis improvement affects short-term cognition, groups 2 and 3, who were asked to exercise, will show minor improvements in their short-term cognitive functions, perhaps not because of the hippocampus but because the increase of neurogenesis in its surroundings may improve this cognition capacity. Evidently, group 1 won't show any improvement.

All groups were asked to memorize a simple but long piano piece before the exercising period and since then they have only had one week in the middle of the three months to practice it again. Now, they will have to play what they do remember without the score. We expect that groups 2 and 3 will remember more of the score and that they will play it with a better technique than group 1 individuals.

Our thought is that aerobic exercise enhances hippocampus functions through stimulating its plasticity and through the release of certain neurotransmitters. Group 2 and 3 subjects will improve their long-term memory, memory consolidation and its phenomenon called long-term potentiation (LTP), which is involved in the capacity to acquire a new skill.

Supposing that the first time they all practiced playing the piano they had the same skill in playing this instrument (null), the only moment left where a difference could be made between groups asked to exercise (2 and 3) and group 1 in memorizing and consolidating this information is in the practice session done in the middle of the three month period. This difference will be due to an improvement in learning and memorizing. We think that as groups 2 and 3 will have done aerobic exercise for more than one month, their memory consolidation and LTP capacity will be enhanced for the second

piano practice session. If at this point their memory consolidation capacity and their capacity to acquire a new skill are better than group 1 subjects thanks to the aerobic exercise, they will memorize more and with a better technique the piano piece they were asked to learn. On the contrary, group 1 won't have enhanced their LTP and memory consolidation for the second piano training session, so their capacity to learn this score will be lower. For all that, in the final piano test, groups 2 and 3 will perform better than group 1 the piano piece practiced the same amount of time by every group.

The last memory test they were asked to do was the one which measured their long-term memory capacity itself, not focusing on long-term potentiation like the piano test. In this measurement we will focus on memory retrieval and their long-term memory capacity.

Once the exercise-training period is finished, they will have to do this test again. They will be asked to memorize a large list of random words for two minutes and when these 2 minutes are finished, 5 minutes without allowing them to study the words will pass. In the meantime, the subjects will undergo an fMRI and will be prepared to be scanned. When the 5 minutes waiting period is finished, they will have to recall the list of words memorized before while their hippocampal area activation and work are scanned.

We expect that groups 2 and 3 (exercised subjects) will show better memory retrieval and long-term memory capacity whereas group 1 individuals won't show any major improvement in their cognitive capacity. We think exercised subjects will have a greater part of their hippocampal region activated with a

stronger neuronal signal thanks to the aerobic exercise effects while the sedentary subjects will show a weaker and shorter activation of the hippocampus. This would mean groups 2 and 3 achieved a significant improvement of their long-term memory retrieval and capacity.

In the long-term memory test these changes will result in groups 2 and 3 having better test scores and results while the non-exercised group's subjects won't show any important differences on their scores.

On the other hand, in the Hamilton Depression score test, we expect group 1 to have the same scores as their first mood test. The results of the HDS test in group 2 subjects, who weren't initially depressed, will indicate that their risk of becoming depressed or having a mental illness is lower whereas the results of group 3, made up of initially depressed people, will show that their grade of depression is milder or that they have recovered from depression.

These improvements in mood state and depression test scores will be due to the exercise-induced neurogenesis of the hippocampal formation. In addition, the increase of the neurotransmitters involved in depression and released during exercise will be another cause of this enhancement. Therefore, groups 2 and 3 from aerobic exercise will have enhanced their hippocampal formation plasticity and also will have maintained or increased their DA, 5-HT and NE neurotransmitter levels, which contributes to a stable and better emotional state.

Conclusions

It is usually thought that exercise only has to do with being in good shape, but that doesn't take into account all its neurological and psychological benefits. Through research and studies in this important yet mainly unknown topic, many irreversible diseases which have no cure, caused by the atrophy of our brain, could be prevented. If we study how exactly memory processes work and the specific role of neurogenesis and the different neurotransmitters to improve the hippocampus function, a huge step forward in neuroscience and even in all of medicine would be achieved. One aim of this proposal was to study the boundless connections and relations that exist between the activities we execute, the chemical substances released in our bodies and the negative and positive effects of them. There is more than one way to obtain a certain effect or improvement, so everything we do possibly concerns many psychological and biological aspects of the human body. Therefore, although it is not generally thought that practicing sports may affect our cognitive performance, we believe exercise has a relevant role in our psychological health because of all the connections that have recently been discovered and yet remain to be further investigated.

As the proverb says, "Prevention is better than cure," and if through aerobic exercise we can not only prevent disorders but also develop our hippocampal formation, and so enhance our cognitive capacities and state of mind, what are we waiting for?

References

1. Edward F. Pace-Schott (epace-schott@mgh.harvard.edu)
2. Suzuki Wendy (wendy@cns.nyu.edu)
3. Kirk I. Erickson, Michelle W. Voss, Ruchika Shaurya Prakash, Chandramallika Basak, Amanda Szabo, Laura Chaddock, Jennifer S. Kim, Susie Heo, Heloisa Alves, Siobhan M. White, Thomas R. Wojcicki, Emily Mailey, Victoria J. Vieira, Stephen A. Martin, Brandt D. Pence, Jeffrey A. Woods, Edward McAuley and Arthur F. Kramer (2010) Exercise training increases size of hippocampus and improves memory; vol. 108 (7)
4. Frank-Gerald Pajonk, Thomas Wobrock, Oliver Gruber, Harald Scherk, Dorothea Berner, Inge Kaizl, Astrid Kierer, Stephanie Müller, Martin Oest, Tim Meyer, Martin Backens, Thomas Schneider-Axmann, Allen E. Thornton, William G. Honer, Peter Falkai, (2010). Hippocampal plasticity in response to exercise in Schizophrenia; *JAMA Arch Gen Psychiatry*, vol.67 (2), 133-143.
5. Kirk I. Erickson, Ruchika S. Prakash, Michelle W. Voss, Laura Chaddock, Liang Hu, Katherine S. Morris, Siobhan M. White, Thomas R. Wojcicki, Edward McAuley, and Arthur F. Kramer (2009) Aerobic fitness is associated with hippocampal volume in elderly humans; *Hippocampus*, vol. 19 (10), 1030-1039
6. Kuljeet Singh Anand and Vikas Dhikav (2012) Hippocampus in health and disease; *Annals of Indian Academy of Neurology* vol. 15 (4), 239-246
7. "Hippocampus" (2014) *The anatomical Travelogue*, www.thevisualmd.com

8. Tulving E. and Markowitsch HJ (1998) Episodic and declarative memory: Role of the hippocampus; *Hippocampus* Vol. 8(3), 198–204.
9. Brenda Milner and Suzanne Corkin (1968) *Neuropsychologia* Vol.6 (3), 215-234.
10. John P. Welby (welby.john@mayo.edu)
11. Richard H. Hall (1998) “Long Term Potentiation”
12. Jesús Pujol, Pere Vendrell, Carme Junqué, Josep L. Martí Vilalta and Antoni Capdevila (1993) When does human brain development end? Evidence of corpus callosum growth up to adulthood; *Annals of Neurology* vol.34 (1), 71-75.
13. Abi Berger (2002) Magnetic Resonance Imaging; *BMJ: British Medical Journal* Vol.324 (7328)
14. David J. Heeger and David Ress (2002) What does fMRI tell us about neuronal activity; *Nature Reviews Neuroscience* vol.3, 142–151
15. "What Causes Depression? *Harvard Health*. Harvard, 9 June 2009
16. Nutt DJ (2008) Relationship of neurotransmitters to the symptoms of major depressive disorder; *J Clin Psychiatry* vol.69
17. Bernarda Zenker “Neurotransmitters and depression” *Trinity Integrative Medicine*, 24-25
18. “Brain PET Scan: MedlinePlus Medical Encyclopedia.” *U.S National Library of Medicine*
19. Rashmi Nemade, Natalie Staats Reiss and Mark Dombeck (2007) Biology Of Depression – Neurotransmitters; *Mental Help Disorders and Issues*

20. "What Causes Depression? - Harvard Health." *Harvard Health*. Harvard University
21. C Anacker, PA Zunszain, A Canttaneo, L A Carvalho, M J Garabedian, S Thuret, J Price and C M Pariante (2011) Antidepressants increase human hippocampal neurogenesis by activating the glucocorticoid receptor; *Molecular Psychiatry* vol. 16, 738–750
22. Stephanie Lou (2010) "Glucocorticoids" *web.stanford.edu, hopes*
23. Robert M. Sapolsky (2000) "Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders; *JAMA Arch Gen Psychiatry*, vol. 57 (10), 925-935
24. R.M. Sapolsky, H Uno, C.S: Rebert and C.E. Finch (1990) Hippocampal damage associated with prolonged glucocorticoid exposure in primates; *The Journal of Neuroscience* vol. 10 (9) 2897-2902
25. Meeusen R. and De Meirleir K. (1995) Exercise and brain neurotransmission; *Sports Med* vol. 20 (3) 88-160
26. Hiroshi Katsuki, Yukiitoshi Izumi, Charles F. Zorumski (1997) Noradrenergic regulation of synaptic plasticity in the hippocampal CA1 region; *Journal of Neurophysiology* vol.77 (6)
27. Charles F. Murchison, Xiao-Yan Zhang, Wei-Ping Zhang, Ming Ouyang, Anee Lee, Steven A. Thomas (2004) A Distinct Role for Norepinephrine in Memory Retrieval; *Cell* vol. 117 (1), 131-143
28. Keith Tully and Vadim Y. Bolshakov (2010) Emotional enhancement of memory: how norepinephrine enables synaptic plasticity; *Molecular Brain* vol. 3 (15)
29. E. Coulthard, R. Bogacz, D Tsivos, L Hasson, A. Whone (2012) "109 The role of dopamine in memory consolidation: evidence from

- patients with Parkinson's disease" *J Neurol Neurosurg Psychiatry* vol. 83 (1)
30. Bruno Giros (2014) Dopamine, hippocampus and psychiatric diseases: Clarifying their relationships; *Biological Psychiatry, Douglas Mental Health University Institute*
31. Watanabe Y., Sakai R.R., McEwen B.S., Mendelson S. (1993) "Stress and antidepressant effect on hippocampal and cortical 5-HT_{1A} and 5-HT₂ receptors and transport sites for serotonin" *Brain Res* vol. 615, 87-94
32. Elizabeth Gould (1999) Serotonin and Hippocampal Neurogenesis; *Neuropsychopharmacology* vol. 21, 46-51
33. Jonathon D. Brown, Judith M. Siegel (1988) Exercise as a buffer of life stress: A prospective study of adolescent health; *Health Psychology* vol.7 (4), 341-253
34. Mark A. Smith, Shinya Makino, Richard Kvetnansky and Robert M. Post (1995) Stress and Glucocorticoids affect the expression of Brain-derived neurotrophic factor and Neurotrophin-3 mRNAs in the Hippocampus; *The Journal of Neuroscience* vol. 15 (3), 1768-1777
35. Suri D., Vaidya V. A. (2013) Glucocorticoid regulation of brain-derived neurotrophic factor: relevance to hippocampal structural and functional plasticity; *Neuroscience* (239), 196-213
36. Benno Roozendaal (2002) Stress and Memory: Effects of Glucocorticoids on Memory Consolidation and Memory Retrieval; *Neurobiology of Learning and Memory* vol. 78 (3), 578-595
37. Majid Fotuhi, David Do and Clifford Jack, (2012) Figure 2: Pathways leading to hippocampal growth or atrophy. Modifiable factors that

alter the size of the hippocampus with ageing; *Nature Reviews Neurology* Vol 8; 189-202

38. Ronald S. Duman, (2004) Figure 5: Hippocampal plasticity; Neural plasticity: consequences of stress and actions of antidepressant treatment (Pharmacological aspects) *Dialogues Clin Neuroscience*; vol 6; 157-169

Acknowledgements

This project required a great deal of work, research and dedication. Still, fulfillment would not have been possible without the support of many individuals. I would like to extend my sincere gratitude to all of them. I give my sincere appreciation to the following who never ceased in helping until this paper was completed:

Professor George Gescheider, Emeritus Professor of Psychology at Hamilton College, Ph.D. University of Virginia.

Hui Ling Wang, Program Management Lead at Pioneer Academics

Luc á Murillo Arroyo, International Representative at Pioneer Academics

Andrew Tonsing, Program Facilitator at Pioneer Academics

Saul Trejo, Program Facilitator at Pioneer Academics

The Pioneer Writing Center

I would also like to express my gratitude to the following professors and researchers for their kind cooperation and encouragement as well as for their helpful advice:

Edward Franz Pace-Schott, M.S., M.A., Ph.D.
Assistant Professor of Psychiatry, Harvard Medical School Associate
Researcher, Psychiatric Neuroscience.
Department of Psychiatry, Massachusetts General Hospital.

Amar Sahay, Ph.D.
Principal Faculty, Harvard Stem Cell Institute
Assistant Professor, Center for Regenerative Medicine, Psychiatric and
Neurodevelopmental Genetics Unit.
Center for Human Genetic Research, Massachusetts General Hospital,
Richard B. Simches Research Center.

Emery N. Brown, M.D., Ph.D.
Professor of Computational Neuroscience and Health Sciences and
Technology. Department of Brain and Cognitive Sciences
MIT-Harvard Division of Health Sciences and Technology

Wendy Suzuki, Ph.D
Professor of Neural Science and Psychology in the Center for Neural
Science, New York University. Suzuki Lab

John P. Welby, Ph.D.
Research Assistant Mayo Clinic

Mi-Hyeon Jang, Ph.D., MS, BS.
Assistant Professor of Biochemistry/Molecular Biology and Neurosurgery,
Mayo Clinic.
Research Associate Johns Hopkins School of Medicine

Robert Stickgold, Ph.D.
Associate Professor of Psychiatry, Harvard Medical School
Associate Professor, Department of Psychiatry, Beth Israel Deaconess
Medical Center.
Director at Center for Sleep and Cognition



The Dynamics of Movie Selection: A Study of Social Contagion Using the Revised SI Model Based on Word of Mouth

Xiangting Chen

Author background: Xiangting grew up in China and currently attends The Affiliated High School of South China Normal University in Guangzhou. His Pioneer seminar program was titled "Networks: from Searching the World Wide Web to Slowing the Spread of Infectious Diseases."

Abstract: Information related to movies, as a kind of popular culture information, spreads easily in social networks. Most of the previous studies of the movie industry focused on the influence of word of mouth, which is the passing of information about goods and services through oral communication among consumers, on box-office income. However, few of them model the dynamic processes of consumers' movie selection. In this study, we revise the classical *SI model* for epidemics to imitate such processes, by considering the potential moviegoers as the susceptible (S) and the actual moviegoers as the infectious (I). In order to examine the awareness effect and persuasive effect, which is caused by the volume and valence of word of mouth, respectively, we further divide the actual moviegoers into the Positive-Infectious (I+) and the Negative-Infectious (I-) based on whether their experiences of watching a particular movie met their expectations. Through simulations under different conditions, we observe and trace the dynamic changes in the population of moviegoers, as well as the number of positive-infectious reviewers and negative-infectious

reviewers for each time step. Simulation results show that the model can effectively explain some basic phenomena in the movie market.

Keywords: SI model; social contagion; word of mouth; box-office income

1 Introduction

There are clear connections between the spread of infectious diseases and the diffusion of new ideas within a community, which is often referred to as “social contagion.” The two processes have similar structural mechanisms in that they both require one to pass the “infection” to another. However, the nature of the two processes is very different. With the spread of infectious diseases, individuals cannot choose to be infected or not infected. Thus, being infected by pathogens is a passive process, which does not fully depend on the will of the individual but rather on the properties of the pathogen and the network structure of the population. With social contagion, however, individuals are constantly making decisions about whether to adopt new ideas or not. Therefore, the assimilation of new ideas is an active process which not only depends on the nature of the idea (i.e., whether it’s appealing or not) but also depends on the character and predilection of each individual (i.e., how the individual feels about the idea).

The distinction between the nature of the spread of epidemics and social contagion results in different research approaches. Probabilistic models based on random processes in networks, such as the *SI model*, the *SIR model*, and the *SEIR model*, are employed to study basic qualitative issues in the spread of infectious disease, like the synchronization in infection and the concurrency in the transmission of diseases. However, in the context of

social contagion, most studies focus on how new behavior and ideas spread from one person to another through social networks (e.g., the diffusion of innovations studied by sociologists). Researchers often assume that individuals' actions are interdependent, and use information cascades, network effects, and rich-get-richer dynamics to model the process of diffusion (Easley and Kleinberg, 2010).

This paper is interested in studying the process through which the idea of watching a movie diffuses within a community. Previous studies related to the movie industry indicated that Word of Mouth (WOM), which is the passing of information about products and services through oral communication among consumers, may strongly affect an individual's movie selection (Austin, 1989; Bayus, 1985; Faber and O'Guinn, 1984; Neelamegham and Chintagunta, 1999; Liu, 2006). Most studies tended to explore the influence of WOM on box office revenues, and seldom modeled the dynamic process of consumers' movie selections.

However, in the case of selecting a movie, the underlying decision-making processes are complex and dynamic. On one hand, features of the movie, such as the presence of stars and a sufficient budget, would influence the actions of moviegoers, and thus, affect box-office revenues. On the other hand, critiques or reviews of the movie released into social networks may also influence the decisions of the consumers. Sometimes, the behavior of family members or friends should also be taken into consideration, as they may exert certain degrees of influence on the individuals around them. Moreover, because the price of a movie ticket is relatively low, watching a movie is often an impulsive action. Therefore, potential moviegoers may not need meticulous considerations, during which they evaluate a large amount of WOM information before they decide to watch a movie (Liu, 2006).

When the decision-making processes of the individuals are hard to model, probabilistic models abstracted as random are more useful in studying them (Easley and Kleinberg, 2010). Therefore, by revising the classical *SI model* for epidemics base on the theory of word of mouth, this study proposed an innovative method to model the dynamics of movie selection.

We first divide the population into two groups, with the potential moviegoers as the susceptible (S) and the actual moviegoers as the infectious (I). The process of converting a potential moviegoer to an actual one is simplified to a random process, where the probability depends on the features of the movies. Then, we further divide the infectious (I) into the positive-infectious (I^+) and the negative-infectious (I^-). The differences in the attitudes of the moviegoers are underscored, because word of mouth is especially influential to consumers' behavior. In social networks, individuals can not only acquire word-of-mouth information from others around them, but can also pass on related information to others, thereby affecting their decisions. Thus, the probabilities which determine individuals' behavior are considered carefully in this study to simulate the real-world situations facing moviegoers.

The paper is organized into the following sections. In section 2, we introduce the theories and models related to this study. In section 3, we propose the revised *SI model* based on the theory of word of mouth. The methods of setting the initial conditions and parameters are introduced in section 4. The simulation process is also specified in section 4. Section 5 includes the simulation results under different conditions. Meanwhile, we compare the simulation results with real-world data. Conclusions and limitations are presented in the last section.

2. Relevant theories and models

2.1 Epidemic models

The process of the spread of infectious diseases is difficult to capture fully. The underlying obstacles in modeling the establishment and the spread of viruses mainly arise from the fact that the process has many interacting components, which are hard to isolate and examine individually. Factors like the size of the population that is exposed to the pathogen, the properties of the pathogen, the environment in which the pathogen and the host are situated in, and the structure of the population which is affected, may all combine to influence the power of the disease.

In the beginning of the twentieth century, scholars started to devise mathematical models to study the dynamic processes of the spread of epidemics. Since 1927, Kermack and McKendrick have devised multiple classic epidemic models, including the *SI*, *SIR*, and *SIS* models, with basic compartments such as Susceptible (*S*), Infectious (*I*), and Removed (*R*) (Kermack and McKendrick, 1927). Recently, studies on the dynamics of infectious diseases have been fruitful. Other compartments, such as Exposed (*E*) and Passively Immune Infants (*M*), are added to investigate diseases with more complicated properties and to provide more accurate simulations.

All of the epidemic mathematical models are based on the theory of probability, using calculations to determine each individual's status. Because human beings are passive participants in an epidemic outbreak (i.e. they do not have full control over their own physical well-being), their behavior can be randomized using probabilities. All of the factors related to the spread of

epidemics can be combined and quantified into the probability of being infected by viruses.

Figure 1 illustrates the *SI model* for epidemics. The model assumes that individuals in the community are in either group *S* or group *I*. Thus, the total population, denoted *N*, is equal to the sum of the number of people in group *S*, and group *I*. Group *S* includes individuals who are currently healthy but are susceptible to infection (susceptible), and group *I* contains individuals who have already been infected and thus have the ability to infect susceptible individuals around them (infectious). Because the *SI model* assumes there is no recovery, it's a simplified way of modeling diseases which are still unable to be cured completely through medication and other therapies, such as AIDS, leukemia, and ALS (Kermack and McKendrick, 1927).



Figure 1. Flow diagram of the *SI* mathematical model for epidemics.

Because the number of individuals in the two compartments may fluctuate over time, they can be presented with functions of time (*t*): $N(t)$, $S(t)$, $I(t)$. Since the dynamics of birth and death are omitted in this simple compartmental model, the relationship between the total population, the susceptible, and the infectious can be expressed using equation 1:

$$N(t) = S(t) + I(t) \quad (1)$$

Moreover, susceptible individuals are assumed to be infected at the infection rate β , if at least one of their neighbors are in the I group.

Using the above description, a simulation model based on random networks can be constructed to describe the dynamics of the spread of epidemics. In the following, $rnd()$ denotes a function that returns a random number between 0 and 1.

During simulation processes, in a random network with n nodes and k edges, the network evolves as follows:

- *For a random network with n nodes and k edges:*
- *For each time step:*
- *For every node:*
- *If at least one of the node's neighbor is infectious, if $rnd() < \beta$, the node becomes infectious.*

2.2 Word of mouth in the movie market

In the area of new product development, WOM is especially important because awareness and recognition of the public must be built, and consumers need information when deciding to purchase a product they are not familiar with (Mahajan, Muller, and Kerin 1984). Meanwhile, WOM differs from other information sources, in that it's usually perceived as a more credible and trustworthy one, and it's relatively easily accessible through social networks (Banerjee, 1992; Brown and Reingen, 1987; Murray, 1991). A McKinsey & Company study found that 67% of the sales of consumer goods are based on WOM (Taylor, 2003).

Two characteristics of the movie industry contribute to WOM's influence over moviegoers' actions (Liu, 2006). First, as a kind of popular culture

product, movies tend to receive great public attention and interest, both before and after their release. Therefore, active interpersonal communications about the topic are expected to exist. Since WOM information is more conveniently accessed through the internet, according to the theory of information accessibility, it's sensible to suggest that WOM influences the audience easily. Second, the intangible and empirical nature of movie consumption makes it difficult to judge a movie's quality. Whether moviegoers enjoyed the movie or not largely depends on that individual's own preferences. When the alternatives are difficult to evaluate before purchase, consumers often engage in WOM interactions to gather more information (Bristor, 1990; Harrison-Walker, 2001). Accordingly, moviegoers also actively participate in WOM interactions to evaluate the worth of watching a movie.

In addition, extant studies have found that volume and valence are among the most important attributes of WOM. Volume measures the total amount of WOM interactions. Some evidence showed that volume of WOM is strongly correlated with consumer behavior and market outcome (Anderson and Salisbury, 2003; Bowman and Narayandas, 2001; Van den Bulte and Lilien, 2001), and can effectively boost consumers' awareness of a product (Godes and Mayzlin, 2004). In the movie market, volume results in the *awareness effect*: the greater the volume, that is, the more conversations there are about the movie, the more likely a potential moviegoer will be to hear about it.

Valence captures the nature of WOM messages, and it can greatly influence consumers' attitudes toward a certain product. In the movie market, positive WOM typically gives a direct or an indirect recommendation of the movie, while negative WOM may involve privately complaining to or giving

denigratory reviews to the movie. Valence results in the *persuasive effect*: positive WOM enhances expected quality, and, thus, consumers' attitudes toward a product, whereas negative WOM reduces it. In the movie market, however, there is no sufficient evidence to support the postulation that positive WOM enhances consumer expectation and vice versa. Nevertheless, the valence of WOM may still cast some extent of influence on consumers' behavior (Liu, 2006).

3 The revised *SI* model based on word of mouth

From the perspective of marketing, there are only two kinds of costumers: the ones who have bought the product and the ones who have not yet bought it. The same applies to the movie market: there are only the individuals who have seen the movie and those who have not yet seen it. Naturally, to model the dynamics of the diffusion of movie information in social networks, the *SI* model for movies places individuals into two basic compartments: the susceptible (*S*) and the infectious (*I*), which represent the potential moviegoers and the moviegoers, respectively.

In modeling the spread of movie information among individuals, two effects may influence individuals' decision as they shift from group *S* to group *I*: the *awareness effect* and the *persuasive effect* (section 2.2). Since the awareness effect is caused by the volume of WOM and the persuasive effect is caused by the valence of WOM, it's necessary to take both valence and volume into consideration.

As a result, we propose the revised *SI model* based on WOM to describe the dynamics of movie selection (Figure 2). We innovate the classical model by dividing group *I* into two sub-compartments: the Positive-Infectious (*I+*), which represents the moviegoers who give positive WOM and the Negative-

Infectious (I^-), which represents the moviegoers who give negative WOM. In reality, though there are individuals who produce neutral WOM or give no comments, we overlook them in our model because they essentially cast no influence on others.

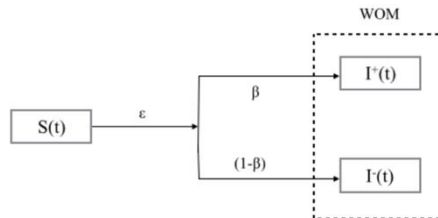


Figure 2. Flow diagram of the SI mathematical model for Movie based on WOM.

In this model, we assume that there is a social network consisting of N individuals. Each of them may have chances to communicate with their neighbors on a random basis. In the *SI model* for movies, the relationship between group S , I^+ , I^- , and N can be expressed with functions of t (time) using equation 2, because the number of individuals in each compartment may fluctuate over time.

$$N(t) = S(t) + I^+(t) + I^-(t) \quad (2)$$

In addition, we assume that susceptible moviegoers become “infected” with the probability of ε ($0 < \varepsilon < 1$). There are two critical factors to determine the value of ε . First is the number of friends around an individual who give positive WOM. Generally, it is believed that friends cast significant influence on an individual’s behavior. Thus, when individuals in group I^+ recommend the movie to their friends who have not seen the movie during

daily communications, the potential moviegoer's chance of watching the movie may increase significantly.

Second is the total number of individuals in the community who have seen the movie. This is important because the social benefits of talking about a popular subject tend to be greater than talking about a less popular one (Brown and Reingen, 1987). In the movie market, potential consumers also tend to discuss popular movies, and they can use the box-office incomes as indicators of popularity: when the box-office income of a certain movie is high, it may be a hit at the moment and many moviegoers may have seen it. This phenomenon may galvanize potential moviegoers to go to the cinemas, because after watching the movie, they can better join the heated discussion.

Therefore, when determining the value of ε , we need to consider the collective power of all the actual moviegoers in the population, as well as the influence of "positive-infected" individuals on their friends. And, we quantify the two factors by taking the ratio of infected individuals to total population, and the ratio of "positive-infectious" friends to total infectious friends around an individual. The value of ε may be different for every individual in the community. Because we assume that potential moviegoers consider the recommendations from friends more credible and trustworthy, WOM from friends is valued more. Thus, the value of ε is the weighted average of the two ratios, and the influence from neighbors (friends) is weighted twice the influence from the public. Combining the two factors together, we can derive equation 3:

$$\varepsilon = \frac{\left(\frac{I(t)}{N(t)} + \frac{I^+(t)_{Neighbor}}{I(t)_{Neighbor}} \times 2 \right)}{3} \quad (3)$$

In addition, an individual joins group I^+ with probability β ($0 < \beta < 1$), and conversely joins group I^- with probability $(1 - \beta)$. After a moviegoer sees the movie, we cannot decide for him or her whether he or she like it or not, because we could not postulate each individual's preference (i.e., we cannot calculate a β for each individual). Thus, in the model, we simplify the individualized process and base the value of β on four important features of movies, and the value varies from movie to movie but is the same for every individual. The features include star effect (SrE), movie genre (Ge), critical reviews from professional critics (CR), and user reviews posted on movie-reviewing websites (OLR).

Firstly, as they are the frequent topics of interpersonal communication, the power of stars in the movie market is undeniable. Thus, the presence of popular stars in a movie is likely to make the movie more preferable. Secondly, certain movie genres are proved to be more popular than others in terms of box-office income, such as Adventure and Comedy (*Most Popular Movie Genres*, Nielsen). Thirdly, critical reviews published on magazines or online also reflect the quality of a movie. However, some may argue that, because WOM comes from the public, it may be perceived as more trustworthy and can better reflect popular taste than critical reviews do (Holbrook, 1999). Thus, Critical reviews may only serve as a complementary information source, and their effect should not be weighted heavily. Lastly, user reviews posted online are often seen as the most authoritative information source to potential moviegoers during their movie selections. When combined together, these four features can let us, to the largest extent, predict the valence of WOM of a certain movie.

By taking the average and weighting the features differently according to their importance, we can derive equation 4:

$$\beta = 0.3 \times SrE + 0.2 \times Ge + 0.1 \times CR + 0.4 \times OLR \quad (4)$$

4 Simulation

4.1 Initial condition

Using the randomized mathematical method described in section 3, we can build a simulation model to describe the dynamics of movie selection.

First, we build the network, assuming that there is a medium-sized city where every resident has at least one friend, and that in this city there is a random social network consisting of 100,000 nodes (i.e., 100,000 individuals) and 100,000 edges (i.e., 100,000 social ties). We broadly define a social tie between two individuals as a relationship that requires regular communication, which includes the relationship between friends, family members, work mates, and schoolmates. Therefore, individuals are frequently exposed to WOM information. We also assume that all individuals are susceptible at the beginning.

Next, we establish the initial condition of the network. Let k be the percentage of the initial moviegoers in the total population. In this study, we discuss the effects of a successful publicization, such as advertisements, social media, and grand movie premieres, on the total amount of moviegoers or on the aggregate box-office income. In addition, we set the value of k to 0.05 and 0.1 in this study, because the effectiveness of the publicization can be reflected by the amount of initial moviegoers. As a result, we assign $100,000 \times k$ individuals to group $I+$ or $I-$ randomly based on the value of β . Therefore, the network initially has $100,000 \times k$ moviegoers, with

probability β of being assigned to group $I+$ and probability $(1 - \beta)$ of being assigned to group $I-$, using the following algorithms:

For node $i = 0$ to $i = 100,000 \times k$:

(a) Take randomly a value, rnd , between 0 and 1.

(b) If rnd is less than β , assign node i to $I+$. Else, assign node i to $I-$.

4.2 Parameter β

In section 3, we introduce the method used to calculate the value of β for each movie based on its features, including star effect, movie genre, critical reviews, and online user reviews, and the data are collected from various public sources. For star effect, we consider a movie, in which both the male and female leading and supporting actors and actresses are well-known by the public, to have a SrE of 1 ($0 < SrE \leq 1$). One-over-four points is deducted from SrE each time a main character fails to appear in “The Power List,” a star list published annually by the magazine *Premier*. Take *The Dark Knight* (2008) for example: the leading actor Christian Bale, leading actress Maggie Gyllenhaal and supporting actor Heath Ledger all appeared in “The Power List.” Additionally, because the movie has no supporting actress, its SrE can be seen as 1.

As for movie genre, we base the value of Ge ($0 < Ge \leq 1$) on the statistical report, *Most Popular Movie Genres*, from Nielsen (<http://www.nielsen.com>), a company which measures consumers’ selection of goods and services around the world. The report says that 50 percent of moviegoers indicate that Comedy is their favorite type of movie to see in a theater. Information about the movie genre is gathered from the *Internet Movie Database*

(<http://us.imdb.com>). Take *Home Alone (1990)* for example: being a comedy, its value of Ge is 0.5.

For critical reviews and online user reviews, we extract the data from the popular movie review website *Rotten Tomatoes* (<http://www.rottentomatoes.com>). The website contains average ratings given by *Tomatometer*, a group of professional critics from newspapers, magazines, and the movie industry; and online users. *Tomatometer's* ratings are scored from 0 to 10, which are converted into CR ($0 < CR \leq 1$). The marks given by online users are scored from 0 to 5, which were converted into OLR ($0 < OLR \leq 1$). Take *The Dark Knight (2008)* for example: its average rating from the *Tomatometer* is 8.6, making its value of CR 0.86. Its average rating from online users is 4.4, making its value of OLR 0.88. Combine the above calculations together, and we can get the specific value of β for each movie.

However, for the sake of simplicity, during simulation, we directly set β to three different values, 0.1, 0.6, and 0.9. These three values represent an unattractive movie, a median movie, and a good or attractive movie, respectively. If β is large (i.e., an attractive movie), more individuals will be assigned to $I+$, and vice versa.

4.3 Simulation process

Next, we define the model's evolution process. While the parameter ε , the probability of becoming a moviegoer, is constantly updating during simulations and is calculated differently for each node, the known parameters we are going to use as inputs are:

- k , percentage of initial audience in the total population;
- β , the probability of a moviegoer giving positive WOM.

According to the statistical data gathered from the *Top Movies in terms of box office income* from 1982 to 2014 (www.boxofficemojo.com), The average amount of time a movie is shown in cinemas is about 30 weeks. Therefore, the simulation is carried out for 30 steps, with each step representing a week. However, movies with smaller β (i.e., are less popular or with more negative WOM) might experience much shorter life spans, usually within 10 weeks, than movies with a higher β .

To ensure that the results can reflect the mean value, we run the simulation for 1,000 times and take the average of the results. For each simulation, we create a new random graph with 100,000 nodes and 100,000 edges, ensuring the density of the graph is unchanged. In the following, $rnd()$ denotes a function that returns a random number between 0 and 1.

Thus, based on the above descriptions, the network evolves as follows:

- *For 1,000 times:*
- *Generate a gnm random graph with 100,000 nodes and 100,000 edges, for each graph:*
- *For 30 time steps:*
- *For every node:*
 - *If at least one of the node's neighbors is infectious, if $rnd() < \varepsilon$, the node becomes infectious. Else, the node stays susceptible.*
 - *Calculate the value of ε for each node. If $rnd() < \beta$, assign the node to $I+$. Else, assign the node to $I-$.*

5 Results and analysis

5.1 Scenario 1 – unattractive movie

We assume the release of a movie whose β equals 0.1. The value of β

indicates that the movie may be characterized by a lack of hit movie stars, a choice of an unpopular movie genre, a bad critical response, or an unpopular appeal to the public. Figures 3 and 4 present the simulation results

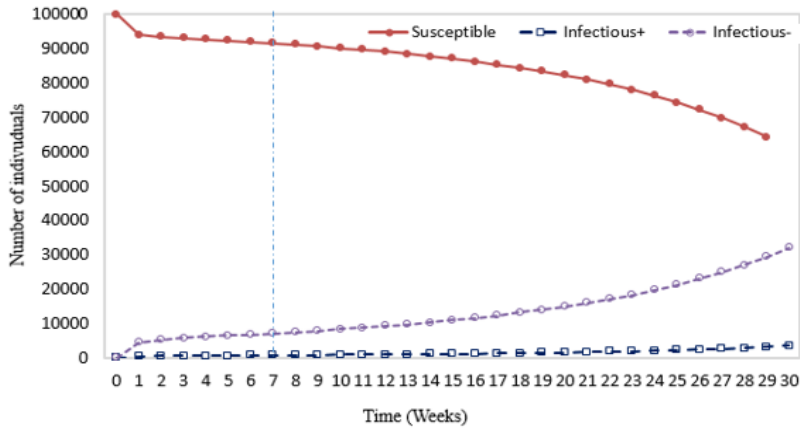


Figure 3. Simulation data with parameters $\beta=0.1$ and $k=5\%$.

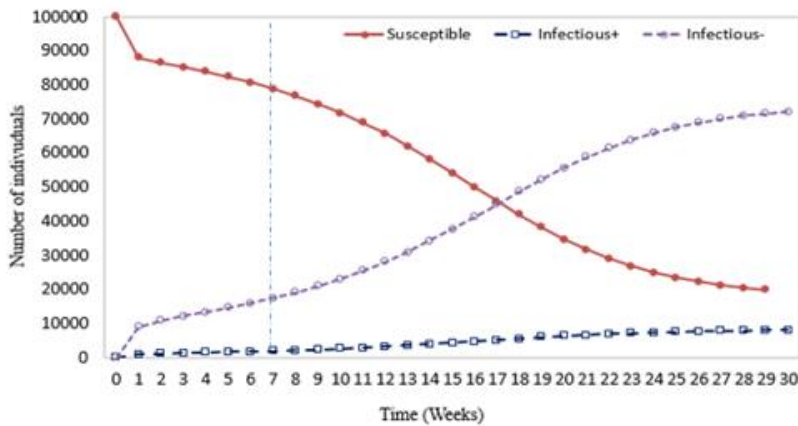


Figure 4. Simulation data with parameters $\beta=0.1$ and $k=10\%$.

The solid-dotted line represents the total population of susceptible, while both the square line and circle lines represent the accumulated population of Positive-Infectious and Negative-Infectious, respectively. The dotted vertical line helps to indicate the value of each group at week 7.

Figure 3 reports the case in which publicization is not very effective, with a low initial audience percentage of 5%. We acquire several facts from the figures. First, the number of positive-infectious individuals is clearly smaller than that of the negative-infectious individuals, which is in accordance with reality and common sense. Second, the rate at which potential moviegoers go to see the movie (i.e., the value of ε) is small. Since each individual's ε is determined by two parts: (1) the percentage of positive-infectious people around a susceptible individual (i.e. how many of your friends recommended the movie to you or gave positive reviews about it), and (2) the percentage of total infectious people in the population (i.e., how many individuals in the social network have seen the movie), the phenomenon presented in the figure can be explained by the fact that either your friends haven't seen the movie or they didn't enjoy it, or most people in the social network haven't seen it. These two conditions give no incentive to potential moviegoers and therefore cannot simulate a large audience.

By comparing Figure 3 to Figure 4, we have a general notion of the effects of a larger initial audience. With more initial audience members, around week 7, around 20,000 more individuals went to see the movie. However, we find that even a movie with $\beta = 0.1$ is capable of attracting a lot of audience members if it is shown in cinemas for 30 weeks, a phenomenon that is especially obvious when percentage of the initial audience is high. This may be understood as, as long as a movie continues to screen, susceptible individuals still have the probability of being "infected," no matter the quality of the movie. However, this is highly unlikely to occur in the real world, because as these movies attract very few moviegoers each week, they usually generate more cost than revenue. Cinemas would rather choose to screen other more attractive movies. Nevertheless, as is stated in

previous sections, the behavior of watching a movie is a complex decision-making process that is influenced by many factors, such as a sudden impulse. Thus, it's still possible that potential moviegoers will choose to view unattractive films.

5.2 Scenario 2 – attractive movie

We assume the release of a movie whose β equals 0.9. The value of β represents a movie, which may be characterized by the presence of many hit movie stars, a choice of a popular movie genre, a positive critical response, or a popular appeal to the public. Figures 5 and 6 present the simulation results.

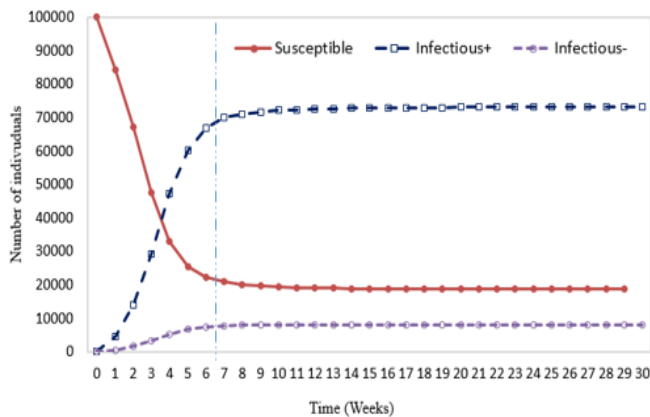


Figure 5. Simulation data with parameters $\beta=0.9$ and $k=5\%$.

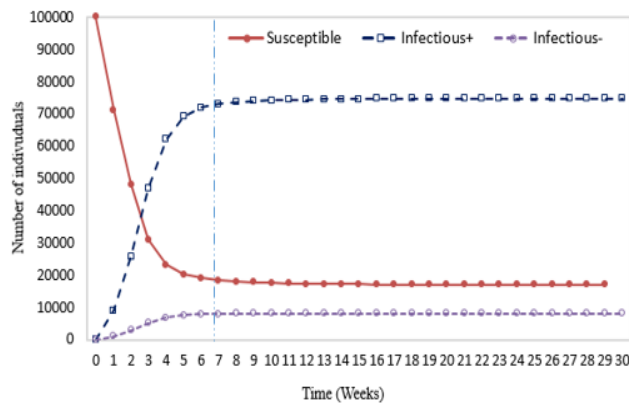


Figure 6. Simulation data with parameters $\beta=0.9$ and $k=10\%$.

The solid-dotted line represents the total population of susceptible, while both the square line and circle lines represent the accumulated population of Positive-Infectious and Negative-Infectious, respectively. The dotted vertical line helps to indicate the value of each group at week 7.

We can see from both figures that the number of “positive-infectious” individuals is more than “negative-infectious” ones. However, unlike what was presented in the results of simulation when β was equal to 0.1, the effects of the variation in the percentage of initial audience are largely diminished, though more initial audience members still speeds up the process of “infection” for about two weeks’ time. This comparison shows that WOM is more influential than the number of initial audience members. Because it took fewer than 8 weeks for the market to saturate, we can also infer that the value of ε is small, a phenomenon which can be explained by both the high percentage of an individual’s friends with positive WOM and the high percentage of total moviegoers in the population.

5.3 Scenario 3 – median movie

We assume the release of a movie whose β equals 0.6. The value of β represents a movie that reaches the median standard. Figures 7 and 8 present the simulation results.

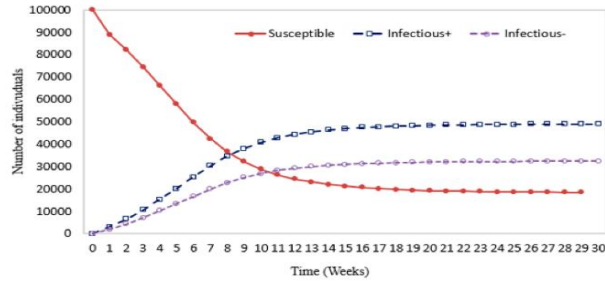


Figure 7. Simulation data with parameters $\beta=0.6$ and $k=5\%$.

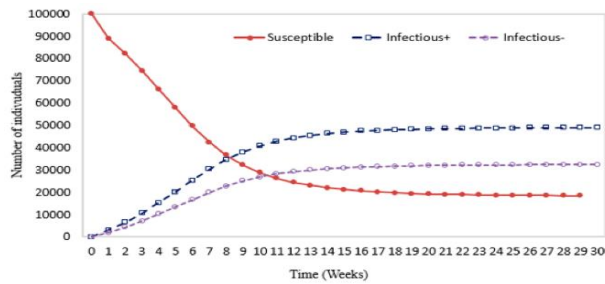


Figure 8. Simulation data with parameters $\beta=0.6$ and $k=10\%$.

The solid-dotted line represents the total population of susceptible, while both the square line and circle lines represent the accumulated population of Positive-Infectious and Negative-Infectious, respectively. The dotted vertical line helps to indicate the value of each group at week 7.

From both figures, we can infer some general information: first, the number of “positive-infectious” individuals is more than the negative-infectious; second, more initial audience members will influence the speed of “infection” by a little; third, the average market-saturation time for a median movie is about 17 weeks.

5.4 Comparison between simulation results with different values of β

We assume that three movies, with different values of β (0.1, 0.6, and 0.9), are released with the same percentages of initial audience members ($k=5\%$). The different values of β represent different qualities of the movies, and

thus, different WOM. The value of k indicates that the publicization of the three movies before release is not as successful as the scenario in which the percentage of initial audience members is 10%. Figure 9 presents the simulation results.

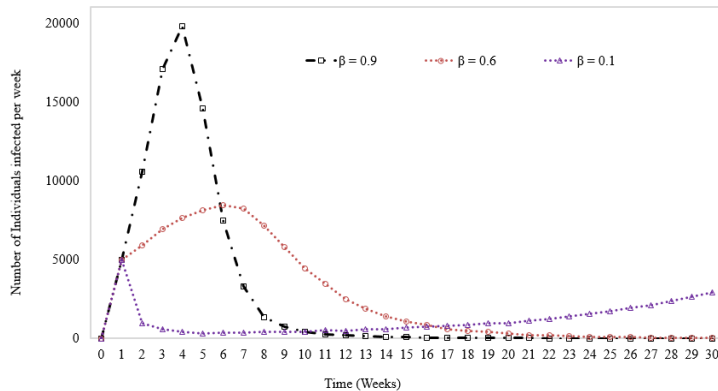


Figure 9. A comparison between three simulations with the same parameter $k=5\%$. The square, circle and triangle lines represent the simulation data when β equals 0.1, 0.6, and 0.9 respectively. Every data point represents the number of infectious individuals (both positive infectious and negative infectious) at each time step.

Under the same initial conditions, we perceive that, for a movie with $\beta = 0.9$, the number of moviegoers per week will peak around week 4 and the market will saturate around week 10. The simulation in which $\beta = 0.6$ produces results with a similar shape and trend. Interestingly, the simulation in which $\beta = 0.1$ presents an upward-sloped graph after week 10, which indicates that the number of infectious individuals per week is increasing as time goes by. This phenomenon can be explained by the fact that, for a movie with $\beta = 0.1$, the number of susceptible individuals is still rather large during week 9.

And, as long a movie is in the cinema and individuals are susceptible, they may still watch the movie even though the value of ε is small. This phenomenon is in accordance with the real world for the reasons stated in section 5.1.

5.5 Comparison between simulation results under different network densities

In section 3, we assumed that the random network in which we conduct simulations has 100,000 nodes and 100,000 edges, meaning that each individual in the community, on average, has one friend. However, since the density of the network (i.e. how many friends each individual has on average) may also affect the process of diffusion, we conduct simulations under different conditions to examine its influence. Assuming that the percentage of initial audience is the same ($k=5\%$), we simulated scenarios in which the average number of friends increases from 1 to 10, and recorded the amount of time needed for the market to saturate (i.e. the time steps needed for the number of susceptible individuals to remain the same). We also compared the differences between the attractiveness of the movie, with β equal to 0.1, 0.6, and 0.9. The results are plotted in Figure 10.

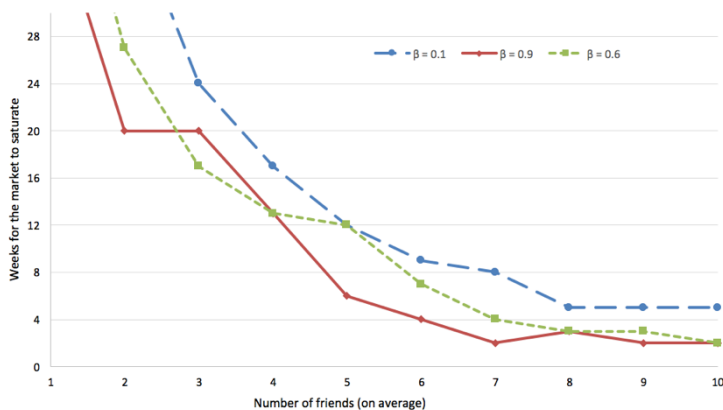


Figure 10. A comparison between three simulations with the same parameter $k=5\%$. The circle, square and diamond lines represent the simulation data when β equals 0.1, 0.6, and 0.9, respectively. The x-axis represents the number of friends, on average, an individual has, which ranges from 1 to 10. The y-axis represents the number of weeks needed for the market to saturate. Each data point represents the number of friends on average, and the corresponding time.

From the figure, we can infer that the number of friends is influential to the diffusion processes. The three lines presenting different attractiveness levels of the movies are all downward sloping, which means that, as the number of friends increases, the time needed for the market to saturate decreases. In other words, as an individual is connected to more people (i.e., more exposed to WOM information), he or she is more likely to be “infected.” In this way, the whole population will be stimulated to watch the movie with greater speed, shortening the time needed for the market to saturate. Surprisingly, for an unattractive movie with $\beta=0.1$, the time needed for the market to saturate is also decreasing as an individual has more friends. While this seems to deviate from reality, this phenomenon can be explained

by the fact that, as individuals are more interdependent in the clustered network, their influence on each other is magnified. Thus, even though the movie is unattractive, under the persuasion or influence of many friends, an individual may still watch the movie.

However, the quality of the movie still matters. From the figure, we can see that, as the the movie becomes more attractive (i.e. with higher value of β), the average amount of time needed for the market to saturate decreases, with the corresponding line shifting inward.

5.6 Comparison between simulation and real world

We then compared the simulation data with the real-world movie data gathered from the website *Box Office Mojo* (www.boxofficemojo.com), which posts weekly box-office income information, to show the effectiveness of the revised *SI Model* for movie selection based on WOM in predicting the behavior of the real movie market. Though the data collected contain weekly information about box-office incomes, they do not directly give us the number of moviegoers. This number is calculated by dividing the box office income by 8, which is approximately the average price of a movie ticket in America. The data plotted are the percentage of the number of “infectious” individuals per week in all moviegoers.

First, we compared our simulation results with real world results for the movie *Terminator 2: Judgment Day* (1991). We collected the movie’s features and converted them to calculate the value of β , which was approximately equal to 0.9. The movie was shown in cinemas for 20 weeks, with a \$201,270,098 Gross Box Office Income. Figure 11 contains the information about the total “infectious” individuals per week for both the simulation whose $\beta = 0.9$ and $k = 10\%$ and the real movie.

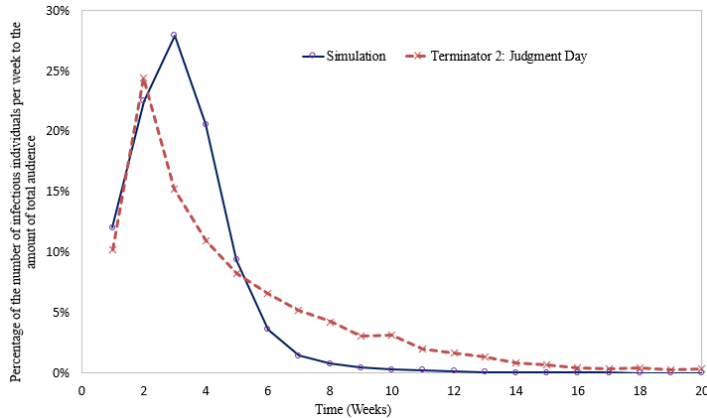


Figure 11. A comparison between the data of the simulation where $\beta=0.9$ and $k=10\%$ and the real world data of the movie *Terminator 2: Judgment Day* (1991). Every data point is the ratio of the number of infectious individuals per week to the total number of audience members.

Next, we compared our simulation results with real world results for the movie *Home Alone* (1990). We collected the movie's features and converted them to calculate the value of β , which was approximately equal to 0.6. The movie was shown in cinemas for 28 weeks, with a \$277,905,624 Gross Box Office Income. Figure 12 contains the information about the total “infectious” individuals per week for both the simulation whose $\beta = 0.6$ and $k = 5\%$ and the real movie.

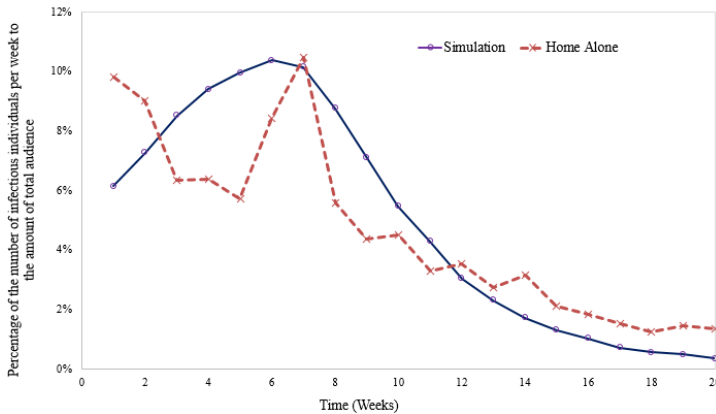


Figure 12. A comparison between the data of the simulation where $\beta=0.6$ and $k=5\%$ and the real world data of the movie *Home Alone* (1990). Every data point is the ratio of the number of infectious individuals per week to the total number of audience members.

Last, we compared our simulation results with real world results for the movie *Focus* (2015). We collected the movie’s features and converted them to calculate the value of β , which was approximately equal to 0.3. The movie was shown in cinemas for 10 weeks, with a \$53,862,963 Gross Box Office Income. Figure 13 contains the information about the total “infectious” individuals per week for both the simulation whose $\beta = 0.3$ and $k = 10\%$ and the real movie.

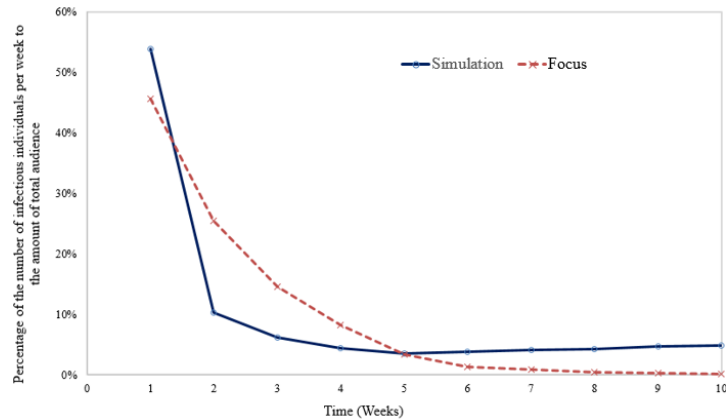


Figure 13. A comparison between the data of the simulation where $\beta=0.3$ and $k=10\%$ and the real world data of the movie *Focus* (2015). Every data point is the ratio of the number of infectious individuals per week to the total number of audience members.

From the results shown above, we can see the general diffusion patterns for different types of movies. For a relatively good movie, the number of moviegoers will rise drastically in a very short period of time and then decline quickly. For a relatively unattractive movie, the number of moviegoers will rise to a small peak and then decline, though it will increase as time goes by. As it may have multiple peaks and oscillations, the pattern for a median movie is more complex, and it is hard to generate a conclusion.

6 Conclusions and limitations

In order to model the dynamic process of consumers' movie selection in a community, we revise the classical *SI* model for epidemics based on two important attributes of word-of-mouth information, volume and valence, both of which can affect an individual's decision. We investigate the effects of volume by manipulating the parameter, k , which is the percentage of

initial audience members in the population. We find that, as the number of initial audience members increases, the time needed for every possible individual in the population to see the movie decreases.

We also investigate the effects of valence by setting the parameters ε and β , which represent the probability of going to the movies and the probability of becoming a positive-infectious individual after watching the movie, respectively. We determine the value of β based on the features and attractiveness of the movie, which directly influence the audiences' movie-watching experience, and indirectly affect the valence of WOM. As a result, the bigger the value of β , the more positive-infectious individuals and positive WOM information. Because the value of ε is partly determined by the number of positive-infectious individuals around a person, it can be influenced by the value β .

In addition, we examine how the density of social network affects the diffusion processes under different conditions. Simulation results indicate that, as individuals have more friends (i.e., the higher the density of social network), they're more likely to be "infected" (i.e., the larger the value of ε).

The revised *SI model* for movie selection is simple yet effective in modeling the complex process of social contagion. The simulation results show that the model can explain some basic phenomena and properties of the movie market.

However, this study still has many limitations. First, 1,000 times of simulation may not be sufficient to produce the average result. Thus, the method of Monte Carlo simulation should be used to ensure the accuracy of the result. Second, although the four movie features used to calculate the value of β are indeed important determinants of the quality of the movie, the

different weightings of each component are determined empirically. Third, we only run the simulations in a random network. In further studies, simulations should be carried out in small-world networks, which better represent the real world. Results can also be compared with those produced by the random network to examine the effect of network structure on the process of diffusion. Fourth, according to the network model proposed in section 3 and the setting introduced in section 4.1, we assume that the social network is an offline friendship network. Considering that the online social networking service (SNS) is an interesting topic to study, in future models, we should assume that individuals are influenced by both the WOM information posted online and the WOM information gained during offline communication. In this way, we can compare the differences in the influence of offline communications and SNS information on individuals.

Reference

- Anderson, and Linda Salisbury (2003), "The Formation of Market- Level Expectations and Its Covariates," *Journal of Consumer Research*, 30 (June). 115-24.
- Austin, Bruce (1989). *Immediate Sealing: A Look at Movie Audiences*. Belmont, CA: Wadsworth.
- Bayus, Berry (1985), "Word of Mouth: The Indirect Effect of Marketing Efforts." *Journal of Advertising Research*, 23 (3), 31-39.
- Bowman, D, and D. Narayandas (2001), "Managing Customer- Initiated Contacts with Manufactures: The Impact on Share of Category Requirements and Word-of-Mouth Behavior." *Journal of Marketing Result* 38 (August), 281-97.

- Bristor, Julia (1990). "Enhanced Explanation of Word of Mouth Communications: The Power of Relationships," in *Research in Consumer Behavior*, E.C. Hirschman, ed. Greenwich, CT: JAI Press. 51-83.
- Brown, Johnson and Peter Reingen (1987), "Social Ties and Word-of-Mouth Referral Behavior," *Journal of Consumer Research*, 14 (December), 350-359.
- Byrnerjee, Abhijit (1992). "A Simple Model of Herd Behavior," *Quarterly Journal of Economics*, 107 (3), 797-817.
- David Easley and Jon Kleinberg (2010), "Networks, Crowds, and Markets", Cambridge University Press.
- Faber, Ronald and Thomas O'Guinn (1984), "Effect of Media Advertising and Other Sources on Movie Selection." *Journal of Consumer Research*, 11 (4), 371-377.
- Godes, David and Dina Mayzlin (2004), "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science*. 23 (4), 545-60.
- Harrison-Walker, L. Jean (2001). "The Measurement of Word-of-Mouth Communication and an Investigation of Service Quality and Customer Commitment as Potential Antecedents," *Journal of Services Marketing*, 4 (1). 60 - 75.
- Holbrook, Morris (1999). "Popular Appeal Versus Expert Judgments of Motion Pictures." *Journal of Consumer Research*, 26 (September), 144-155.
- Kermack W O, McKendrick A G. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London A*:

Mathematical, Physical and Engineering Sciences. The Royal Society, 1927, 115(772): 700-721. Social Networking Sites.

Liu, Yong (2006), "Word of Mouth for Movies: Its Dynamics and Impact on Box Office Revenue," *Journal of Marketing* 70 (July), 74-89.

Mahajan. Vijay, Eitan Muller. and Roger Kerin (1984), "Introduction Strategy for New Products with Positive and Negative Word-of-Mouth," *Management Science*, 30 (12). 1389-1404.

Mizerski, Richard (1982). "An Attribution Explanation of the Disproportionate Influence of Unfavorable Information," *Journal of Consumer Research*, 9 (3). 301-310.

Murray, Keith (1991), "A Test of Services Marketing Theory: Consumer Information Acquisition Activities," *Journal of Marketing*, 55 (January), 10-25

Neelamegham, Ramya and Pradeep Chintagunta 1,1999), "A Bayesian Model to Forecast New Product Performance in Domestic and International Markets," *Marketing Science*, 18 (2), 115-36.

Taylor, John (2003), "Word of Mouth Is Where It's At." *Brandweek*, (June 2), 26.

Van den Bulte, Christophe and Gary Lilien (2001). "Two-Stage Partial Observability Models for Innovation Adoption," working paper, Wharton School, University of Pennsylvania.

Zufryden, Fred (1996), "Linking Advertising to Box Office Performance of New Film Releases: A Marketing Planning Approach," *Journal of Advertising Research*, 36 (July-August), 29-41.



Adolescent Decision Making: Neuroscientific Basis of Risk-Taking Behavior in the Developing Brain

You Jin Jung

Author background: You Jin grew up in South Korea and currently attends Korea International School in Seoul. Her Pioneer seminar program was titled “The Socioemotional Brain and Decision Making.”

Abstract

Adolescence is a period where different regions of the brain are still maturing, and is associated with risk taking and impulsivity. The connection between neurobiological factors of human brains—namely, gender and personality—and the tendency to take risks was studied in terms of the developing adolescent brain. Based on the prior research papers on adolescent neurobiological development that state that the emotion-based limbic regions mature at a faster rate than the impulse-managing prefrontal regions, I introduce possible explanations of individual differences in risk-taking in adolescents. Personality was researched under the assumption that the Big Five model has biological evidence supporting its theories. Additionally, gender studies used in this paper do not necessarily fall under adolescent studies, but from basic male and female adult brain differences, it may be assumed that such differences also manifest themselves in adolescent brains.

BACKGROUND

It is widely known that adolescents take risks comparatively more than young children or fully grown adults. Adolescents are associated with a

lack of impulse control, and evidence supports this association. There are over 13,000 adolescent deaths in the United States per year, 70% of them from vehicle crashes, homicide, and suicide. The 2005 National Youth Risk Behavior Survey suggests that adolescents are more likely to drive under the influence of alcohol, disregard seat belts, use drugs, and take part in unprotected sex (Casey et al. 2008). Because of the extreme statistics, it has been commonly assumed that adolescents are unable to properly judge themselves and danger. For the last 25 years, there has been much research done by developmental scientists, and the studies have eliminated several assumptions about adolescents. Adolescents are actually not “irrational or deficient in their information processing”; they have risk-reasoning abilities identical to those of a fully grown adult; they do not believe they are “invulnerable” nor are they “less risk-averse” (Steinberg 2008). In fact, they are prone to overestimating risks, akin to adults. However, questions still remain regarding the possibility of impulsivity control and development of the ability to manage spur-of-the-moment behavior.

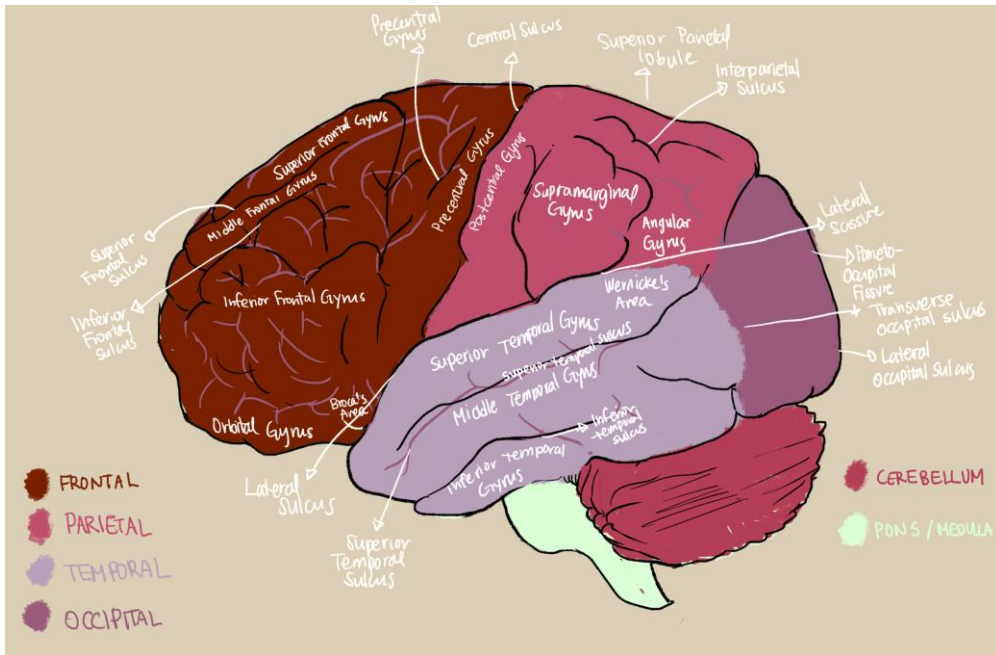


Figure 1-1 illustrated an exact division of the lobes to give more information about smaller divisions of the brain.

To understand impulsivity control and its development, it is essential to first understand the anatomy of the brain and its mechanisms. The cerebral cortex, a crucial portion of the brain required for various skills such as motion and reasoning, is divided into four “lobes”: frontal, temporal, parietal and occipital as shown in Figure 1, all of which play essential roles in the functioning of the brain. The frontal lobe, which will be discussed extensively later on in this paper, is responsible for judging, deciding, planning, problem solving, and movements. The parietal lobe is associated with the senses--movement and orientation--and contains the sensory cortex. The temporal lobe is linked to perceiving sound and making speech. Lastly, the occipital lobe is associated with vision. Other basic divisions include the new brain, mid brain and the old brain. The terms referring to age simply

indicate evolutionary age. The old brain initially evolved along with amphibians, while the new brain is the most advanced level, present in humans, dolphins and apes. The new brain is the aforementioned cerebral cortex, responsible for higher-order functions. On the other hand, the brain stem, medulla, thalamus and the cerebellum are all parts of the old brain. They regulate basic survival mechanisms, including breathing, sleeping and coordination of movement. More specifically, the thalamus, which acts as a sensory switchboard, receives sensory information from throughout the brain and transmits senses to the appropriate locations. Lastly, the midbrain is the limbic system, consisting of the hippocampus, amygdala and the hypothalamus. The hippocampus regulates memory, the hypothalamus maintains the body's homeostasis, and the amygdala is particularly associated with emotions of fear and aggression. Beyond this anatomical organization of the brain, there are bridges in the brain at the cellular level. Each brain cell, or neuron, transmits electrical signals along with the glial cells that help "maintain an appropriate chemical environment for neurons" (Purves et al. 2001). Between each neuron is a synapse, where neurotransmitters bind to receptors on the next neuron. This fundamental arrangement of the brain and the interactions between various components—both anatomically and chemically—are what allow organisms to make decisions for both survival and everyday life.

One of the primary developing regions of the brain is the frontal lobe in the anterior part of the brain. Increased activity in this area is considered evidence of cognitive control over impulsivity (Casey et. al 2008). Furthermore, organisms' decision-making is powered by a reward seeking, punishment avoiding system that seeks maximum gain and minimum risk. In

contrast with adults, adolescents have a high emotional reactivity which results in the imbalance between the control system and emotional system in the process of seeking novelty and sensations (Casey et al. 2008). This is especially demonstrated by adolescent animal research. Animals such as rats (Kellogg et. al 1989) are shown to share a pubertal change trend similar to that of humans in development while adolescent mice show more heightened novelty seeking than adult mice (Adriani et. al 1998).

The current model of reward processing in the brain focuses on adults, and little research has been done with adolescents. The Simon study found that, even when behavior seems similar across different ages, the brains respond differently, especially in the orbitofrontal cortex and the striatum, which is involved in learning, reward processing and movement (Simon et al. 2014). Schramm-Sapyta observed rats and concluded that organisms go through a distinctive developmental period and are vulnerable to disorders and addictions (Schramm-Sapyta et. al, 2009). In other rat studies, stress and gender were found to have significant impact on risk-taking behavior specifically related to forced swims and mazes. In the process of synapse pruning and neurons' myelination in the adolescent phase, the varied rate of maturation of emotional regulation and cognitive function may be the cause of the increased risk (Toledo-Rodriguez et al. 2011) and novelty seeking in periadolescent mice (Adriani 1998). Although this adolescent period may seem disadvantageous and inconvenient in everyday modern situations, it is indeed a naturally-selected and advantageous phase in many mammals. Adolescence allows an individual to acquire skills necessary for independence from the family. It is a period characterized by independence, novelty and sensation seeking, grouping

with similar age groups and conflicts with parents as a young individual strives to adapt to mating and learning new abilities (Spear 2000).

THESIS

Along with the theories stated above, I hypothesize that personality and gender predispositions strongly interact with the brain regions involved in decision-making during the natural maturation process of the brain. Statistics related to the amount of brain activity before monetary reward and during risk-taking between people of ages of 7 to 29 do suggest that there are some adolescents who take even fewer risks than adults or children (Galvan 2007), despite the general adolescent population's inclination towards risk-taking. To test this hypothesis, the adolescents must be grouped according to their level of impulsiveness. After the categorization, each group must be tested for their tendency to indulge in risk-taking behavior. Since within the context of this course, it is not possible to conduct original experiments, I summarize and analyze experiments conducted by other scientists, synthesizing diverse aspects of their research and supplementing it with my findings from other fields including psychology.

NEUROANATOMY AND NEUROCHEMISTRY

The developing brain goes through a myriad of changes at various levels--anatomically, synaptically and functionally. These changes can be detected by brain scanning devices, such as the MRI, DTI, and fMRI. MRI, Magnetic Resonance Imaging, is a test that uses a magnetic field and radio waves to generate pictures of organs, tissues and bones. The DTI, Diffusion MRI, creates a picture of the diffusion process of molecules in biological tissues. Lastly, the fMRI, the functional MRI, measures brain activity by

detecting changes in blood flow. Scans show that the gray matter and the white matter of the brain develop in different ways in adolescence (Casey et al. 2008), as shown in Figure 2. Gray matter is the exterior layer where cell bodies of neurons are located, and white matter is the portion containing myelinated fibers (Walker 2002). Their neural synapse connections continue to be fine-tuned throughout development, with gray matter gradually reducing in volume, and white matter expanding in volume (Squeglia et al. 2009).

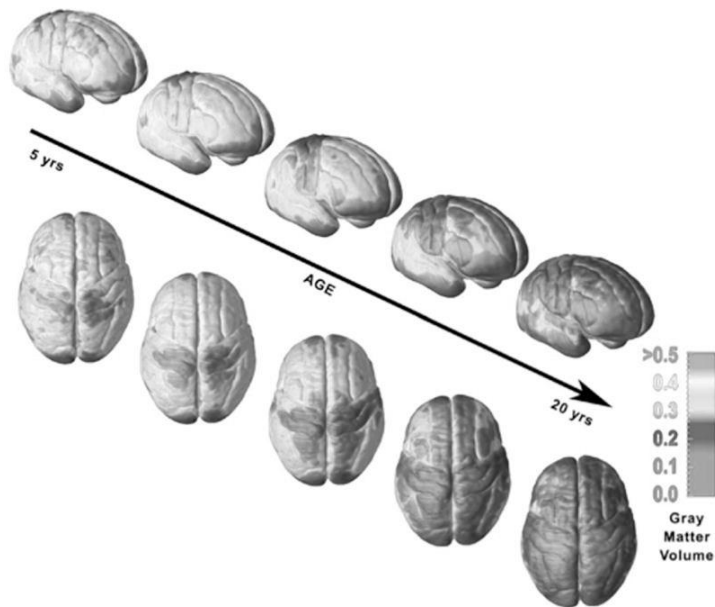


Figure 2- This graphic depicts the development of gray matter volume; it is clearly seen that the gray matter volume decreases with age.

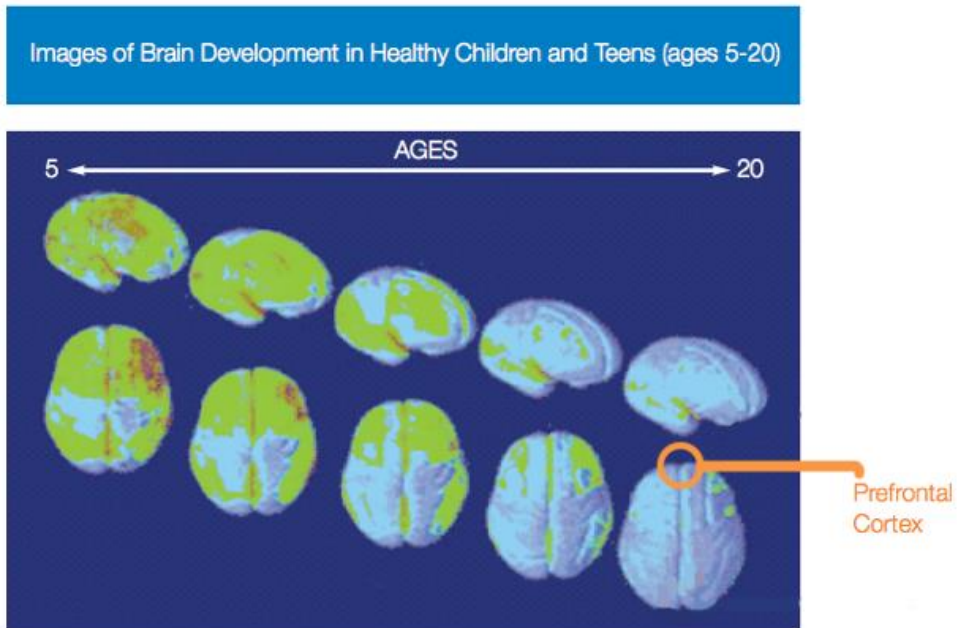


Figure 3 - This graphic shows the development of the prefrontal cortex in healthy child and adolescent brains from ages 5 to 20. The blue indicates the mature portion, and the green indicates the immature portion.

These developmental studies suggest a connection between volumes and the ability for cognitive management, indicating that they have the full ability to choose appropriate responses over unfavorable ones (Casey et al. 2008). This is proof that structural and physical transformations of the brain could be linked to cognitive changes, and emphasizes the influence of subcortical and cortical (prefrontal cortex) development on certain cognitive abilities. Its significance lies in the fact that changes in emotion-controlling subcortical regions (limbic system), such as the accumbens (part of the striatum that plays a big role in the reward system) or amygdala interact with the prefrontal cortex to produce behavioral changes in impulsivity and risk-taking. This major transformation from child to adult has been the subject of

many studies, and an important speculation has arisen: If only the immaturity of the prefrontal cortex alone was the core cause of increased risk-taking in adolescents, as many people commonly believed, then children with even more underdeveloped prefrontal cortexes should undergo an exacerbated risk-taking phase. However, that is not the case, which suggests a more complex developmental pattern of the brain. A critical ability undergoing development is to control impulses—disadvantageous thoughts and behaviors—for long-term gain. This ability to judge and make smart decisions is governed by the aforementioned prefrontal cortex (Casey et al. 2008), which develops last as indicated by Figure 3.

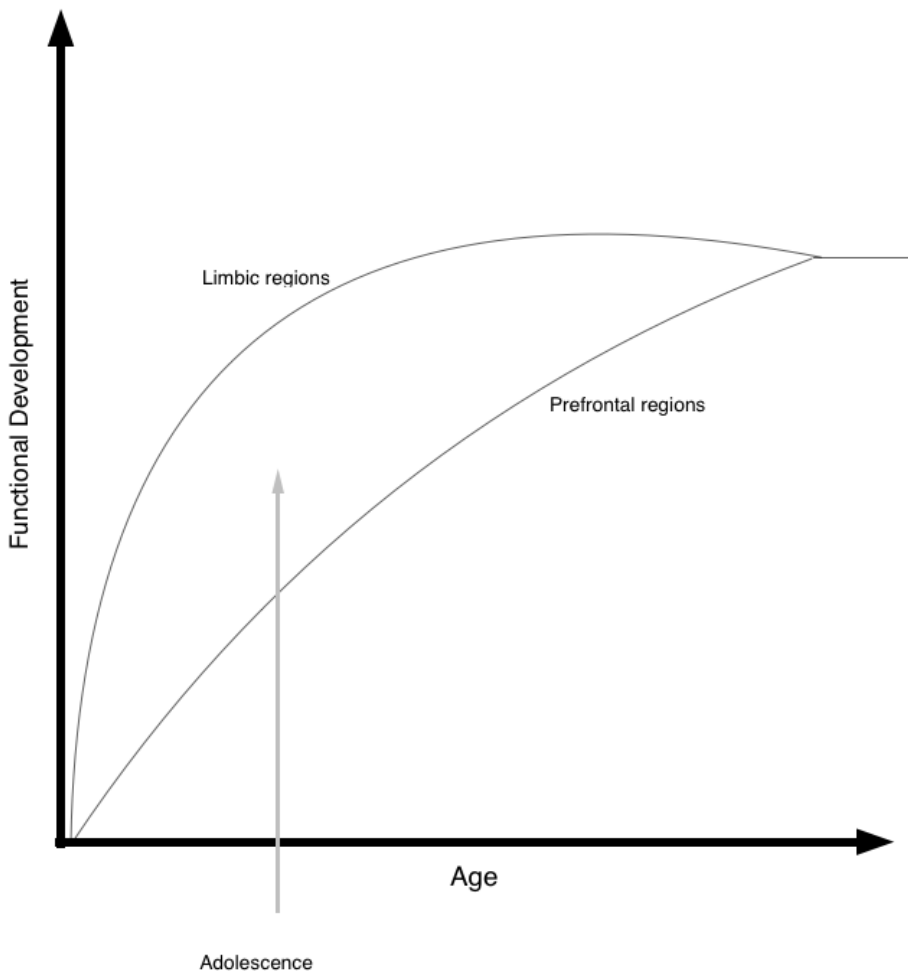


Figure 4- This figure is a graph illustrating the development of two different regions of the brain. The x-axis is a representation of age, and the y-axis indicates degree of development. The shaded area is where risk-taking is, theoretically, at the maximum level as the emotional regions—limbic region—develop faster than the impulse controlling prefrontal regions.

Figure 4 shows that the limbic system regions which encourage emotional decision-making develop a few years earlier than the prefrontal regions, and therefore an adolescent is influenced by the imbalance between the mature limbic system and the still-developing prefrontal regions. On the

other hand, an adult's impulse control system is fully mature along with the limbic system, so adults have some measure of control over impulsivity and responsiveness to rewards. Due to the disparity between developmental rates, adolescents show signs of seeking short-term gain and an aggravated accumbens response in excitement of reward (Casey et al. 2008). Additionally, damage to the amygdala results in the lack of an "autonomic response to reward and punishment" and the inability to "utilize cues" to direct future decision-making. There is also evidence that amygdala damage in men and women causes "deficits in decision making and social behavior" (Gupta et. al 2011). This result helps explain that the lack of development and the resulting instability of the amygdala may be one of the influential factors causing adolescents to often be more vulnerable to affective disorders and risk taking.

NEUROTRANSMITTER MECHANISMS

On top of that, another reason for adolescents' particular increase in risk-taking may be partly explained by the dopaminergic analysis of the brain. Dopamine, a neurotransmitter associated with cognition, reward seeking, and certain psychological disorders, is a vital part of decision-making in an organism's brain. Dopamine mainly originates from the striatum, which is connected to the frontal and limbic cortexes. The neurotransmitter dopamine behaves rather like a signal between "go"—associated with positive outcomes—and "no go"—associated with avoidance of punishment. This decision between go and no go is determined by the dopamine in the cortical region in the brain. The balance in the adolescent brain between cortical and subcortical dopamine systems transforms so that there are greater levels of dopamine in the cortical areas

during adolescence (Spear 2000). Furthermore, it is evidenced by primate studies that dopamine levels increase in the prefrontal cortex, whose neurons undergo an extended transformation throughout childhood and adolescence, interacting with other circuitry in the brain (Rosenberg & Lewis 1995). Moreover, there are two dopaminergic levels: one is a slow and gradual release of dopamine via D2 receptors on the striatal neurons, and the other is a series of dopamine firing, signaling both D1 and D2 receptors. D1 receptors have a low attraction to dopamine and allow one to learn to approach rewarding stimuli, whereas D2 receptors have a high attraction to dopamine and allow one to avoid negative outcomes (Clark & Dagher 2014). Another study suggests that, in adolescents, D1 over-expression leads to greater immediate self-satisfaction, and reduces D2 receptors, aggravating hedonistic tendencies (Sonntag et al. 2014).

Furthermore, “Prospect theory” by Kahneman and Tversky suggests that many people have loss aversion, which is the tendency to view possible losses as larger than possible gains (Clark & Dagher 2014). Two possible models of the mechanism that controls such risk taking consider whether choices are the result of “weighting of gains or losses along a single value axis” or from a “competitive interaction between separate systems for gains and losses” (Clark & Dagher 2014). FMRI evidence implies that there are “bidirectional” responses to losses and gains in the striatum and the prefrontal cortex (potential gains), the amygdala and insula (potential losses) (Clark & Dagher 2014). The loss of striatal dopamine neurons was linked with reduced risk-taking behavior in Parkinson’s disease patients (Brand et al, 2004), perhaps interfering with the judgment of potential gains.

Moreover, the aforementioned striatum plays a large role in adolescent decision making. A rat study by Sturman and Moghaddam hypothesized that the age-related changes influence the processing of key brain regions of adolescents in terms of dealing with rewards in learning and habit formation. Adolescents had a larger proportion of neurons in the dorsal striatum than adults, which further suggests that mechanisms of the brain may differentiate an adolescent's behavior from that of other age groups. In this study, adolescent and adult neural unit activity from the striatum was recorded as the subjects learned to associate an action with a reward. The task was performed in a box with holes in one wall and a food bowl on the opposite side. The experiment started as a light was shone through a center hole. If the rat poked into the hole, then the light turned off, then turned on above the food bowl, and food was given. When the rat ate the food, the light turned off again. As a result, adolescent rats were shown to exhibit more reward-anticipation neuron firing that lasted all the way to the reward, while adult neural units were activated earlier and returned to baseline before reward. Also, adults had much larger proportions of inhibited units, and adolescents had more activated units into the food bowl (Sturman & Moghaddam 2012). Figure 5 shows the increased activation in adolescents compared to adults.

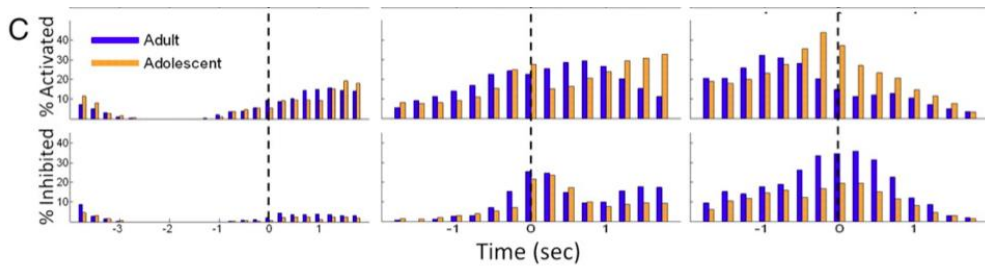


Figure 5- *This graph illustrates the percent of inhibited and activated neural striatum units in adults and adolescents. It can be clearly seen that adults have a higher percentage of inhibited units, and adolescents have a higher percentage of activated units.*

LIMITATIONS

Increased tendencies to take risks may be explained by the natural neuroscientific mechanisms' unbalanced pattern of development and an excess of activity of dopamine. However, there exist exceptions to the general fact that adolescents tend to take risks more than adults and children. Although cases of adolescents taking risks such as binge drinking, driving without seat belts, and having unprotected sex are highlighted over and over again in studies, media and daily life, there do not exist many studies that involve a large portion of adolescents that tend to abide by the rules and do not take risks. Figure 6 describes individual differences in decision-making of adolescents, even in a small scale study of binge drinking from ages 12 to 23. My goal is to identify some possible causes of that individual difference from a neuroscientific perspective.

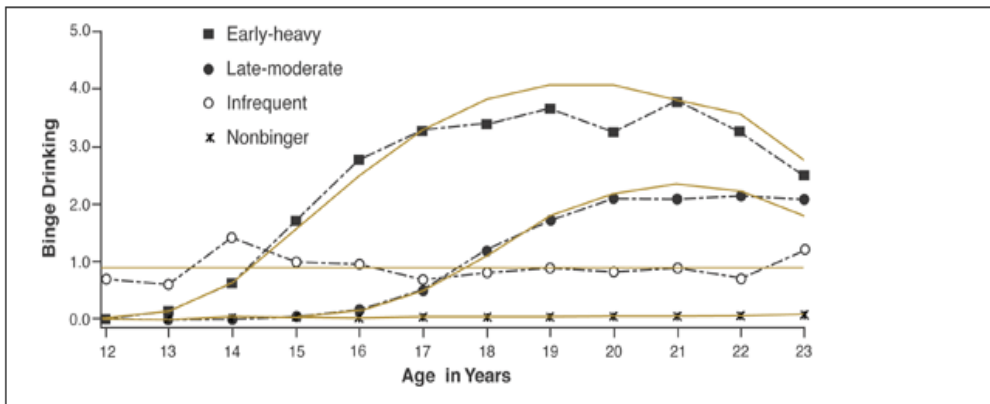


Figure 6- *The y-axis indicates the frequency of binge drinking, and the x-axis the age of students. It can be seen through this graph that there are certain types of risk takers in drinking: The “early-heavy” group tends to have the highest binge drinking count, and there is a group at the opposite end of the spectrum: “nonbinger.”*

In fact, as can be seen from Figure 6, many adolescents are non-risk-takers, and two of the major influential factors that can explain this puzzling phenomenon are personality and gender.

PERSONALITY

There is a neuroscientific basis for the manifestation of certain types of behavior: personality. In studying the individual-specific brain mechanisms behind personality, behavior, motivation, emotion and cognition must all be considered and measured. The field “Personality Neuroscience” focuses on the relatively consistent patterns of behavior and cognition: traits. In order to study personality neuroscience, a method of classifying personalities must be established first. One of the most frequently used models is the Five Factor Model, or the Big Five, which organizes

major traits into five “domains”: Openness, Conscientiousness, Extraversion, Agreeableness and Neuroticism. The Five Factor Model can be used in scientific studies because the model suggests that there is a shared underlying cause, possibly biological. A biological factor is a “reasonable hypothesis” to make, because individual differences in the Big Five traits are genetically impacted, and this genetic “factor structure” of the Big Five seems to be “invariant across European, North American, and East Asian samples” (DeYoung 2010). The Big Five domains can be associated with actual brain mechanisms. Broadly defined, Extraversion and Neuroticism are the demonstrations of sensitivity to reward and punishment; Agreeableness is associated with the tendency to be altruistic. Conscientiousness is the top-down control of impulses, and Openness is the tendency to utilize the information available.

Personality Dimension	Code	Description
Agreeableness	A	One's propensity to be trusting, altruistic, compliant, modest, genuine, and sympathetic toward others.
Extraversion	E	One's propensity to be gregarious, assertive, friendly, excitement-seeking, fast-paced, and high-spirited.
Openness to Experience	O	One's propensity to be intellectually curious, appreciative of novelty, interested in art, music, and beauty, and imaginative and creative.
Conscientiousness	C	One's propensity to be dutiful, orderly, self-disciplined, deliberate, and achievement striving.
Emotional Regulation (Neuroticism)*	R	One's propensity to be resistant to stress, anxiety, depression, anger/hostility, and self-consciousness.

Figure 7- This illustration shows a clear description and categorization of characteristics into the five domains of the Big Five Model.

Furthermore, Plasticity and Stability, which are metatraits (combinations of Big Five factors), seem to be connected to certain neurotransmitters in the brain. Stability is the tendency to hold back disruptive feeling or action, and Plasticity is the tendency to experiment and explore prospects. Stability is shown to be related to serotonin, and Plasticity

to dopamine. Serotonin can inhibit mood and behavior (Zuckerman 2014), and such traits are linked to domains of the Big Five. Aggression, lack of impulse control, and depression, which are all characteristic of the negative end of the spectrums of Neuroticism, Agreeableness, Conscientiousness, can indicate low serotonin. Plasticity is connected to dopamine, which mainly controls exploring tendencies of an individual—Extraversion and Openness. Dopamine is involved in the nucleus accumbens, amygdala, and the prefrontal cortex, strongly implying the neuroscientific basis for the Big Five (DeYoung 2010).

Extraversion characteristics such as talkativeness are key to understanding the link between the domain and the neurotransmitter dopamine. For example, speech is usually used in society to achieve social reward and serves as a sign of extraversion. Dopamine regulates the neural mechanisms that manage the brain's reward system and many studies with brain imaging showed the direct association between Extraversion and the dopaminergic cortical arousal. FMRI's show that Extraversion has a positive relationship with brain activity of the orbitofrontal cortex, nucleus accumbens, amygdala and striatum during positive rewards (Canli et al. 2001). Further, MRI scans proved that Extraversion is related to a larger orbitofrontal cortex, which implies that an individual scoring high on Extraversion may give greater value to rewards.

Neuroticism, on the other hand, symbolizes the tendency to experience the pessimistic emotions of punishment and threat, and is characterized by anxiety, depression, anger, and vulnerability (DeYoung 2010). Brain scans illustrate that Neuroticism may be linked to brain activity

in response to aversive stimuli in the amygdala and insula. Neuroticism has also been linked to smaller volumes in the prefrontal cortex and brain activity indicating poor emotional regulation (Deckersbach et al. 2006). Moreover, higher basic levels of cortisol, a stress hormone, and decreased cortisol in stress situations is linked to higher scorers in Neuroticism (DeYoung 2010).

Higher scores in Agreeableness are associated with altruistic tendencies, and such scorers are careful of others' feelings and needs and stay clear of aggression. An MRI study of the temporal sulcus, posterior cingulate cortex and fusiform gyrus—regions that control social information processing and empathy—linked greater brain volumes with Agreeableness.

Because humans are a unique species that set explicit rules and control their own behavior to achieve long term goals, there must be a brain system that controls such tendencies. Conscientiousness describes the ability to organize and control one's own impulses and actions. This was evidenced by an MRI study that found that Conscientiousness was associated with a larger volume of the middle frontal gyrus in the prefrontal cortex involved in maintaining goal information and the execution of plans (DeYoung et al. 2010).

Openness/Intellect reflects the engagement of intellectual and sensory information. It can be associated with intelligence and working memory volume; Intellect is linked to brain activity, especially during challenging working memory tasks, mainly in the prefrontal cortex, which is

involved in monitoring goal-achieving behavior and reduction of error during cognitive challenges.

A study by McGhee, Ehrer, Buckhalt, and Phillips sought to identify the relationship between Big Five traits and risk-taking in preadolescents. The Big Five personality test was administered and scored for all fifty participants, and the participants performed a risk-taking measuring task involving possibility of failure. Three options the participants had were guaranteed success with small reward, while the other two had high failure potential but yielded greater reward. When the data were analyzed, close relationships were found between high risk-taking and some Big Five traits. For instance, high Extraversion and high Openness indicated high risk-taking. Additionally, because the concept of Openness also included the tendency to seek new experiences, the researchers hypothesized that some participants viewed the riskier choice as more engrossing. On the other hand, children with high levels of Conscientiousness showed less risk-taking, suggesting that they did not desire to risk failure, even in insignificant situations. The result from this study was that while Openness and Extraversion caused the greatest variance in the predictability of the Big Five in risk-taking, Agreeableness and Conscientiousness were not as significant (McGhee et al. 2012). There are varying degrees of differences in each individual that may cause inaccuracies in rather abstract testing such as personality tests. There indeed is another much more definite and well-studied component—gender.

GENDER DIFFERENCES AND SITUATIONAL FACTORS

Gender is another factor that sets apart the risk-taking group from the relatively non-risk-taking group. Males are generally assumed to be more risk-taking than females, and there are certain mechanisms that indeed differentiate the male brain from the female brain, such as neurotransmitter processing and hormones.

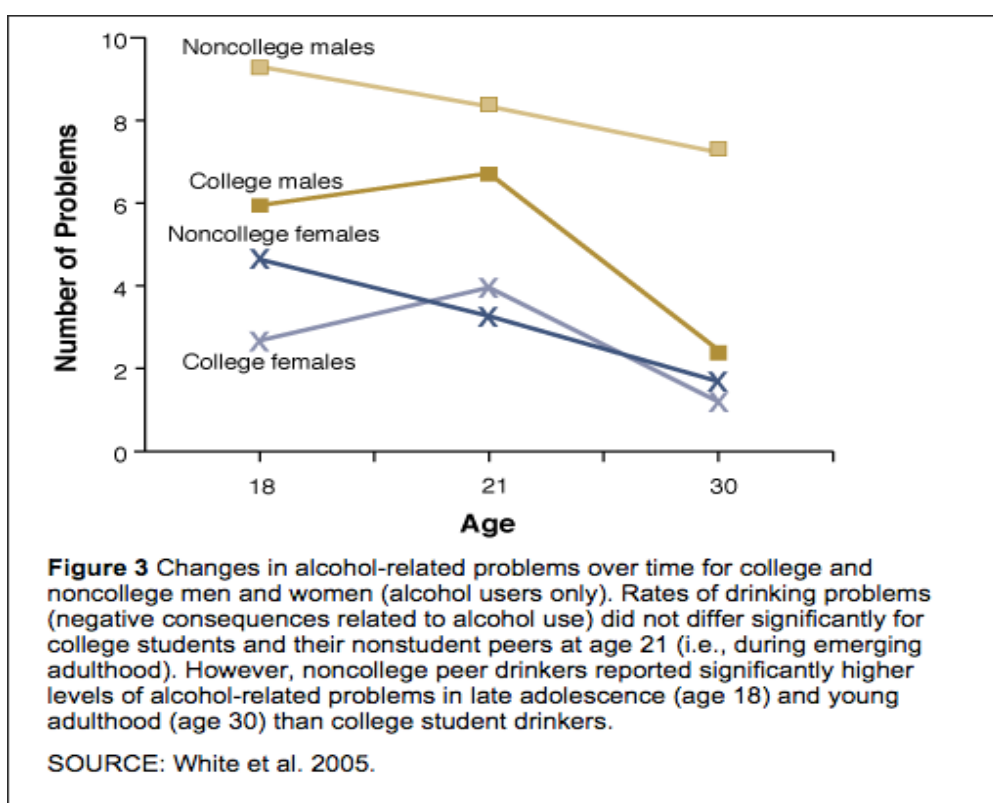


Figure 8- This is a graph illustrating the difference in number of drinking problems that males and females of late adolescence have. Males seem to have more problems than females in general.

A study by Barbara A. Morrongiello sought to answer the question of whether younger girls and boys viewed injury and injury risk differently.

Participants of ages 6 to 10 were recruited and presented with pictures from low risk (an example would be a person riding a bicycle with safety gear) to high risk (person riding a bicycle with one hand, carrying a bag, without a helmet). Then, they were asked how much they engaged in the activity. It was found that girls perceived the “evidence” of possible risk in a situation more than boys. On the other hand, boys had a tendency to assume that possible risk did not have implications for themselves, and engaged in risk-taking activity more than girls, even if they had seen a peer get injured from the activity (Morrongiello 1998).

Further, the fMRI study by Galli et al. demonstrates another neurological difference between males and females. Interoception is the ability to perceive one’s own physiological state. Activities such as breathing are carefully monitored by structures in the brain in an effort to maintain homeostasis. Galli’s study explored the neurological sex-related differences in the presence of an aversive respiratory stimulus by having participants rate the level of difficulty of breathing on a computer while they breathed normally or with a respiratory load. Galli hypothesized that men are more sensitive than woman to hedonistic rewards and sensation seeking. Sex differences did arise in the insula activity level after the termination of the respiratory load that acted as the aversive stimulus. Men especially reported the highest perceived unpleasantness, and showed the strongest activation of the reward system in the brain. This finding indicates that in men, the more aversive the stimulus, the greater the reward system activation after the termination of the stimulus.

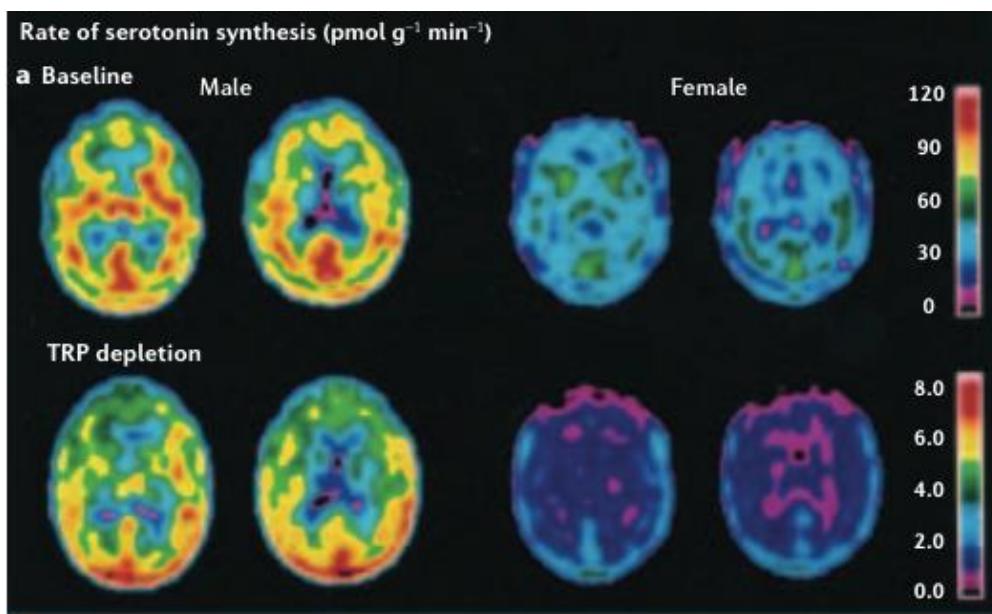


Figure 9- This PET scan of male and female brains show that males have a much higher rate of serotonin synthesis than females.

To speak more biochemically, neurotransmitters in the brain may also influence male and female decision-making. Figure 9 depicts the positron emission tomography of male and female rates of serotonin synthesis in the brain. The scans were taken before and after the depletion of plasma tryptophan, which creates serotonin. Serotonin, in fact, is a calming neurotransmitter; a lack thereof may result in overstimulation. This helps explain why adolescent males may engage in more risk-taking behavior—their serotonin depletion rate is higher than females' (Cahill 2006).

Neurotransmitters impact male and female decision-making and risk-taking greatly, but there are many more factors such as hormones that accentuate their differences. For example, stressful situations seem to have opposite effects on male and female risk-taking. In a stressed state, the male body releases testosterone, which increases aggression. For females, there is

no change in aggression; in fact, the female body releases oxytocin, which has calming effects. fMRI activity scans showed that women had greater activation in the amygdala, ventromedial prefrontal cortex and anterior cingulate, but men had much greater activation in the dorsolateral prefrontal cortex and insula. Due to this difference, stress is more likely to activate the emotional networks in decision making for women, and in the medial prefrontal regions in males, stress increases strategic processing rather than risk avoidance or somatic cues. As a result, under acute stress, the contrast between women and men is heightened (Lighthall et al. 2009), as illustrated by Figure 10.

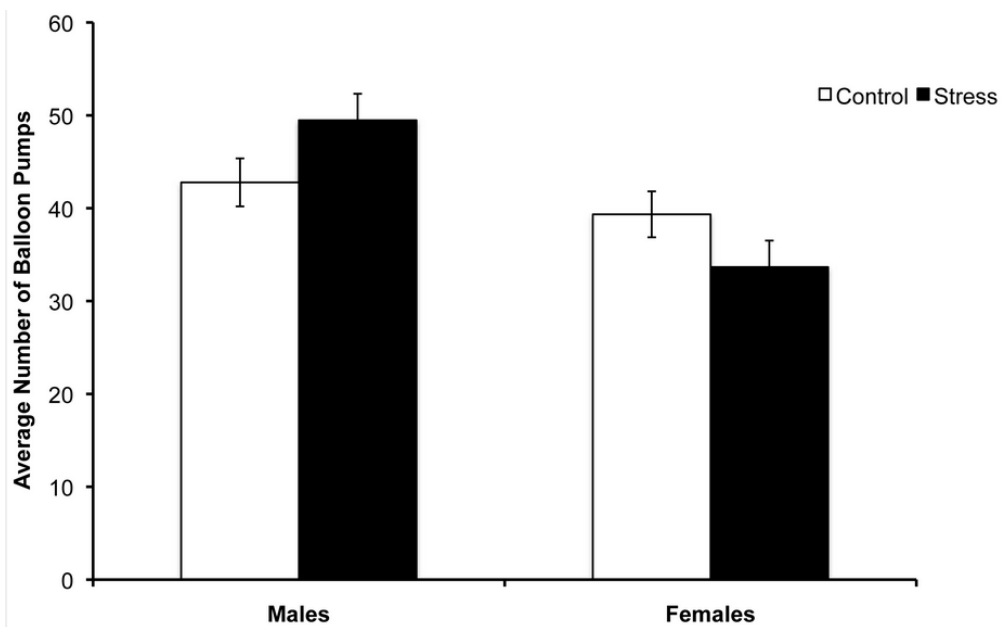


Figure 10- In this study, the level of risk taking was measured by the number of pumps into the balloon in stressed and controlled situations. Males have an increase in risk-taking in a stressed state, and women have a decrease in risk-taking.

IMPLICATIONS AND ANALYSIS

My hypothesis that personality and gender predispositions strongly connect with the neural networks and mechanisms involved in decision-making during adolescence is, for the most part, correct. Some regions of the brain, particularly the frontal cortex, limbic system and reward system, are influenced by genetic individual differences in terms of temperament, behavior and natural gender differences. If my hypothesis is true, there are indeed certain qualities in personality that lead to more risky behavior and impulsive thoughts, as well as certain biological differences in sex.

The information presented in my paper indicates that the notion that adolescents are universally prone to risk-taking is actually much more complex, with multifaceted details underneath a simple generalization. In some ways, it is true that the maturation of neural networks is often the source of adolescents' risk-taking and impulsive decision-making. In that aspect, there is plenty of evidence to support that biological developments cause an increase in risk-taking. However, some evidence provided in my paper may be too vague for a concrete statement that personality is an entirely determining factor in adolescent decision-making. For example, the fact that three out of the Big Five domains failed to predict preadolescent risk-taking may be a sign that more research is needed. On top of that, people tend to act differently around peers or in stressful situations. Additionally, how different hormones cause either reduced or increased risk-taking behavior in men and women should be explored, possibly with adolescent participants. Lastly, learning more about how each brain region, especially the striatum and prefrontal cortex, exactly influences adolescent risk-taking will be vital in understanding the mechanisms of adolescent behavior and tendency to make rash decisions. Adolescents are a greatly

understudied group as compared to children and adults, and there are a myriad of paths that adolescent research could take in the future.

CONCLUSION

This paper presents the neuroscientific basis for risk-taking in the developing brain. Despite the fact that development patterns apply to every adolescent universally, there are many pieces to consider when analyzing the level of risk-taking. Whether they be neurotransmitters, hormones, gender, personality dispositions or situational factors, every adolescent's decision-making differs slightly due to individual variation. With enough experimental data, it is probable that adolescent behavior might be clarified in the near future, since they are currently an understudied population.

Although there could exist innumerable unaddressed aspects due to the broad nature of structures and systems involved in decision-making, it is evident from a variety of studies that certain types of personalities, as well as gender, influence adolescent risk-taking, from the anatomical level to the molecular. As observed by McGhee et al. in the 2012 research paper on the Big Five, "Risk taking is a multidimensional phenomenon involving personality correlates, self-regulating processes, and most likely cognitive aspects of decision making." All of these processes are controlled by extensive, complex interactions between brain mechanisms involving the prefrontal cortex, limbic system and the dopamine signaling system. The sudden increase in risk-taking is a temporary phenomenon that is driven by the imbalance in maturation from child to adult.

BIBLIOGRAPHY

- Adriani, W., Chiarotti, F., & Laviola, G. (1998). Elevated novelty seeking and peculiar d-amphetamine sensitization in periadolescent mice compared with adult mice. *Behavioral Neuroscience*, *112*(5), 1152–66. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9829793>
- Bonnie, R. J., O’Connell, M. E., & Drinking, N. R. C. (US) and I. of M. (US) C. on D. a S. to R. and P. U. (2004). *Consequences of Underage Drinking*. National Academies Press (US). Retrieved from <http://www.ncbi.nlm.nih.gov/books/NBK37591/>
- Brand, M., Labudda, K., Kalbe, E., Hilker, R., Emmans, D., Fuchs, G., ... Markowitsch, H. J. (2004). Decision-making impairments in patients with Parkinson’s disease. *Behavioural Neurology*, *15*(3-4), 77–85. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/15706051>
- Cahill, L. (2006). Why sex matters for neuroscience. *Nature Reviews. Neuroscience*, *7*(6), 477–484. <http://doi.org/10.1038/nrn1909>
- Canli, T. (2004). Functional brain mapping of extraversion and neuroticism: learning from individual differences in emotion processing. *Journal of Personality*, *72*(6), 1105–32. <http://doi.org/10.1111/j.1467-6494.2004.00292.x>
- Casey, B. J., Jones, R. M., & Hare, T. A. (2008). The adolescent brain. *Annals of the New York Academy of Sciences*, *1124*, 111–26. <http://doi.org/10.1196/annals.1440.010>

- Clark, C. A., & Dagher, A. (2014). The role of dopamine in risk taking: a specific look at Parkinson's disease and gambling. *Frontiers in Behavioral Neuroscience*, 8, 196.
<http://doi.org/10.3389/fnbeh.2014.00196>
- Deckersbach, T., Miller, K. K., Klibanski, A., Fischman, A., Dougherty, D. D., Blais, M. A., ... Rauch, S. L. (2006). Regional cerebral brain metabolism correlates of neuroticism and extraversion. *Depression and Anxiety*, 23(3), 133–8. <http://doi.org/10.1002/da.20152>
- DeYoung, C. G. (2010). Personality Neuroscience and the Biology of Traits. *Social and Personality Psychology Compass*, 4(12), 1165–1180.
<http://doi.org/10.1111/j.1751-9004.2010.00327.x>
- Galli, G., Shukla, A., Simmons, A. N., Davenport, P. W., & Paulus, M. P. (2013). Sex differences in the neural processing of aversive interoceptive events: the benefit of relief. *PloS One*, 8(12), e84044.
<http://doi.org/10.1371/journal.pone.0084044>
- Le Bihan, D. (2003). Looking into the functional architecture of the brain with diffusion MRI. *Nature Reviews. Neuroscience*, 4(6), 469–80.
<http://doi.org/10.1038/nrn1119>
- Lighthall, N. R., Mather, M., & Gorlick, M. A. (2009). Acute stress increases sex differences in risk seeking in the balloon analogue risk task. *PloS One*, 4(7), e6002. <http://doi.org/10.1371/journal.pone.0006002>

- McGhee, R. L. (2012). The Relation between Five-Factor Personality Traits and Risk-Taking Behavior in Preadolescents. *Psychology, 03*(08), 558–561. <http://doi.org/10.4236/psych.2012.38083>
- Morrongiello, B. A., & Rennie, H. (1998). Why do boys engage in more risk taking than girls? The role of attributions, beliefs, and risk appraisals. *Journal of Pediatric Psychology, 23*(1), 33–43. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9564127>
- Primus, R. J., & Kellogg, C. K. (1989). Pubertal-related changes influence the development of environment-related social interaction in the male rat. *Developmental Psychobiology, 22*(6), 633–43. <http://doi.org/10.1002/dev.420220608>
- Purves, D., Augustine, G. J., Fitzpatrick, D., Katz, L. C., LaMantia, A.-S., McNamara, J. O., & Williams, S. M. (2001). *Neuroglial Cells*. Sinauer Associates. Retrieved from <http://www.ncbi.nlm.nih.gov/books/NBK10869/>
- Rosenberg, D. R., & Lewis, D. A. (1995). Postnatal maturation of the dopaminergic innervation of monkey prefrontal and motor cortices: a tyrosine hydroxylase immunohistochemical analysis. *The Journal of Comparative Neurology, 358*(3), 383–400. <http://doi.org/10.1002/cne.903580306>
- Schramm-Sapyta, N. L., Walker, Q. D., Caster, J. M., Levin, E. D., & Kuhn, C. M. (2009). Are adolescents more vulnerable to drug addiction than adults? Evidence from animal models. *Psychopharmacology, 206*(1), 1–21. <http://doi.org/10.1007/s00213-009-1585-5>

Simon, N. W., & Moghaddam, B. (2015). Neural processing of reward in adolescent rodents. *Developmental Cognitive Neuroscience, 11*, 145–54.
<http://doi.org/10.1016/j.dcn.2014.11.001>

Sonntag, K. C., Brenhouse, H. C., Freund, N., Thompson, B. S., Puhl, M., & Andersen, S. L. (2014). Viral over-expression of D1 dopamine receptors in the prefrontal cortex increase high-risk behaviors in adults: comparison with adolescents. *Psychopharmacology, 231*(8), 1615–26.
<http://doi.org/10.1007/s00213-013-3399-8>

Spear, L. P. (2000). The adolescent brain and age-related behavioral manifestations. *Neuroscience and Biobehavioral Reviews, 24*(4), 417–63.
Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/10817843>

Squeglia, L. M., Jacobus, J., & Tapert, S. F. (2009). The influence of substance use on adolescent brain development. *Clinical EEG and Neuroscience, 40*(1), 31–8. Retrieved from
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2827693&tool=pmcentrez&rendertype=abstract>

Steinberg, L. (2008). A Social Neuroscience Perspective on Adolescent Risk-Taking. *Developmental Review : DR, 28*(1), 78–106.
<http://doi.org/10.1016/j.dr.2007.08.002>

Sturman, D. a., & Moghaddam, B. (2012). Striatum processes reward differently in adolescents versus adults. *Proceedings of the National Academy of Sciences, 109*(5), 1719–1724.
<http://doi.org/10.1073/pnas.1114137109>

- Toledo-Rodriguez, M., & Sandi, C. (2011). Stress during Adolescence Increases Novelty Seeking and Risk-Taking Behavior in Male and Female Rats. *Frontiers in Behavioral Neuroscience*, 5, 17.
<http://doi.org/10.3389/fnbeh.2011.00017>
- Walker, E. F. (2002). Adolescent neurodevelopment and psychopathology. *Current Directions in Psychological Science*, 11(1), 24–28.
<http://doi.org/10.1111/1467-8721.00161>
- Weiner, I. B., & Greene, R. L. (2008). Handbook of personality assessment. *John Wiley and Sons*, xviii, 677. Retrieved from
http://books.google.com.library.gcu.edu:2048/books?hl=en&lr=&id=HcLSuFEYRo8C&oi=fnd&pg=PR7&dq=the+means+of+determining+the+reliability+of+personality+tests&ots=RbV0lzmJb7&sig=htIde7hHkVccD_dPqJS6iSwZu00
- Zuckerman, M., & Kuhlman, D. M. (2000). Personality and risk-taking: common biosocial factors. *Journal of Personality*, 68(6), 999–1029.
Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11130742>

FIGURES

- Casey, B., Jones, R., & Hare, T. (n.d.). The Adolescent Brain. Retrieved September 27, 2015.
- Gogtay, N. (n.d.). Adolescent Brain Booklet. Retrieved September 27, 2015, from
http://www.mentorfoundation.org/uploads/Adolescent_Brain_Booklet.pdf
- NIAAA Publications. (n.d.). Retrieved September 27, 2015.

Somerville, L., Jones, R., & Casey, B. (n.d.). A time of change: Behavioral and neural correlates of adolescent sensitivity to appetitive and aversive environmental cues. Retrieved September 27, 2015.

Squeglia, L., Jacobus, J., & Tapert, S. (n.d.). The Influence of Substance Use on Adolescent Brain Development. Retrieved September 27, 2015.

"Striatum Processes Reward Differently in Adolescents versus Adults." *PNAS*. N.p., n.d. Web. 7 Jan. 2016.

The Big- 5 Theory Of Personality: The O-C-E-A-N Of Human Behavior - PsyTreasure. (2015, June 25). Retrieved September 27, 2015.



The Effect of Diffraction Aperture Structure on Reciprocal Space

Eren Kafadar

Author background: Eren grew up in Turkey and currently attends Robert College, a high school in Istanbul. Her Pioneer seminar program was titled "Optical Diffraction Patterns."

Abstract

Reciprocal space can be investigated by looking at diffraction patterns of different aperture structures. The Digital Fourier Transform is used to compute these diffraction patterns through the software MATLAB. The simulations confirm that the dimensions of the aperture are inversely proportional to the dimensions of the diffraction pattern. This is true for double slit structures as well as single slit ones. The positions of the apertures, however, is only a factor when there is more than one opening. Reciprocal space can also be seen through observation of matrices created by simulations on MATLAB. These matrices reveal that the inverse proportionality stems from the process through which the diffraction pattern is found.

1. Introduction:

Diffraction of light waves occurs when light passes through a small opening on a large surface (a slit). Light waves bend as they pass through this slit, and the result can be observed on a screen behind the object.ⁱ When the distance between the slit and the screen (D) is relatively larger than the width of the slit (d) it is called the Fraunhofer diffraction. The following equation gives the relationship between the diffraction and the slit widthⁱⁱ.

$$m \times \lambda = \frac{y \times d}{D} \quad (1)$$

where m is the wave number, λ is the wavelength, d is the slit width, y is the distance from the central fringe, and D is the distance between the slit and the screen.

To simulate the diffraction patterns of different shaped apertures, the Digital Fourier Transform (DFT) is used. The DFT analyzes the structure, which can be represented as a function, to find which frequencies are present, and how much. The intensity of the diffraction pattern is directly proportional to the square of the DFT's absolute value. However, the structure of the slit cannot be found using the diffraction pattern simulated by DFT, because the phase information for the light waves is lost when we take the absolute value of the DFT.

The DFT of any function is another function that uses the reciprocal variable of the original function: eg. if the original function is given in terms of time(t), the DFT would be given in terms of frequency(t^{-1})ⁱⁱⁱ.

The Fourier transform translates real space into the reciprocal space, using two conjugate variables commonly denoted as p and x . In the case of diffractions, these variables correspond to (respectively) the dimensions of the aperture (d) (where light passes through) and the diffraction angle divided by the wavelength ($\sin \theta / \lambda$).ⁱⁱⁱ As the slit gets narrower, the diffraction angle gets bigger and thus the resulting pattern has wider maxima.

The diffraction pattern is the "reciprocal lattice" of the object we use to obtain this diffraction. This idea is often used in crystallography to figure out the lattice.^{iv} The closer the atoms are in the lattice, the wider spread the diffraction pattern.

I used MATLAB, a matrix-based computing program, to simulate these different diffraction patterns to observe the effects of different slit structures on reciprocal space. Because we cannot directly translate the diffraction pattern back into the slit structure due to the lost phase information, reciprocal space can help predict the slit structure for a given diffraction pattern.

2. Simulations of single slit structures

The first factor that is analyzed is the position of the slit.

In a one-dimensional simulation the width of the slit is kept constant while the position is changed. DFT is used to simulate the resulting diffraction pattern. The diffraction patterns for two identical slits at different positions are identical as seen in **Figure 1(a)**.

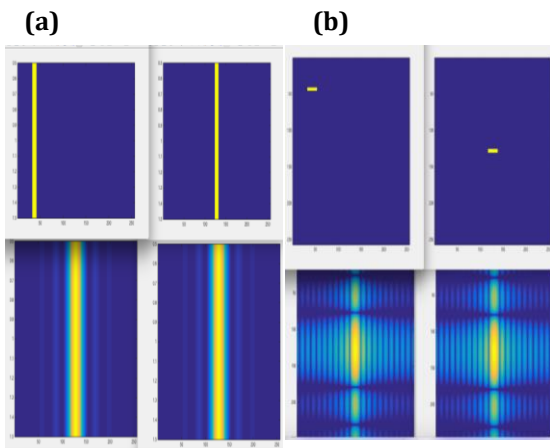


Figure 1 In the figure above we can see that in both the 1D(left) and 2D(right) simulations, even though the positions of the slits are different, the resulting diffraction patterns are identical.

The same process is repeated with two-dimensional slits, this time with both dimensions kept constant, and the diffraction patterns are again identical **Figure 1(b)**. This observation is

consistent with the mathematics behind Fourier transforms, $F(t)$, because Fourier Series are summations of different frequencies present in the original object, or function $f(t)$, which is independent of the position of the objectⁱⁱⁱ.

$$F(x) = \sum_{-\infty}^{+\infty} a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \quad (2)$$

The second factor for single slit simulations is the dimensions of the slit. The dimensions of the slit have inverse proportionality to the dimensions of the diffraction pattern. As seen in **Figure 2** the slit with bigger dimensions has smaller fringes compared to the slit with smaller dimensions. This is true for both one-dimensional and two-dimensional analysis.

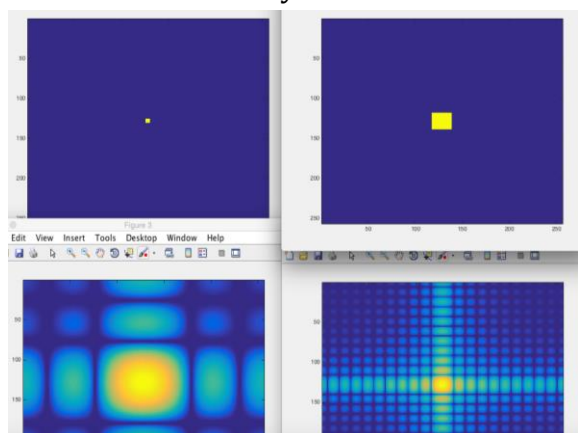


Figure 2 The figure shows that when the aperture is bigger (right), the diffraction pattern has smaller fringes.

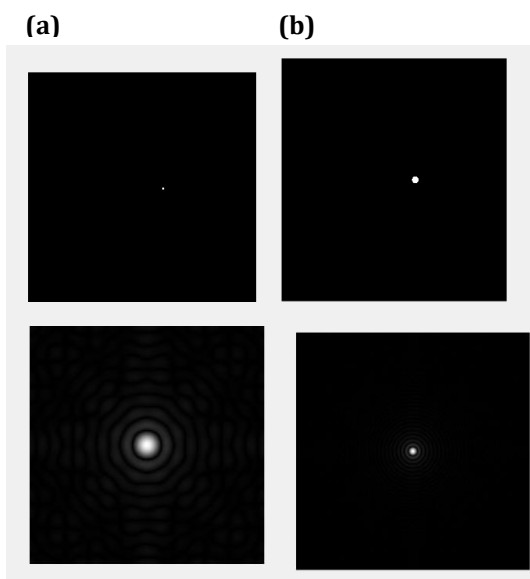


Figure 3 *The above figure shows the inverse relationship between the diameter of the slits (upper row) and the diameter of the diffraction pattern (lower row).*

When the horizontal dimensions are kept constant and the vertical dimensions are changed, it is seen that these changes are translated inversely onto the diffraction pattern. This proves the idea of reciprocal space. Referring back to Equation (1), the $(m \times \lambda)$ value and the distance from the screen (D) is constant for both simulations. When the width of the slit (d) is increased, the width of the fringes (y) has to decrease in order to satisfy the equation. The shape of the aperture in this case is rectangular; however, the

idea of reciprocal space holds when the apertures have different shapes. When the shape of the aperture changes, the shape of the fringes of the diffraction pattern changes in the same way.

When we use circular apertures instead of rectangular ones, we see a central bright disc surrounded by concentric circles gradually decreasing in intensity, which can be seen in **Figure 3 (a)**. When the diameter of the original circle is increased, the diameter's circular fringes in the diffraction pattern get smaller, which is shown in **Figure 3(b)**.

3. Simulations of double slit structures

Double slit diffractions follow the same Equation (1) but this time the variable d corresponds to the distance between the two slits rather than the width of each slit. This difference is also responsible for the differences in how the reciprocal space works for double slit diffractions. Once the light beam passes through the slits, the two slits start acting as two different light sources, and as a result the light beams from these two slits interfere to create an interference pattern. When the two light waves are in phase we see constructive interference resulting in a bright fringe, and when the light waves are out of phase we

see destructive interference

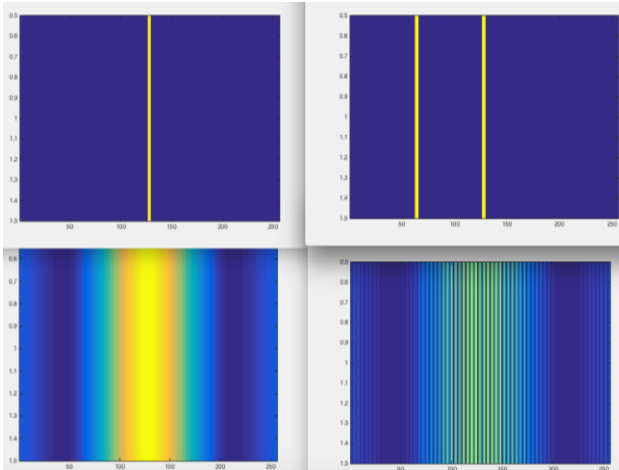


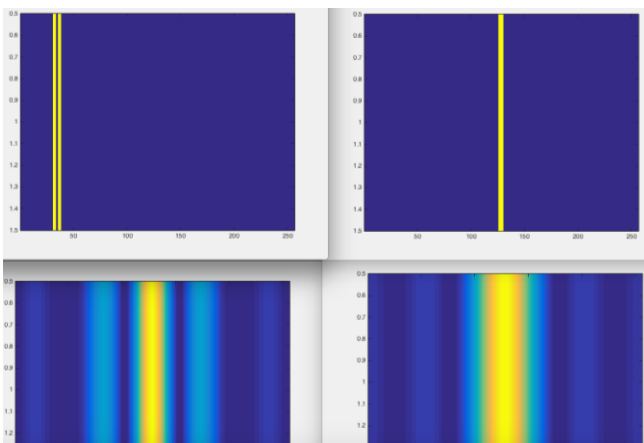
Figure 4 In the figure above both diffraction patterns exhibit the same general shape. However, the double-slit diffraction has narrow, vertical strips of light and dark fringes within this general pattern.

resulting in a dark fringe.

On the screen, these bright and dark fringes create equally spaced bright and dark spots, which are enclosed in a general pattern of diffraction for a single rectangular slit^v. This phenomenon can be seen in **Figure 4**.

Unlike single-slit simulations, the positions of the two slits has a huge impact on the diffraction pattern. In **Figure 5 (a)** we see a double slit diffraction, where each slit has width n , and on the right there is a single slit diffraction where the slit has width $2n$. The slits on the left are very close together, so they almost act as if they were a single source of light (a single slit) instead of two different sources of light. As a result, the interference pattern is almost invisible. This idea is proven in **Figure 5(b)** where the diffraction on the left is again showing two slits; however, now they are so close that MATLAB imaging cannot differentiate between them. As a result, the diffraction pattern for the double slit interference on the left is identical to the diffraction pattern for the single slit interference on the right, because they are both treated as

(a)



(b)

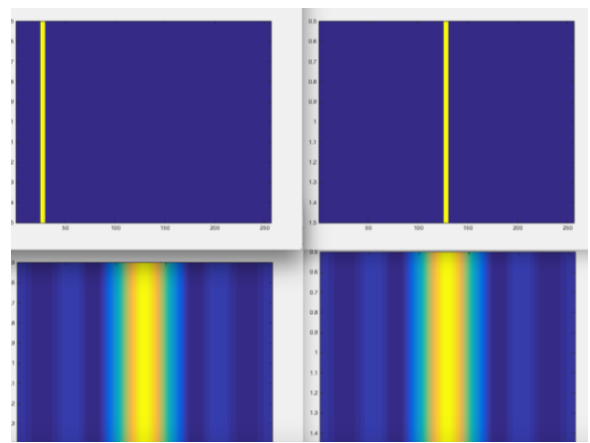


Figure 5 As the slits in double slit diffractions get closer, the interference becomes less visible and the pattern becomes similar to a single slit diffraction.

if they were single slit simulations.

4. Explanation of reciprocal space

The DFT uses the following formula for the phase matrix of one dimensional simulations^{vi}:

$$X_k = \sum_{n=0}^{N-1} e^{-\frac{2\pi i k n}{N}} x_n \quad (3)$$

where X_k is the spectrum of x at spatial frequency $2\pi k$, N is the number of samples, n is the sampling instant, and k is the order of sample. In a MATLAB 8 x 8 matrix, this formula is evaluated for every point and gives the following phase matrix:

Table 1 The phase matrix on MATLAB

	1	2	3	4
1	0.7071 + 0.7071i	-0.0000 - 1.0000i	-0.7071 + 0.7071i	1.0000 + 0.0000i
2	-0.0000 - 1.0000i	-1.0000 + 0.0000i	0.0000 + 1.0000i	1.0000 + 0.0000i
3	-0.7071 + 0.7071i	0.0000 + 1.0000i	0.7071 + 0.7071i	1.0000 + 0.0000i
4	1.0000 + 0.0000i	1.0000 + 0.0000i	1.0000 + 0.0000i	1.0000 + 0.0000i
5	-0.7071 - 0.7071i	0.0000 - 1.0000i	0.7071 - 0.7071i	1.0000 + 0.0000i
6	-0.0000 + 1.0000i	-1.0000 - 0.0000i	0.0000 - 1.0000i	1.0000 + 0.0000i
7	0.7071 - 0.7071i	-0.0000 + 1.0000i	-0.7071 - 0.7071i	1.0000 + 0.0000i
8	-1.0000 - 0.0000i	1.0000 + 0.0000i	-1.0000 - 0.0000i	1.0000 + 0.0000i

	5	6	7	8
1	-0.7071 - 0.7071i	-0.0000 + 1.0000i	0.7071 - 0.7071i	-1.0000 - 0.0000i
2	0.0000 - 1.0000i	-1.0000 - 0.0000i	-0.0000 + 1.0000i	1.0000 + 0.0000i
3	0.7071 - 0.7071i	0.0000 - 1.0000i	-0.7071 - 0.7071i	-1.0000 - 0.0000i
4	1.0000 + 0.0000i	1.0000 + 0.0000i	1.0000 + 0.0000i	1.0000 + 0.0000i
5	0.7071 + 0.7071i	0.0000 + 1.0000i	-0.7071 + 0.7071i	-1.0000 + 0.0000i
6	0.0000 + 1.0000i	-1.0000 + 0.0000i	-0.0000 - 1.0000i	1.0000 - 0.0000i
7	-0.7071 + 0.7071i	-0.0000 - 1.0000i	0.7071 + 0.7071i	-1.0000 + 0.0000i
8	-1.0000 + 0.0000i	1.0000 - 0.0000i	-1.0000 + 0.0000i	1.0000 - 0.0000i

There are two objects, A and B, with different widths being analyzed and they have the following 1 x 8 object matrices:

Table 2 Object matrix for object A

	1	2	3	4	5	6	7	8
1	0	1	1	1	1	1	0	0

Table 3 Object matrix for object B

	1	2	3	4	5	6	7	8
1	0	0	1	1	1	0	0	0

These object matrices are multiplied by the phase matrix and the result is the DFT, a complex matrix which is then squared to give the intensity of the diffraction pattern. The intensity matrices are seen below:

Table 4 DFT² of object A

	1	2	3	4	5	6	7	8
1	0.1716	1.0000	5.8284	9	5.8284	1.0000	0.1716	1

Table 5 DFT² of object B

	1	2	3	4	5	6	7	8
1	0.1716	1.0000	5.8284	25	5.8284	1.0000	0.1716	1

As we can see, the DFT² table of object B, which is twice as wide as object A, has a faster drop in values as we move away from the first maximum (column 4). This translates to the diffraction pattern as a narrower bright fringe because the fringe right next to the central one is very dark in comparison. Since the square of the DFT is only proportional to the intensity of the diffraction pattern, and doesn't give the actual intensity values, what is important for the simulation is the relative values within each object rather than the relative values between objects. This means that even though the central fringe of object B has a higher value than the central fringe of object A (25>9), the bright

fringe of object A is wider than the bright fringe of object B.

For an 8x8 matrix, the values except for the central one are identical for both objects; however, this is not always the case. When the matrices and objects are greater (16x16, 32x32), there are different arrangements of values for objects of different widths.

When the object is wider, there are more "1"s and fewer "0"s on the object matrix. This means that as we multiply the object matrix by the phase matrix, there are fewer values that are canceled out by "0"s. Since the phase matrix has negative values as well as positive ones, some values of the DFT matrix end up being smaller than what they would've been had the "0"s cancelled out negative values from the phase matrix (this would happen if the object was narrower and therefore had more "0"s in the object matrix). Having smaller values in the DFT matrix means that these fringes are darker than what they would've been if the slit had been narrower. This results in narrower central bright fringes for wider slits, and hence in reciprocal space.

The process explained above also works for dark fringes. If the value on the object matrix is "1" instead of "0", certain values on the DFT end up being greater because fewer positive values on the phase matrix are now cancelled by "0"s

on the object matrix. Therefore, the darker fringes are also narrower because fringes next to them are now brighter (have greater values) than they would've been had the object been narrower. As a result, the whole diffraction pattern for wide slits is narrower than the diffraction pattern for narrow slits.

5. Conclusion

The initial sections of the paper discuss how the reciprocal space is affected by structure of the aperture. We observed that while the position of the aperture isn't a factor in single slit diffractions, it plays a significant role in double slit diffractions. The effects of the dimensions of the aperture were also analyzed, and we saw that an increase in a certain dimension of the aperture is translated as a decrease in that dimension of the diffraction pattern. This proved the idea of reciprocal space in Fraunhofer diffractions.

The final part of the paper discusses the explanation behind the reciprocal space by investigating matrices used in computing the simulations for the paper. Wider apertures result in DFT^2 matrices where the change in values is steeper, resulting in a faster change from light to dark fringes which means the diffraction pattern is narrower.

References

- ⁱ Magnes, Jenny, Kathleen Susman, and Rebecca Eells. "Quantitative Locomotion Study of Freely Swimming Micro-Organisms Using Laser Diffraction." *Journal of Visualized Experiments : JoVE* 68 (2012): 4412. PMC. Web. 19 Sept. 2015.
- ⁱⁱ Walker, James S. "Physical Optics: Interference and Diffraction." *Physics*. Upper Saddle River, NJ: Pearson/Prentice Hall, 2004. 979-93. Print.
- ⁱⁱⁱ James, J. F. "1. Physics and Fourier Transforms." *A Student' Guide to Fourier Transforms: With Applications in Physics and Engineering*. Cambridge: Cambridge U, 2002. 9-11. Print.
- ^{iv} Williams, David B., and C. Barry Carter. "12 Thinking in Reciprocal Space." *Transmission Electron Microscopy: A Textbook for Materials Science*. New York: Springer, 2009. 211. Print.
- ^v Duffy, Andrew. "25 Interference and Diffraction." *Essential Physics*. 1st ed. N.p.: Boston U, 2012. N. pag. Advanced Instructional Systems. Web. 19 Sept. 2015.
- ^{vi} Smith, J.O. "DFT Definition." *Mathematics of the Discrete Fourier Transform (DFT) with Audio Applications, Second Edition*, online book, 2007 ed. N.p. 19 Sept. 2015.

The Hidden Key to Success: How Saudi Arabia's Economic Needs are Expanding Women's Rights

Mae Kirkpatrick

Author background: Mae was born in Holland, grew up in Singapore and currently attends Hockerill Anglo-European College, a high school in Hertfordshire, United Kingdom. Her Pioneer seminar program was titled "Women's Rights in the Middle East and South Asia."

1. Introduction

Saudi Arabia has long been a country known for its strong religious belief system, Wahabi Islam, on which its society, laws and economic system are based. With a *Shura* Council having heavy input in all political and legal decisions, and religious police, the *Mutawe'en*, roaming the streets enforcing compliance with the country's strict interpretation of shari'a law, the daily lives of the population of Saudi Arabia are severely impacted. For women, this impact has mostly been negative, with Saudi Arabia currently standing as one of the worst countries in regard to women's rights, ranking 130th of 142 countries in the Global Gender Gap Index for 2014 (Global Gender Gap Report 2014). Some ways in which women are unequal to men in Saudi Arabia include the guardianship rule, which forbids women from medical treatments, work and travel without male permission, and the infamous driving ban, which restricts women from driving and forces them to find other means such as drivers and male guardians for transport. Saudi Arabia is also one of the only countries where women are not able to vote in national elections. Saudi women must always dress in public in at least the traditional *abaya*, a long black cloak with a black headscarf. Lastly, Saudi

women are segregated from men in almost all everyday activities. This repression has plagued women for many years, and even as many neighboring Gulf countries, such as Bahrain, Qatar, the UAE and Kuwait, continue to make progress towards gender equality, Saudi Arabia seems to be falling farther and farther behind (Women's Rights Gain Traction 2009).

Women in Saudi Arabia have continued to advocate for their rights and have been gaining supporters from every area in their battle for equality. Many of their allies have been those in the business community, who see economic benefits in the expansion of women's rights. In this paper I will explore how the economic needs of Saudi Arabia correlate with the expansion of women's rights and seek to answer the following question: In what ways have the country's economic concerns driven the expansion of women's rights in Saudi Arabia? I will argue that Saudi Arabia's need to nationalize the work force, diversify the business and industrial workforce, and improve the Saudi image abroad have been the driving forces behind the expansion of women's rights, specifically through being allowed into more areas of work and power, expanded educational opportunities, and increased publicity in relation to women's rights.

In order to argue further that economics is a driving force in the expansion of women's rights, I will use a variety of sources. I will first look at past research that has examined the economic concerns of Saudi Arabia and how women's rights have been connected to them. I will also analyze the National Development Plans of Saudi Arabia, specifically looking at its economic goals and goals related to women's rights, and speeches of previous and current Saudi Arabian rulers. This will help me further find what correlation people in positions of power see between Saudi Arabia's economic needs and the expansion of women's rights. Additional data from

the Economic Intelligence Unit, the United Nations and Human Rights Watch will also give some insight into the topic that will be useful in answering my question. The variety of evidence presented will all contribute to support the argument that follows.

The paper will begin by sketching the socio-economic history of Saudi Arabia and how both its economy and the rights of women have progressed or regressed over the years. This section of history will also delve deeper into which people have had a large role in both these areas and which people have been significant in highlighting correlation between them, showing powerful, professional opinions for the connection between women and economic needs. The history will be significant in creating a clear base for determining what has led to advances in women's rights in the past and what have been the past impacts of the economic needs of Saudi Arabia. It will also be significant in creating a clearer understanding of the current state of Saudi Arabia's economy and what current economic concerns Saudi Arabia is faced with. The paper will then move on to consider the economic benefits of integrating women into the workforce and other outcomes that are primarily beneficial to Saudi Arabia's development goals and plans. It will also look into how the expansion of women's rights has allowed the entrance of women into the public sphere, more specifically, into higher education and the labor market. The paper will conclude with final remarks and thoughts about Saudi Arabia's progressive development as a whole.

This research is significant because in order to speed up the progress in women's rights in Saudi Arabia, it is vital to build upon the correlation between economic needs and gender equality and use them as a key, inarguable justification in the fight for women's rights.

2.A Brief Socio-Economic History of Saudi Arabia

Saudi Arabia was established as a unified country in 1932. Before becoming Saudi Arabia, it was populated with tribes loyal to empires such as the Ottoman Empire or eventually the British Empire. Abdulaziz al-Saud wanted to regain what had previously been a small empire or state ruled by his family, and using the help of the Wahhabis, an ultra-conservative Islamic group, created a band of militias called the *Ikhwan*. The *Ikhwan* fought for al-Saud under the agreement that they could enforce and impose their ultra-conservative interpretation of Islam on those they conquered. Wahhabism has evolved since then from opposing liberal Islam to opposing any forms of modernization, with women's rights in particular being seen as a symbol of Westernization and modernization (Fisher 2015).

Since the first discovery of oil in Saudi Arabia in 1938, Saudi Arabia's economy has made rapid progress, with an average annual GDP growth rate of 5% from 1969 to 2015 (Saudi Arabia GDP Growth Rate 2015). Oil was first produced under the US-backed oil company Aramco, otherwise known as the Arabian American Oil Company. Between 1964 and 1975, King Faisal began modernizing Saudi Arabia. With oil revenues increasing by 1600%, he was able to bring about rapid developments in the economy, accompanied by rapid modernization and industrialization, leading to an increase in relations with Western countries. This displeased the *ulama*¹, as it went against much of what they had been opposed to. In order to appease the ultra-conservative clergy on whose support the monarchy's legitimacy rested, King Faisal tolerated extremist Muslims from other Arab countries, such as Egypt and Syria, and thus made Saudi Arabia a sanctuary for fundamentalist scholars and professionals. Being seen as a

¹ The *ulama* are a group of Muslim scholars and religious leaders in Saudi Arabia

devout Muslim, King Faisal was able to begin expanding female education, believing that tradition and development could work together. He justified his actions with the idea that no aspect of the *Quran* forbids the education of women while many support it (Hamdan 49, 2005). He also allowed for the first television broadcast, which was a recitation of the *Quran*. Both sparked many protests from the religious conservatives but King Faisal continued to protest and advocate for the *Quran's* acceptance of both changes (A Chronology: The House of Saud 2005).

In 1972, the reliance on the US was lessened as Saudi Arabia was able to gain 20% control of Aramco. They used their oil to advantage in boycotting Western countries that were supporting Israel in the 1973 October Yom Kippur war. In 1979, one of the most dramatic events was to occur in Saudi Arabia. On November 20, hundreds of Saudi fundamentalists took over al-Haram, the Great Mosque of Mecca. Their actions were in protest against the modernization that had taken place in Saudi Arabia. The radicals accused the royal family of being illegitimate and corrupt, due to their modernization and connections with the West. After the event, the royal family went on to further enforce religious law and behavior, and created a new Islamic agenda for Saudi Arabia. In the same year, the Soviet Union invaded Afghanistan, and Saudi Arabia, seeing communism as a godless ideology, allied with the US to finance the Afghan war against the Soviet Union. Saudi Arabia also supplied thousands of Saudis to fight with the *Mujahedeen* in Afghanistan against the communists. Between 1975 and 1982 King Khalid ruled, and he emphasized development of the country. This allowed Saudi Arabia to reach a period of wealth and prosperity, which caused a slight rise in tension between those who wanted development and

modernization and the religious fundamentalists (Saudi Arabia Profile 2015).

In 1991, 47 Saudi women attempted to drive in order to challenge the driving ban, before being arrested by religious police. The justification for banning women drivers was that conservative Muslim clerics felt that promiscuous behavior would follow if women could drive. In addition, religious scholars claimed that through driving, women would lose their virginity and be led deeper into a life of sin (Associated Press 2014). Despite an order by Prince Salaman to release the women, the royal family later responded to the anger of religious conservatives by reasserting the ban on women driving. In 1992, King Fahd announced a reformed system of government, which included the proposal of a Consultative Council, later to be known as the *Shura* Council. In the late 1990's came another significant change in the public sphere for the Saudis, with the rise in Arab media. They began broadcasting across Saudi Arabia, and for the first time Saudis witnessed televised reports on their lack of civil rights, political freedoms, and the corruption among the royal family (A Chronology: The House of Saud 2005).

Who drove these changes? While history and politics in Saudi Arabia have always been very male-centric, women's contributions should not be ignored. It is true, however, that the contributions of women were primarily those of elite women, connected to men in power. For example, Princess Effat, who began advocating for girls education alongside her husband, King Faisal, became a significant figure in this particular issue, with Effat College in Jeddah being named after her. Many other princesses used their elite position and connections to take action in starting programs and organizations for women's issues. Princess Hussah, who was married to

King Abdullah, founded the National Charitable Institute for Home Health Care, while Princess Moodi bint 'Abd al-'Aziz became actively involved in organizations created for women, such as one on art preservation (Zuhur 2011, 372). There has also been progress in the expansion of women's rights due to kings wanting to create positive social change for Saudi Arabia. For example, in 2009, King Abdullah took major steps, firing many heads of the religious police and appointing the country's first woman minister, Nora Al Faiz, to be Deputy Minister of Education (The History of Saudi Arabia). Along with this, he also gave women the right to vote and run in elections starting in 2015 (Epatko 2011). The power elite have also, therefore, been significant in slowly building the rights of all women, even those without power or wealth. This makes Saudi Arabia one of the few places which expand women's rights through power from above as opposed to power from below, such as everyday women's advocacy and protests. There have been instances of everyday women taking action in calling for change, whether verbally or through direct action, but even without specific legal restrictions, both social views and the power of religious police have caused major negative repercussions for those who made these attempts. For example, Loujain al-Hathloul and her friend Maysaa al-Amoudi were arrested after attempting to drive into neighboring United Arab Emirates in 2014. Despite the fact that no specific Saudi laws prohibit women driving - the driving ban is solely based on socio-religious custom - they were still arrested and referred to the terror court because of the social media posts they had made previously (Saul 2015). The women were later released yet still faced death threats, both online and in public (Dubai 2014). Along with the negative consequences for these women came no form of progress; the women's actions caused no change in the system and only highlighted the

government and religious police's willingness to take serious action to prevent women from taking such stands. This event also made clear that for progress in the expansion of women's rights in Saudi Arabia to be made, it must hold some form or benefit for Saudi Arabia as a whole.

3. The Current State of Women's Rights in Saudi Arabia

In regard to the driving ban, there have been some tentative signs of progress. The Saudi King's advisory council suggested that the government lift the ban, on condition that women be over 30, off the road by 8pm, and drive in conservative clothing without makeup (Henderson 2014). Despite this, the recommendation does not seem to have been pursued further and the driving ban continues to be enforced. As a result, women or their families are forced to hire drivers, or if they are not able to afford that, they are instead forced to rely on male relatives for their everyday transport.

Many times in Saudi Arabia, studies making radical claims without sufficient evidence and without supporting studies showing similar answers are released to justify these restrictions placed on women. One example was a study released after the Minister of Labor permitted Saudi women to work in malls as cashiers (Dhiyab, 2012). As would be expected, the Wahhabi religious authorities were against this action due to the interaction between men and women. After the religious protests were unsuccessful, a study was released which claimed a woman working was forbidden under Islam as it was seen as exploitation. It was also claimed that work was degrading for women (Dohman 2013). In addition a French study was translated which claimed that 2.5 million French women did not want equality and instead wanted to follow the lifestyle and lives of Saudi women (al-Bishr 2013). The study was seen to be highly inaccurate due to the miscommunication to

French women on the specifics of the daily restrictions Saudi women face, along with the unlikely statistic of 2.5 million. The reliability of the study is unknown, yet studies like this one continue to be released to argue against women's rights (Dohman 2013).

Apart from legal restrictions, women in Saudi Arabia are also often prevented from work, travelling and other personal freedoms due to cultural restrictions. The traditional beliefs of many Saudis, both men and women, are that women in the family should be wives, mothers and caregivers, while the men of the family should work and earn money. These traditional beliefs, therefore, also become a main reason that many women are prevented from working in their family, as well as a reason that many Saudi women do not choose to work even when there may be an option. In Saudi households these pressures may not only come from the men in the family but rather from mothers, sisters and other female relatives. Therefore, in Saudi Arabia the pressure for modernization runs parallel to the pressure to maintain the Saudi traditions and beliefs.

As mentioned before, many of the restrictions women face in Saudi Arabia, including the guardianship law, driving ban and strict gender segregation, are due to the strictly conservative interpretations of religious law on which Saudi society is based. Legal decisions of all kinds in Saudi Arabia are based on *Sharia*, or Islamic Law, which is founded on the highly traditional school of Islam known as Wahhabism. As mentioned above, Saudi Arabia's connection and reliance on Wahhabism dates back to the foundation of Saudi Arabia, when the royal family agreed to allow this ideology to be the main belief for all those the Wahhabis conquered. Although this agreement was made nearly 100 years ago, its influence has continued to impact the lives of Saudis today. Unlike many perceptions in

the Western world, the interpretation of the *Quran* and *sharia* that Wahhabism proposes does not represent the only interpretations of Islam nor does it represent the beliefs of all Muslims in the world. Despite many claims, the *Quran* explicitly addresses the issue of equality, stating, “All people are equal, as equal as the teeth of a comb. There is no claim of merit of an Arab over a non-Arab, or of a white over a black person, or of a male over a female” (Heyneman 51). This seems to clearly disprove any need of laws which treat women as minors and unequal to men, yet Wahhabism is used as reason to further enforce these laws, seeing itself as the true version of Islam. An example would be one of Saudi Arabia's laws which requires all women to cover up when in public in an *abaya* (a long black cloak) and Muslim women must also wear a headscarf, or *hijab*. In the *Quran*, this conservative clothing is not a requirement, as there is “no compulsion in religion” (*Surah* 2:256), which is why many Muslim women around the world do not always choose to wear *hijabs*, *abayas*, *niqabs* or *burqas*. The *Quran* simply requires that both men and women dress modestly (*Surahs* 24 and 31). Despite this, Wahhabism and many traditional Muslims believe all areas of a woman's body are *awrah*², which implies that a woman's body is provocative and must be hidden from unrelated men, therefore requiring women to cover up at all times in public (Women’s Rights in Saudi Arabia 2013). This is only one of the many examples in which Wahhabism, sharia law and conservative religious leaders of Saudi Arabia have influenced many of the repressive laws against women in Saudi Arabia, a basis deeply engraved in Saudi Arabia’s history and difficult to move away from.

² *awrah* is the Islamic term for intimate body parts that must be covered from others with clothing.

4. The Connection Between Economic Developments and Women's Rights

In this section I will briefly highlight the specific economic needs of Saudi Arabia, specifically the nationalization of the workforce, diversification of the workforce and improvement of the Saudi image abroad. I will evaluate their importance and how each issue impacts Saudi Arabia economically. I will then evaluate four key areas of women's rights that all have occurred in correlation with the specified economic needs. The four areas that will be discussed are the expansion of educational opportunities for women, the expansion of women in the workforce, legal changes in favor of women's rights, and the expansion of publicity about women.

In Saudi Arabia, one of the most important economic objectives is the desire to nationalize the workforce, otherwise known as "Saudization." In 2014, about 56% of the workforce in Saudi Arabia was made up of foreign workers (Jones and Al Omran 2015). This is mainly due to the fact that foreign workers coming into Saudi Arabia have a lack of restraining legislation on their schedules and salaries, making them cheaper workers for Saudi businesses (Le Renard 624, 2008). In 1993, with a workforce mainly made up of foreign workers, there was a \$3.73 billion transfer of funds overseas in the form of remittance. With the cheaper and easier to manage expatriate workforce, competition for employment is greater for Saudi nationals in employment, leading to an overall 2.8% unemployment rate among Saudi nationals in 2014, with women having an unemployment rate at 20.10% ("Saudi Arabia Unemployment Data" 2013). According to Saudi Arabia's National Development Goals (NDG), some of the objectives for future years are to enforce Saudization to reduce the unemployment rate and

give a competitive advantage to Saudi nationals. In addition, the NDG of Saudi Arabia highlights the importance of not only the young population of Saudi Arabia but also women in the Saudization of the workforce and increasing overall income and economic development (Millennium Development Goals 34-35, 2013). The nationalization of Saudi Arabia has thus become one of Saudi Arabia's most important issues to be tackled in the coming years, and women, being highly educated potential workers and making up about half of the population, have been one of the main focuses to tackle the issue. Women's involvement in Saudi Arabia's economic growth is significant not only in the workforce but also in the home, for increased income is likely to be invested in children's health and education (Jamjoom, 2013). Women are also high consumers, especially with an increased income, and will consume more products such as handbags, perfumes, and other luxury items. The high consumption rate, therefore, also becomes significant in helping the Saudi economy. Women hold high potential in the economic growth and stability of the nation and many are beginning to realize the importance of harnessing the possibilities they offer.

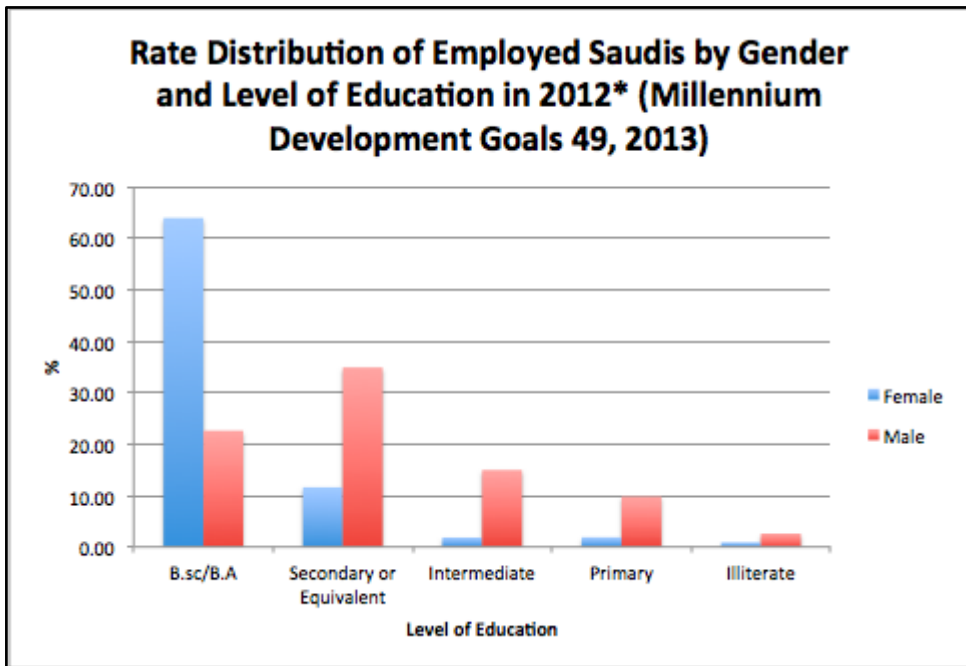
The second significant issue in Saudi Arabia is the need to diversify the business and industrial workforce. As with all businesses and companies, there is always a need to have ability and knowledge to appeal to all age ranges and sexes, with women also being large consumers for many products, which means the need to include female influence and output. Furthermore, with Saudi women making up 60% of college graduates, there is a large proportion of highly educated students whose potential is not being accessed, limiting the potential development of companies and businesses. Women's potential includes creativity, new ideas for product design, marketing knowledge, etc., and can assist companies in expanding product

reach and accessing a greater consumer base. For companies producing feminine products such as tampons and pads, women are essential in providing ideas in marketing, practicality and design. Although women make up 60% of college graduates, very few of them are involved in work or education, which are seen as 'men's' professions. Therefore, there has developed a need for diversification and an increase of women in all professions, including those previously dominated by men or involving interaction with both genders.

Aside from issues related to the Saudi workforce, many other issues less obviously impacting the Saudi economy have become a focal point in the development of Saudi Arabia. With the rise of social media across the world came the rise in world pressure on Saudi Arabia. The issues and inequalities women in Saudi Arabia face daily have become further publicized, with more and more people abroad highlighting the oppression. Improving the rights of women and lifting the oppression they face has become important for Saudi Arabia to improve its image abroad. Improving the image acts not only to lower the world pressure they face from global organizations but also to improve their partnerships with countries for trading. For example, the Swedish government recently stopped an arms deal with Saudi Arabia, a defense agreement that had lasted a decade, after news that the Swedish Foreign Minister Margot Wallstrom was being prevented by the Saudis from talking about democracy and women's rights at a meeting in Cairo with the Arab League (Tharoor 2015). Although some acts are said to be only for the sake of improving the Saudi public image, while not providing any real improvement in equality for women, many have led to the slow but important expansion of their rights.

As mentioned above, education has clearly been one of the major areas in which women have been empowering themselves, with about 60% of college graduates being female. The large number of educated females makes the group a good target for potential workers in the nationalization of Saudi Arabia. There is a clear correlation between women's education and their participation in the workforce, and as shown in figure 1 below, 94.2% of working women hold secondary school certificates or higher.

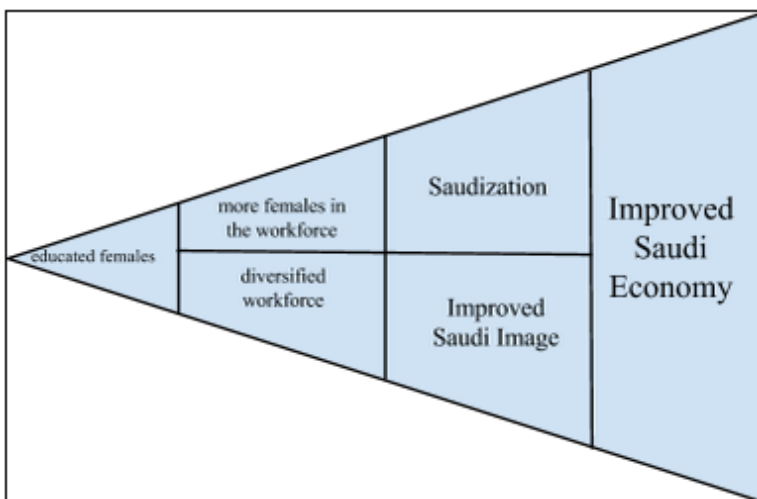
figure 1.



* “% share of employed males to total male population, and % share of working females to total female population.” (Millennium Development Goals 49, 2013)

This high level of female education has been a main motivator for Saudi Arabia to invest more in education so as to increase labor force participation,

specifically with women. In the second development plan, 16% of the \$142 billion saved was dedicated to education (Alsuwaigh 69, 1989). As mentioned above, although the percent of females in university is high, very few of them are studying areas that are usually considered men's, such as sciences, engineering and architecture, thereby limiting their access into many fields of work (Baki 5, 2004). Between 1995 and 1999, out of 114,000 graduates only 10,000 graduated with engineering degrees, leading the engineering workforce to be made up mainly of expatriates (Baki 6, 2004). In order to combat this problem and increase diversification in the workforce, King Abdullah established the first mixed gender university, King Abdullah University of Science and Technology in 2009, which, through scholarships, promoted female participation in areas of science and technology previously dedicated to males, such as biological and environmental science and engineering; computer, electrical and mathematical science and engineering, and physical science and engineering (Laessing and Alsharif, 2009). Scholarships for education abroad have also encouraged an increase in education for all females as opposed to mainly the elite and wealthy females. Additionally, throughout Saudi Arabia, the development of technology has enabled females to study courses with male students through video conferencing, allowing them to work in gender-separated classrooms (Le Renard 616, 2008). The need for Saudization and a diversified workforce, therefore, began with creating more educational opportunities for women in all areas of education. Figure 2 below highlights the connection between establishing good female education and an improved national economy.

figure 2.

As discussed above, educated females in all areas of study allow for a larger and more diversified workforce, which in turn, to be discussed in the following paragraph, leads to Saudization and a better Saudi image, all contributing to improving the Saudi economy.

Ever since 1980 it has been part of Saudi Arabia's development plans to create more areas for women's work (Le Renard 624, 2008). Since 1980, Saudi Arabia's GDP has dropped by roughly \$16,500, leading more and more Saudis to question the handicap placed on half the population and thereby limiting their full involvement in improving the Saudi economy (Coleman 90, 2004). The expansion of women's work has also become important due to the flood of educated Saudi women increasing each year. In the year 2000, approximately 78,000 women graduated from universities in Saudi (Doumato 570, 1999). In addition, Saudi women are seen to have a generally better work ethic than Saudi men, due to the importance to them of

being able to go out and be active in everyday life, as opposed to working just for the sake of earning money (Le Renard 624, 2008).

Despite this, the Saudi government has continually struggled with creating jobs for women while still following the strict Islamic beliefs of gender segregation and the protection of women's dignity and virtue. The government thus created movements that were able to find balance through establishing women-only work areas. A significant improvement came in 1980, when Riyadh opened its first women's branch of Al Rajhi bank, with 13 female branches opening by 1982. This created new job opportunities for women working in the banks and new economic opportunities for women joining the female branches. Since 2004, women have been able to open firms under their own name and have been creating their own enterprises in which they only hire women (Le Renard 624, 2008). Saudization has also been seen to become a strong influencer in the expansion of women's work in the Kingdom. King Fahd stated during a meeting in 1997 with the Council of Ministers that 79% of the 660,000 jobs planned for Saudization should be dedicated to women under the provision they follow Islamic standards (Doumato 571, 1999). Although not completely successful, the Minister of the Economy launched a program in June 2005 that aimed to replace all foreign workers in lingerie shops with Saudi women. Banning men from lingerie stores would allow Saudi women to work segregated from men. Despite this, the government postponed the measure because women would be working in mixed gender malls and because Saudi shop owners felt Saudi women would be too expensive to hire over immigrants. (Le Renard 625, 2008). Nevertheless, because Saudi women are known for being good workers and being highly educated, work opportunities for them are continuously being expanded.

Economic needs have also brought about legal changes for women's rights. King Abdullah has made major improvements in women's rights through incorporating them into various areas of government, giving them the right to vote for the Chamber of Commerce and loosening the guardianship rule for business women. In order to help diversify the government, while also improving the Saudi image abroad, the government appointed 30 women, or 20% of the 150 members to the Consultative Council, or *majlis al-shura*, each with an academic background and from a variety of regions of the Kingdom. Not only has this action improved the Saudi image drastically in respect to gender equality, it has also allowed for the Council to consult with women on topics related to family and women, such as dowries, marriage and breastfeeding.

There are still improvements to be made in the power and equality of women in the Council. Many of them have had their roles limited to women-related aspects, and they have stated they are there for symbolic reasons as opposed to being real members (Le Renard 620, 2008). However, many have been hopeful about the changes their involvement can bring. Starting in 2015, women have also been given the right to vote in municipal elections. One of the main propellers towards this action has been once again the desire to transform the image of Saudi Arabia abroad. In a speech to the Shura, King Abdullah stated, "We refuse to marginalize the role of women in Saudi society," highlighting the importance of allowing women to have the basic rights other countries have already granted women (MacFarquhar 2011). Another slight improvement in Saudi Arabia has been the loosening of the guardianship rule, specifically for businesswomen. Perhaps in an attempt to expand business opportunities for women, King Abdullah issued a decree permitting women to open businesses without male guardian

approval, also expanding on the process of Saudization by producing more Saudi-run businesses in the Kingdom (“Looser Rein” 16, 2010). Saudi Arabia's National Development Plans highlighted the importance of women in investment and business administration in different economic activities, with 85,197 women-owned commercial enterprises by the end of 2012. The development plan highlighted the importance in further improving these areas to increase economic participation (Millennium Development Goals 50, 2013). All have been significant in the economic empowerment of women, while still assisting in the improvement of Saudi Arabia's own national economy.

Since September 11, 2001, when 15 of the 19 terrorists who destroyed the World Trade Center were Saudi, and since the Saudi bombings in 2003 in which suicide bombers targeted three compounds for expatriates in Riyadh, Saudi Arabia's government has been battling the rise of terrorism and extremism, while also attempting to promote reform and modernization in Saudi Arabia (CNN Library, 2014). The increase in terrorism influenced Western countries to begin pressuring Saudi to make reforms and prevent future terrorist attacks. To protect the legitimacy of the Kingdom, the government began reforms towards rebuilding the nation, targeting areas of the population affected by particular issues. Women in Saudi Arabia became a primary focal point, leading to an increase in not only the steady expansion of their rights but also an increase in publicizing the issues and the need to improve their rights. State institutions and national media have now been extensively discussing women's rights with Saudi men and women in positions of power, such as members of the government, official and unofficial *ulama* and female and male intellectuals with different orientations (Le Renard 617, 2008). As a result, the increase of both the

overall discussion about the issue and elite women advocating for equality has somewhat erased the taboo on women's rights which was previously common in Saudi Arabia. An extensive national dialogue was held in Medina in which half of the 70 participants involved were women. Topics discussed included women's rights and duties, women and work, women and education and women and society. The debate resulted in 19 final recommendations in which the male participants presented to the Crown Prince Abdullah, and the female participants presented to his spouse (Le Renard 618, 2008). Although overall progress is still slow, a female participant from the dialogue felt that the fact that women's issues in Saudi Arabia are becoming so widely discussed, not just abroad but in the Saudi media, is a significant representation of the change that has occurred so far (Le Renard 619, 2008). While some say the increase within Saudi media worldwide is only to make the Saudi image better without real improvements, the impacts thus far on Saudi women have been significant in propelling the expansion of their rights.

Conclusion

In conclusion, many of the improvements in the rights of Saudi Arabia's women have occurred in response to the country's economic needs. With nationalization, diversification and improvement of the Saudi image have come changes in legal rights, educational rights, work opportunities and increasing images of Saudi women in publicity and the media. Over the past few years these economic needs have allowed for women to gain access to new areas of study, vital positions in the Chamber of Commerce, a decrease in female unemployment rates and more extensive publicity about women in Saudi media, all significant steps in the expansion of their rights.

These correlations have become vital in moving Saudi women away from an image of oppression to an image of economic opportunity. In Saudi Arabia, where women have become objects of conflict between religion and Westernization, the argument for Saudi women being important in the improvement of Saudi Arabia is allowing for the religious and traditionalists to increase needed support for their rights.

Bibliography

- "A Chronology: The House of Saud." PBS. August 1, 2005. Accessed June 22, 2015.
- Al-Bishr, Badria. "Saudi Women, Objects of French Envy?" Al Arabiya News. March 21, 2013. Accessed July 6, 2015.
- Alsuwaigh, Siham A. "Women in Transition: The Case of Saudi Arabia." *Journal of Comparative Family Studies* 20.1 (1989): 67-78. JSTOR. Web. 13 June 2015.
- Associated Press. "Easing of Saudi Driving Ban Possible." Daily Mail Online. November 8, 2014. Accessed June 22, 2015.
- Baki, Roula. "Gender-segregated Education in Saudi Arabia: Its Impact on Social Norms the Saudi Labor Market." *Epa Education Policy Analysis Archives* 12.0 (2004): 28. Web.
- CNN Library. "September 11th Hijackers Fast Facts - CNN.com." CNN. September 8, 2014. Accessed July 6, 2015.
- Coleman, Isobel. "The Payoff from Women's Rights." *Foreign Affairs* 83.3 (2004): 80. JSTOR. Web.
- "Millennium Development Goals" *Millennium Development Goals*. Information Office of the Royal Embassy of Saudi Arabia, n.d. Web. 08 June 2015.
- Dhiyab, Rahmah. "Saudi Arabia Orders Factories To Hire More Female Workers." Al-Monitor. April 24, 2012. Accessed July 6, 2015.
- Dohman, Ahmad, and Steffi Chakti. "In Saudi Arabia, Science and Religion Collide Against Women." Al-Monitor. June 23, 2013. Accessed June 22, 2015.

- Doumato, Eleanor Abdella. "Women and the Stability of Saudi Arabia." *Middle East Report* No. 171. *The Day After* (1991): 34-37. *JSTOR*. Web. 07 June 2015.
- Dubai, Tahira. "One Brave Woman's Protest Against Saudi Arabia Female Driving Ban." *Daily Mail Online*. December 1, 2014. Accessed June 22, 2015.
- Epatko, Larisa. "Saudi Arabia: Women Can Vote, Starting in 2015." *PBS*. September 26, 2011. Accessed July 3, 2015.
- Fisher, Max. "9 Questions about Saudi Arabia You Were Too Embarrassed to Ask." *Vox*. January 26, 2015. Accessed June 22, 2015.
- "Global Gender Gap Report 2014." *World Economic Forum*. 2014. Accessed June 22, 2015.
- Hamdan, Amani. "Women and Education in Saudi Arabia: Challenges and Achievements." *International Education Journal*, 2005, 42-64.
- Henderson, Barney. "Saudi Arabia Considers Lifting Ban on Women Drivers." *The Telegraph*. November 8, 2014. Accessed July 9, 2015
- Heyneman, Stephen P. *Islam and Social Policy*. Tennessee: Vanderbilt University, 2004.
- Jamjoom, Mounira. "Make Saudi Women Visible -- Boost the Economy - CNN.com." *CNN*. January 23, 2013. Accessed July 3, 2015.
- Jones, Rory, and Ahmed Al Omran. "Saudi Arabia Puts Squeeze on Foreign Workers." *The Wall Street Journal*. March 6, 2015. Accessed July 7, 2015.
- Laessing, Ulf, and Asma Alsharif. "Saudi Arabia Opens First Mixed-gender University." *Reuters India*. September 23, 2009. Accessed July 6, 2015.

"Looser Rein, Uncertain Gain: A Human Rights Assessment of Five Years of King Abdullah's Reforms in Saudi Arabia." *Human Rights Watch*, 2010.

Macfarquhar, Neil. "Saudi Monarch Grants Women Right to Vote." *The New York Times*. September 25, 2011. Accessed July 6, 2015.

Renard, Amélie Le. "'Only for Women:?' Women, the State, and Reform in Saudi Arabia." *The Middle East Journal Middle East J* 62.4 (2008): 610-29. Web.

"Saudi Arabia GDP Growth Rate | 1969-2015." Saudi Arabia GDP Growth Rate | 1969-2015 | Data | Chart | Calendar. Accessed June 22, 2015.

"Saudi Arabia Profile." *BBC News*. BBC, 29 May 2015. Web. 13 June 2015.

"Saudi Arabia Unemployment Data." Quandl. December 1, 2013. Accessed July 7, 2015.

Saul, Heather. "Saudi Women Jailed for Driving 'released from Prison' after Two Months." *The Independent*. Independent Digital News and Media, 13 Feb. 2015. Web. 13 June 2015.

Tharoor, Ishaan. "At Last, a Western Country Stands up to Saudi Arabia on Human Rights." *The Washington Post*. March 12, 2015. Accessed July 7, 2015.

"The History of Saudi Arabia." *The History of Saudi Arabia*. Information Office of the Royal Embassy of Saudi Arabia, n.d. Web. 13 June 2015.

"Women's Rights Gain Traction in Gulf Arab States, but Inequality Persists." Freedom House. February 11, 2009. Accessed June 22, 2015.

"Women's Rights in Saudi Arabia: Wahhabism vs. Islam." Morocco World News RSS. April 16, 2013. Accessed June 22, 2015.

Zuhur, Sherifa. *Saudi Arabia*. Santa Barbra: Middle East in Focus, 2011.

Print.



Which Model is Feasible for Hong Kong? A Comparison among the Democracies in Britain, Mainland China, and Hong Kong

Zhimin Lin

Author background: Zhimin grew up in China and currently attends The Affiliated High School of South China Normal University in Guangzhou. His Pioneer seminar program was titled "Comparative Political Institutions of the World."

I . Introduction

Since the administration of Hong Kong was returned to China on July 1st, 1997, its democracy has been a unique system in the world. On one hand, it is controlled by the Chinese central government—the elected governor must be approved and appointed by the central government; on the other hand, it has been granted a high level of autonomy; for example it remains capitalist and has the right to legislate. This sort of fusion of Chinese and Western democracy makes Hong Kong a special case to look at, and has also led to a long dispute which has intensified in recent years. The focus of controversy is mainly on suffrage: some pan-democratic political organizations use many approaches to rebel in an attempt to force the central government to permit the Legislative Council (LegCo) and Chief Executive to be elected in general election, and to demand that the candidates for Chief Executive not be limited by the central government. They also launch activities which aim to oppose any policy from the central government, and claim that a pure Western democracy should be conducted. However, much evidence, research and history indicate that Hong Kong is a unique place which

requires a new innovative model of democracy that is different from both Chinese socialism and the former British democracy—the Hong Kong Model.

In this research, I shall discuss the four different democracies that are available to Hong Kong, and try to approach the conclusion that only a combination of Chinese and British democracies is feasible in Hong Kong. I shall start with the history of Hong Kong beginning from when it was ceded to the UK in 1840, and discuss the democracy of Hong Kong during its colonized period. After that, I shall introduce and discuss the three different democracies which could have been applied in Hong Kong when it was returned to China, trying to find their advantages and disadvantages. After that an argument will be made based on the previous chapter to clarify why the combination is the only viable democracy for modern Hong Kong. In the next chapter, democracy in Macau will be briefly discussed and compared to democracy in Hong Kong since Macau shares the same political position in China nowadays, and it experienced a history similar to that of Hong Kong: they both had been colonized and then returned in the late 1990s. Why does the system in Macau seem less controversial and produce more harmony than the one in Hong Kong? Lastly, a conclusion will be drawn to finally point out which democracy system is suitable for Hong Kong, and is the right path for it to remain a stable and growing society.

II. The Concept of Democracy

Before going into the topic of this essay, it is necessary to briefly discuss the concept of democracy, as it is such a widely used and defined word that it has multiple levels of meaning in different parts of the world, and

differences in the definition of democracy have long been a source of conflict between the Western society and the rest of the world. The lack of a congruent definition of democracy in the world means that the discussion of this topic requires people to take the different values and real situations on which the specific democracies are based into consideration.

In traditional Western values, there are some commonly recognized elements for a real democracy. In Robert A. Dahl's book *On Democracy*, he proposed the following factors as the minimum requirement for a democratic state: elected officials; free, public, regular elections; freedom of speech; various sources for information; autonomy of organizations; and the inclusive rights of citizens. It is generally believed that a democracy without these elements—or some of them—will not be truly democratic.

When people in most parts of the world—including east and south Asia—are asked about their attitudes toward the factors mentioned above, most of them will more or less support them. There are studies indicating that Asian people's agreement with these factors and the general Western idea of democracy is increasing (Larry Diamond, 2013, Pp. 20-22). This is not astonishing, as it is not hard to understand the great disadvantage caused by the high centralization of political power in one individual or family and the lack of the former values: corruption, low efficiency, abuse of violence to suppress opposite voices, etc. There have been so many examples in the dark history of dictatorship.

However, the wide acceptance these creeds receive does not necessarily mean that every regime in the world would have a similar democracy. Though it has long been argued that Western democracy is the system that is most ideal to the development of a nation and the human rights of its

citizens, it is also true that practically this model is not suitable for a broad range of countries—at least for the time being. These countries (or regions) do not have necessary prerequisites, such as the weakness or absence of cultural conflicts and the substantive altitude and flexibility when making compromise under the democratic frame, for a Western democracy to be established (Robert A. Dahl, 2012, Pp.124; Larry Diamond, 2013, Pp.173). With this logic, if there is an alternative model that is beneficial to a state and its people with some extent of democracy and some variation from the traditional model, it would be reasonable and wise for the country to adopt this system.

That is the reason why I also take factors other than political system itself, such as history, economy and stability of a society into consideration when I discuss the three democracies in this paper, especially the ones of Hong Kong and China, and try to explain the influence of them on the Hong Kong model. It is clear that Hong Kong has a special political position: it is a “Special Administrative Region.” This being said, Hong Kong is not an independent country, which is the foundation of all discussion about its democracy. It is also clear that Hong Kong had been colonized for over a century by Britain, so a large part of its perception of democracy is close to the Western concepts, while some Chinese values still remain. These two points form the foundation of the Hong Kong model and are pivotal to the understanding of it.

As for the Chinese political system, although I call it a “democracy,” it is neither equal nor similar to the Western democracy, because China has a recognition of democracy that is different from the Western one. The Western democracy emphasizes the electoral process, political pluralism and

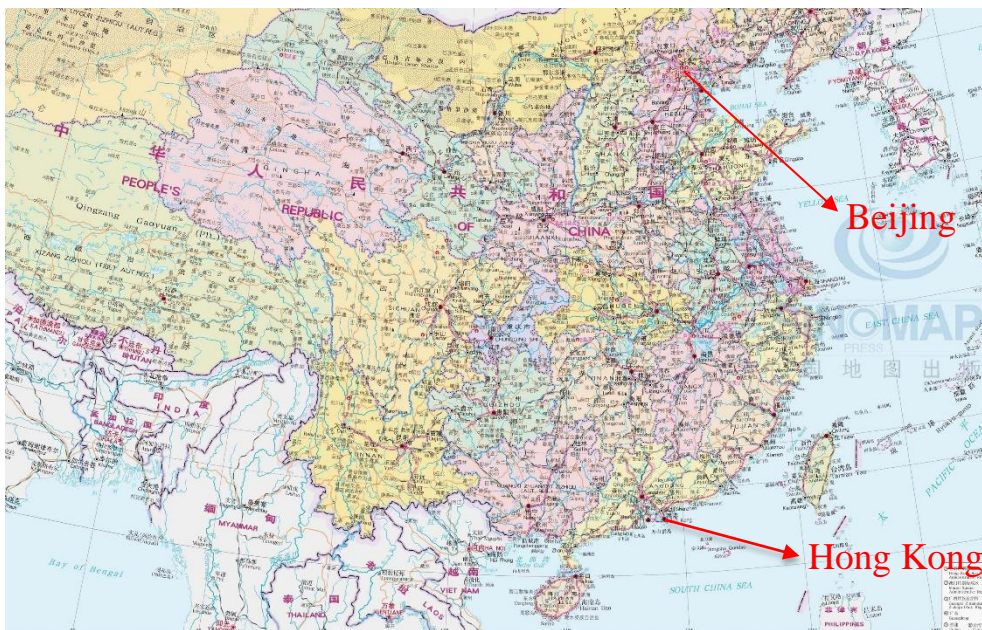
freedom of expression and belief (Freedom House, 2015), while the Chinese system focuses mostly on the output of the system, such as the prosperity of economy and harmony and stability of the society. This is not denying that Chinese people, including political leaders, despise the democratic elements above; instead, it is because they agree that these elements are not the priority at this stage compared to economic growth and development of the society. However, it is probable that the essence of these values increases as the nation becomes more developed. The single-party system determines that the power in China is highly centralized; nevertheless, the country has been approaching its goal in the last three decades and more and more effectively protecting people's basic rights—suffrage, ownership of properties and so on. Hence, the use of the phrase “Chinese democracy” here refers to the model that is currently feasible in Mainland China based on Chinese values and the actual circumstances.

III. A General Introduction to Hong Kong

Hong Kong Special Administrative Region of the People's Republic of China is located on the east bank of the estuary of the Pearl River on the southern China coast line. Adjacent to Shenzhen, Guangdong Province, Hong Kong covers an area of around 11 million square kilometers including Hong Kong Island, Kowloon Peninsula, the New Territories and 230 other small islands. The main topography in Hong Kong is hills. The climate type in Hong Kong is a sub-tropical climate, so in the summer it occasionally suffers from the influence of typhoons. The population in Hong Kong is 7.2 million¹ (Government of Hong Kong, 2015). Most of the population consists

¹ By the end of 2014.

of Chinese people whose native place is the Pearl River Delta. The remaining 5% or so of the population are foreigners, most of whom are Indonesian, Filipino, European, or Indian.



The Location of Hong Kong (sources of maps at the end of references)

Hong Kong's economy is highly developed and concentrates on the service industry. The gross domestic product of Hong Kong in 2014 was 2.25 trillion Hong Kong dollars with a growth of 5.35%, and the per capita GDP was 310 million Hong Kong dollars² (Government of Hong Kong, 2015). Hong Kong's economy is a capitalistic system which pursues a free market, and the government of Hong Kong constantly applies laissez-faire in the economy. For this reason, Hong Kong is considered one of the freest markets in the world. It is also the financial center of Asia and the world, and the third biggest gold exchange globally.

² Calculated by major expenditure component.



Full View of Hong Kong (Some Small Islands Not Shown)

IV . The History of Hong Kong and Its Political Situation before the 1980s

On January 26th, 1841, the commander of the British fleet, Gordon Bremer, seized Hong Kong Island with his British navy. On February 1st Bremer and Charles Elliot, Chief Superintendent of British Trade, declared to the Hong Kong residents that Hong Kong Island had been taken over and was to be governed by the United Kingdom. Though by this time the Treaty of Nanjing, which ceded Hong Kong Island to the UK, had not yet been signed, it was virtually this point when Britain's colonization of Hong Kong began. In 1843, the Constitution of Hong Kong and Hong Kong Royal Instructions were signed by Queen Victoria and announced. In August, 1843, China was

obliged to sign the unequal treaty, Treaty of Nanjing, officially ceding Hong Kong Island to Britain.

The initial colonized Hong Kong government consisted of the Administration Council and the Legislative Council. The role of the Executive Council was to advise the governor on policy, and that of the LegCo was to help him in law making. However, it should be noted that the power to legislate for Hong Kong was retained by Britain. It could be either an act of parliament or prerogative (i.e. an executive act) (Yash Ghai, 1991, 798).

In the next 55 years (before 1898), British colonists gradually shaped the political system in Hong Kong. At the very beginning, all power was monopolized by British colonists—mainly by the governor. For example, the first governor Henry Pottinger was simultaneously the Post of Envoy, Plenipotentiary in China, Chief Superintendent of British Trade and military commander; British army officer William Caine was Chief Magistrate, and also councilor in the Executive Council and Legislative Council. However, as British consortiums developed rapidly in Hong Kong, their demand for political influence and power also increased. In 1849, 1955 and 1894, activities to require political reform and establishment of an autonomic government formed of British constituency took place but all were turned down by the British government. Nevertheless, the British government realized that monopoly was no longer feasible, so from 1850 businessmen started to show up in the LegCo, and this was also the beginning of functional representation (Yash Ghai, 1991, Pp. 798).

In this period, the Legislative Council was also taking shape. Initially, the councilors were appointed by the governor freely, and there was no

disciplinary regulation for them. Only the governor had the power to propose a motion for a bill. From 1865, councilors were determined by their posts and then appointed by the governor. They were also ordered not to oppose the government's bill. But on the other hand, they gained the right to propose a bill as long as it was seconded by at least two councilors and did not relate to financial problems.

In 1860 and 1898, the other two unequal treaties, the Convention of Beijing and the Convention for the Extension of Hong Kong Territory, extended the colony first to Kowloon Peninsula and then to the New Territories. From the 1890s, as various Chinese national democratic revolution activities developed and expanded, the influence of Chinese people in Hong Kong was also increasing. The ruling class in Hong Kong had to open the power institution to some Chinese in the upper class. In 1926, the Executive Council regulated Chinese councilor positions for the first time. In 1929, there were three Chinese councilors in the Legislative Council.

After World War II, due to a series of positive factors, Hong Kong managed to recover from the war quickly and entered a golden age of industrialization and its economy soared. The growth in economy led to the rise of the middle class, so the power system was forced to accept intellectuals from more sectors of the society and became more pluralistic and representative.

There were a few significant changes in the Hong Kong political system between 1945 and the 1980s. First of all, after the 1967 anti-colonial riots, the Hong Kong government decided to implement the Home Affairs Project, which basically encouraged residents to participate in community affairs such as security and health. Communities started to have leaders and committees to manage affairs, and the government also sent representatives

to join in these committees. Secondly, more and more Chinese people became civil servants. From 1952 to 1981, the percentage of local people among civil servants increased from 95.56% to 97.85%. However, it should not be ignored that most of the senior positions in the government were still occupied by foreigners, especially in sensitive departments like the police department and the justice institutions. Thirdly, the most important transformation was the establishment of ICAC, the Independent Commission Against Corruption. ICAC was an independent institution established to counteract the worsening corruption situation in the Hong Kong government in the 1970s, after the shocking corruption of the senior police officer Peter Godber. ICAC was totally independent from any government offices, and was directly responsible to the governor.

From 1841 to the late 1970s, Hong Kong had been a colony of Britain with little democracy. The Legislative Council did not share the power to legislate, and councilors could not even oppose bills from the government. Hong Kong citizens did not have suffrage. Although the proportion of Chinese people in the government had been increasing, and the government seemed to care more about the public opinion over time, the power of decision-making was still controlled by a minority of foreigners. The political system in Hong Kong was totally led by the administration, and most of the population had little chance to vote and make political decisions.

V. The Democracy in Hong Kong from 1980 to 1997

After Governor Crawford Murray MacLehose visited Beijing in March 1979 and learnt that China would resume its sovereignty over Hong Kong in 1997, he immediately launched the arrangement of representative democracy in

Hong Kong. Within 15 years before Britain ended its reign in Hong Kong, the democracy developed rapidly. The main reason for this sharp change from suppressing democracy to enthusiastically advocating it, if not the only one, was that British government wanted to cultivate a democracy that was independent—or friendly and beneficial to the UK—and free from the control of the Chinese central government, so it could still exert influence on Hong Kong (Liu Manrong, 2009, Pp.144-147). Nevertheless, this soaring of democracy in 15 years laid a foundation for the unique, combined system applied after Hong Kong returned to China and played an important role in stabilizing the society. In this chapter, I shall introduce the “over-speeded” (Liu Manrong, 2009, Pp.135) progress of democracy in Hong Kong and the political model before its handover to the PRC. The democracy was mainly formed of three levels of representative institution: the District Council, the Urban Council, and the Legislative Council. Since the governor virtually had the power to make final decisions in legislature, this chapter only focuses on the representative electoral system developed in this period.

District Council

The core of the democratic change in the 1980s was the formation of the district council. In 1982, 18 district councils were established in Hong Kong, and the first election was held. From 1982 to 1994, five district council elections were held (1982, 1985, 1988, 1991, 1994). Similar to the British majoritarian democracy, the election for district councils in Hong Kong also applied the simple majority methods. In each constituency, the candidate (or two candidates) who won the most votes got elected.

In the first election, government officers, appointed councilors and elected councilors took 1/3 of the councils respectively. There were also 57 natural

councilors, who consisted of 27 Chairmen of Rural Committees and 20 councilors of Urban Council. In 1985, government officers no longer held positions in District Councils, and their proportion was devoted to the elected candidates. District Councils were also connected with the Legislative Council. All District councilors were divided into 10 groups, each of which represented about 500,000 citizens, and elected 10 Legislative Councilors. In the 1988 election, Urban Councilors were no longer appointed to be District Councilors and the total number of District Councilors dropped. In the 1991 election, election of Legislative Councilors by the District Council was terminated. From then on, the influence of the District Councils decreased (Liu Manrong, 2009, Pp.365-366). In the last election before Hong Kong's return, all District Councilors were elected except 27 Chairmen of Rural Committees.

Urban Council

The Urban Council was established in January 1st, 1936, mainly supervising food, health, environment and entertainment (Zhou Jianhua, 2009, Pp.43). In 1973, the British Hong Kong government abolished all positions for government officers and made the Council financially independent. From 1983 to 1985, the number of appointed and elected councilors on the Urban Council remained 15. However, in the 1989 election, 10 more councilors were elected by District Councils, and in 1995, all councilors other than 9 from District Councils were all generated from general election. The general election for Urban Councilors had applied the relative majority method between 1983 and 1995. The candidate who won the most votes got elected.

Before the 1980s, the qualification for Hong Kong citizens to become voters for the Urban Council election was high and thus excluded the major

population. A citizen had to have either passed the HKCEE³ or had higher education records or be a businessman or professional to have the right to register as a voter for the Urban Council election. It is very clear that these rigorous limitations caused the constituency to be limited to wealthy and highly educated people, and the Council did not represent the majority of society (Zhou Jianhua, 2009, Pp.45). But in 1981, all these limitations were abolished, and any citizen over 21 who had lived in Hong Kong for seven years could register as a voter except some people with conditions such as commission of crime.

Legislative Council (LegCo)

The Legislative Council was the highest level of representative institution in Hong Kong's democracy. Before 1985, Legislative Councilors were all appointed by the Governor of Hong Kong. In the 1985 election, for the first time 12 Legislative Councilors were elected indirectly from 9 functional groups including business, industry, finance, social service, education, law, medicine and engineering. The functional groups were divided into two types: those "related to economy and society," such as the functional groups of business, and others that were formed with professions, like lawyers, teachers, etc. In the first type, members of the organizations that formed the functional groups automatically became constituents, and each organization was a single constituency. In the second type, members of professional organizations automatically became constituents, but the whole functional group voted as one constituency.

³ Hong Kong Certificate of Education Examination. It was the examination that Hong Kong students took when they finished their middle school education. The exam is still being conducted nowadays.

The election applied a type of absolute majority method. Constituents could list their candidates on their ballot tickets and rank them. In the first round of counting, the first candidates on every ballot were counted, and if one of them won over 50% of votes, this candidate was elected. If not, the second candidates on ballots were counted and the votes were added to the first round. The procedure was repeated until at least one candidate won. In 1988, 1991 and 1995, the election procedure changed to another type. Constituents still ranked candidates on their ballots, but when ballots were counted, if no first-ranked candidates won over 50% of the votes, then the votes for the least voted candidate would be diverted to the second candidate on the ballots.

In addition to functional groups, 12 seats in the Legislative Council were elected by another approach—the election committee. The committee consisted of 12 election groups, which were formed of the Urban Council and District Councils. The method for the election was the same as that for functional groups in 1988 (except the 1985 election).

In 1991, the election for Legislative Council included direct election for the first time. However, due to the contention between different political powers, the methods for direct election in 1991 and 1995 were different.

One third of Legislative Councilors (18 seats) were directly elected in 1991. Candidates only needed to win the highest number of votes (relative majority) to get elected. In each two-member district, each voter could vote for two candidates. However, the unusual point was that the most voted candidates could divert his or her extra votes to another candidate who was linked with him. This kind of design was strongly condemned by public opinion (Zhou Jianhua, 2009, Pp.51), because the democratic parties won 16

out of 18 seats and the institution was considered to favor these parties. Hence, in the 1995 election, the government switched to the simplest single-member district model. In 20 districts, each one shared one seat, and the most voted candidate won the seat.

VI. Hong Kong Model—the Combined Democracy

After the handover of Hong Kong in 1997, the political system was modified. According to the “one nation, two systems” policy proposed by Deng Xiaoping, Hong Kong was granted a high degree of autonomy. In this chapter, I shall discuss the current democracy in Hong Kong—I call it the “Hong Kong Model”—which is a combination of Chinese and British models. The discussion consists of two aspects: an electoral system and executive-legislative relations. In chapter seven, I shall discuss the central-local relations, and how that affects the democracy in Hong Kong.

Electoral System for the Legislative Council

The present Legislative Council in Hong Kong has 70 seats, with 30 of them elected from functional groups, 5 of them elected from the District Council, and 35 of them elected from a general election. The method for general election is List PR, which requires each party to put its candidates in a list, and constituents vote for lists rather than a single candidate. If a list wins a standard number of votes, a single candidate on the list is elected. If the remainder of the votes for the list reaches the standard number again, then another seat is granted to this list. The function for the standard number is:

$$Q = \frac{V}{N}$$

where Q is the standard number, V is the total votes for all lists, and N represents the number of seats in a district.

If the elected candidates are fewer than the number of seats, a maximum remainder way is utilized. Among lists that still have remaining votes, the one with a maximum remainder wins a seat, until all seats are distributed.

Compared to the majority system used by the colonial government, the List PR system reduces the chance of underrepresentation, and better reflects the will of constituents. Besides, the PR system makes it easier for small parties to win seats, because they only need to win part of the votes instead of an absolute majority.

In the 2012 election the system introduced a new section of election—the “super seats.” The five candidates for these seats were nominated by District Councils, and voted for by all constituents except those who voted in functional groups. The five most voted candidates won the seats.

Unlike some other countries such as New Zealand and Holland, the PR system in Hong Kong does not set a threshold for parties to enter the Legislative Council. Any party that wins a seat can take the seat. This defect causes fragmentation within the Council (Xie Yu, *Administrative Law Review*, Vol.16 2013, Pp.211). At present, 16 parties occupy seats in the Legislative Council—with 35 seats in total, and 10 of these parties were formed after the List PR system was first applied in 1998.

The fragmentation caused by the lack of threshold has led to serious problems. First of all, it encourages the formation of radical small parties. These parties don't have to take a moderate path in campaigns, because they only need to win a small portion of votes to get seats (Xie Yu,

Administrative Law Review, Vol.16 2013, Pp.215). Radical parties also have become more aggressive and have organized more and more extreme movements such as “Occupying the Central” and violent assault of the Legislative Council on Nov. 19th, 2014.

Secondly, there is no party in the Council that is close to being dominant. The Council is highly diversified, and coordinating interests of different parties becomes a heavy encumbrance on the daily affairs. Unavoidably, this situation has led to a low efficiency in the LegCo (Xie Yu, *Administrative Law Review*, Vol.16 2013, Pp.212). Though the general election only produces half of the councilors, those from functional groups are likely to be with different parties. The convoluted relationship and conflicts between parties in the Council is a significant defect of the democracy.

Nevertheless, the electoral system had successfully drawn public attention to politics, and encouraged people to participate in elections. In the 2012 election, the turnout was 53% for the general election and 52% for the “super seats” election (Government of Hong Kong, 2012).

Electoral System for Chief Executive

For the time being, the Chief Executive of Hong Kong is elected by an election committee rather than by general election. Candidates for Chief Executive are nominated by no fewer than 150 members of the Election Committee. The Election Committee consists of 1200 representatives, and is divided into four subsectors, with each subsector including 300 representatives. The election for subsector representatives uses a relative majority system.

The election for Chief Executive applies an absolute majority system. If there is only one candidate, the candidate must win over 600 votes to be elected; otherwise, the election procedure will be terminated. In a contested election, if two candidates win the same number of highest votes, or if none of candidates win over 50% of the votes, then another round of election will be held between these two candidates.

After a candidate gets elected, the result is reported to the central government, and the central government can decide whether or not to appoint the elected candidate to be the Chief Executive.

The appointment of Chief Executive is the practical aspect that best demonstrates the combination characteristic of the Hong Kong Model. Unlike the former governors of Hong Kong who were appointed by the monarch of the UK, the Chief Executive is elected using the British majoritarian system, with the result still requiring the approval of the central government. It is an approach that allows a higher degree of democracy, but still allows the central government to apply its influence when necessary.

The Executive-Legislative Relations

The executive-legislative relations in Hong Kong have been somehow tricky and incongruous (Yang Yilong, Liu yanjun and Yin Cunyi, *Studies on Hong Kong and Macau*, Vol.8, 2012, Pp.107). The main reason is that the Chief Executive is elected neither by parties nor by the Legislative Council, but by the Election Committee. Also, any candidates for Chief Executive with a party background must declare disconnection with the party after inauguration, so the Chief Executive and his administrations are “cut apart”

from parties (Zhou Ping, *Thinking*, Vol. 30, 2004, Pp.53), and cannot rely on any party to magnify their influence in the LegCo.

However, acts and laws must be signed by the Chief Executive before coming into force. If the Chief Executive refuses to sign an act (or law), the LegCo has to vote again and get two-thirds positive votes to force the Chief Executive to sign the act, but he has the power to dismiss the LegCo. Nevertheless, if the new Legislative Council still passes the act, the Chief Executive must resign.

Moreover, one fourth of the Legislative Council can put forth a united proposal for investigation on the Chief Executive, and the Council can entrust Chief Justice of the Court of Final Appeal to form an investigation committee. If the charges are supported, a two thirds majority of the LegCo can pass an impeachment and report to the central government for approval.

In addition, there is another way that the power of the legislature is limited. All laws made by the LegCo must be sent to the NPC to be filed. If the NPC refuses to file, it can send the law back without amendment, and the law is nullified immediately. Thus, the NPC has the final power to determine the fate of a law.

Overall, the executive-legislative balance in Hong Kong leans toward the executive side. The central government can abolish laws, and the Chief Executive can dismiss the LegCo, but the chance for the Chief Executive to be impeached or forced to resign is low. This is not hard to interpret: the Chief Executive is approved by the central government, and so are the administrations. It should never be forgotten that Hong Kong belongs to China, so the central government naturally needs an access to apply its influence on the Hong Kong democracy while leaving it with a high level of

democracy. The executive system led by the Chief Executive is the access, and becomes the leading sector of politics (Zhou Ping, *Journal of Yunnan University*, Vol.4, 2005, Pp.15). Again, this relationship illustrates the combination characteristic of the Hong Kong Model.

VII. The Two Different Democracies: British and Chinese

1. The British Democracy

The British Democracy is a very typical majoritarian model, and it was the prototype of the democracy in Hong Kong in the 1980s—it is not difficult to notice that elections in Hong Kong before handover all used the majority method. In this chapter, I shall briefly introduce the British democracy from three aspects: the electoral system, executive dominance, and constitutional flexibility.

Electoral System

One of the most significant traits of the majoritarian democracy is its disproportional election system. The House of Common is elected through a majority method. In a single-member district the candidate with the highest number of votes wins the seat. This system caused highly disproportional results and severe overrepresentation. In all elections between October 1974 and 2005, the winning party won clear majorities of seats with never more than 44% of the vote. The seeming majorities do not reflect the actual situation, and are more like “manufactured majorities” (Douglas W Rae, 1967, Pp.74) which are produced by the system itself rather than by constituents’ will.

Executive Dominance

There are two aspects that reflect the dominant position of the executive institution—the cabinet—in Britain. The first one is its power in the Parliament, and the second one is its centralization of executive power.

The cabinet in Britain is usually formed by the leaders of the majority party in the House of Common, so on most occasions it is relatively easy for its legislative proposals to be passed in the Parliament, and normally there is no worry about distrust of the cabinet. As long as the majority party remains cohesive and predominant in the House of Common, the chance of the cabinet being hampered by the legislature or falling out of office is small.

However, there is a chance that no party becomes the absolute majority in the Parliament. Though this situation rarely occurred in the last century, the cabinet from 2010 and 2015 was formed of the Conservative and Liberal Democrats, because the Conservative was only the relative majority. In this occasion, the cabinet is not as stable as the one-party cabinet, and depends on the level of cooperation between the two parties. Nevertheless, this coalition cabinet is not common, and does not much affect the fact that cabinets in the UK have a dominant advantage in the executive-legislative relations.

The other aspect of dominance of the British government is the centralized power. Theoretically the British government can be involved in any level of administration, and there is no constitutionally guaranteed autonomy for any local government (Lipjhart, 2012, Pp.16). In this aspect, Britain and China are similar. Both systems allow governments to have full power to control the administrative system nationally and give orders to local governments. Even though Scotland and Wales now have their own assemblies and

possess some degree of autonomy, England is still “the West’s most centralized system” (Ganesh, *Economy*, 2010).

Constitutional Flexibility

There is no such “constitution” in Britain that is written and specifies the properties of the regime, basic rights and basic responsibilities of citizens and governments, and other fundamental factors of a democracy. The constitution is instead defined as a series of basic laws such as the Bill of Rights of 1689 that regulates common law principles, customs and conventions (Lipjhart, 2012, Pp.18). The amendments to these laws require only the same majority as other laws do—over 50% of votes at the House of Common, instead of 2/3 for constitutions in many other democracies. This being said, the constitution can be amended like regular laws, so it is much more flexible than a written formal constitution.

Strength and Weaknesses

A very significant strength of the British majoritarian democracy is its relatively unhindered cabinet-parliament relation. In this model, it is more likely for the cabinet—the executive institution—and the House of Common to be congruous, which is beneficial to the efficiency of the executive. The single-party cabinet normally can easily get its legislative proposals adopted in the Parliament and keep the continuity and completeness of its administration roadmap from being frequently interrupted by the protracted debates and seesaw battles between parties.

However, the weakness of the British democracy is also obvious. The majority method with single-member districts largely undermines the credibility of the results of elections. The disproportional system more or

less distorts the true scenario of an election, and leads to over representation of the winning party in the parliament. Because the cabinet is usually formed by the majority party, the effect of the overrepresentation is actually magnified, from the legislature to executive institutions.

Another weakness is the unwritten constitution. The requirement for the informal constitution to be amended is as simple as normal law. Though this unique structure seems suitable to Britain, it would not fit some other plural regimes well, such as the one in Hong Kong. One of the bases of the existence of this setting is the existence of the two generally balanced major parties, the Conservative and the Labor. The similarity of the power of the two parties restricts the arbitrary for explaining or amending the unwritten constitution. Even if the dominant party tends to re-explain those basic laws, it may face great resistance from the other major party. However, the circumstance is totally different in Hong Kong. Hong Kong is much more plural than the UK, and there is no one or two parties in Hong Kong that can influence the political situation there. The consequence of the lack of a formal constitutional document is likely to be chaos. Every party would want to explain laws in its own way, and the legislature would lose its lighthouse. The unwritten constitution only fits Britain itself, with its historical and social background.

2. The Chinese Democracy and Political System

As commonly recognized, Chinese democracy is very different from the Western one, and it is not viable to totally utilize Western criteria to discuss the Chinese model. Thus, in this chapter, I shall introduce the Chinese democracy in a different way. The introduction mainly focuses on the electoral system, the legislative-executive relations, and the role of the CPC,

Communist Party of China. I also include analyses on the advantages and defects of this model.

Electoral System

The Chinese national legislative institution is the NPC, the National People's Congress. In the lower administrative level, there are four levels of People's Congress: Province, City, County (District), and Town. The tenure for all levels of People's Congress is five years. Each year, a plenary session is held to discuss the essential decisions of the Congress: budgets, final accounts, and amendments to some laws. During the adjournment of the Congress most of the year, the daily affairs are taken care of by the Standing Committee.

Among the five levels of People's Congress, the two lowest levels use direct election, and the three upper levels use representative election. For the direct election, the partition of districts is not strictly regulated nationally, but each representative in a district must represent a similar number of people. Each district can have 1 to 3 seats, and the number of candidates must exceed the number of seats for 33% of the seats but be less than 200% of the seats. Registered organizations and parties can nominate candidates. Nomination can also be made if signed by over 10 constituents⁴. In every direct election, if the turnout is lower than 50%, the election is invalidated. The number of candidates voters can vote for on their ballots depends on the total seats in the district. Ballots with equal or fewer candidates than the number of total

⁴ However, before 2003, independent candidates were extremely rare. In the 2003 election, the number was lower than 100 nationwide. Nevertheless, in the 2006-2007 election, the number increased to over 40,000 (John L. Thornton, *Foreign Affairs*, Vol. 87, 2008, Pp.2-22).

seats are valid. All names on one ballot are considered equally. Candidate(s) with over 50% of votes get elected⁵.

For the higher levels of People's Congress, the basic rules for election are similar, but the constituents are representatives in the level lower than the elected level, and candidates do not have to be lower level representatives.

The presidential election in China is constrained within the National People's Congress and is always a single-candidate election. Due to the single-party system, the nomination of the president is usually made by the Communist Party of China after negotiation with other democratic parties. In the last 22 years the General Secretary of the CPC has always been elected to be the President. The NPC also elects the Chairman of the Central Military Commission, President of the Supreme People's Court and the General Procurator of the Supreme People's Procuratorate.

The Legislative-Executive Relations

It is undoubted that the Chinese democracy is a majoritarian type. Article 1 of the Constitution of the People's Republic of China clearly states that "The People's Republic of China is a nation based on...with people's democratic dictatorship." This property determines that China has a central government with highly concentrated power.

The State Council, led by the Premier, is the highest executive office of China. The Premier is nominated by the President and voted on by the NPC.

⁵ There is a possibility that elected candidates are fewer than actual seats. In this scenario, a second election is held. Due to limited space here, this situation is not discussed. However, the Election Law of the People's Republic of China has specific regulation for this situation.

Other ministers and councilors are nominated by the Premier and voted on by the NPC. Generally speaking, the State Council has the authority to manage all aspects of the nation. There are mainly two things that it needs the authorization or permission from the NPC or the Standing Committee of the NPC for: i) budget and final account; ii) proposals to the NPC including amendments to laws. The State Council has the authority to make executive regulations and measures and to abolish executive regulations and orders made by lower level governments that it considers inappropriate. It is very rare that regulations from the State Council are abrogated by the NPC.

The Dominant Party

The party-executive relations are the most important part of the politics in China. Since the CPC is always the party in power, all crucial decisions are discussed within the party, and all national institutions are controlled by the party, including the NPC, the State Council, the Supreme People's Court, the Supreme People's Procuratorate and the CPPCC⁶.

The decision-making body and the core of the collective leadership of the CPC is the Standing Committee of the Political Bureau of the CPC. It is elected by the CPC Central Committee, and consists of the seven most powerful leaders of the nation, including the General Secretary, the Premier, the Chairman of the NPC, the Chairman of CPPCC, the Member of Secretariats of CPC Central Committee, the Secretary of the CPC Central

⁶ China People's Political Consultation Conference. It is an institution established to allow more citizens to participate in politics and express their opinions and suggestions to the government. It is also an important platform on which the CPC communicates and negotiates with other democratic parties and organizations and unites people from different regions, nations, occupations and classes to form the Patriotic United Front.

Commission for Discipline Inspection (CCDI) and the First Vice Premier. In most occasions, the committee is not involved in daily affairs. Its job focuses on discussing the national strategy and framework, such as political reform and Macro-economy formulating. Because the bureau includes the leader of every essential institution in China, most decisions are well developed and coordinated before being announced. The decisions will then be executed by the State Council and other related institutions.

In the local level of government (i.e. province and lower level), there are also branches of the Party that hold the power of final decision. Instead of the governors, mayors and chief executives in these levels, it is the Secretary of the Provincial (Municipal/District) Committee of the CPC that holds the final executive power, though this is not stated in the laws. Moreover, officers in every level of government, except a minority of members of democratic parties⁷, are all members of the CPC.

Advantages and Defects

The Chinese majoritarian democracy is definitely more concentrated than the British one, if they are to be compared. The main decisions are made by the CPC Central Committee. From a Western view this is severely against the principle of democracy, which is the division of power. Admittedly, the defects are disturbing. However, there are also some indisputable advantages with this model, and the result is significant.

⁷ This arrangement is to guarantee that parties other than the CPC have representatives in the government and hold part of the power, so there is a close relation between the CPC and these democratic parties. Nevertheless, officers who belong to these parties have no chance to be the chief executive in any level of government, and are normally in charge of some relatively “unimportant” sectors of the society like cultural affairs.

The most outstanding advantage of the Chinese concentrated model is its efficiency. In the last 37 years, the explosive growth of the Chinese economy has been considered a miracle, and the most important pillar that supported this kind of soaring was the collective leadership and centralization of power. The Party concentrates most resources nationwide and coordinates the distribution. Normally all local governments have to make executive decisions from the upper level with whatever methods they have, and opposition rarely matters. All needs can be satisfied within a short time, whether it is a 2000-kilometer railway constructed at an average altitude of 13,000 ft., or a stadium for an audience of 100,000 in the capital. The concentrated power reduces a vast amount of time for major national programs being debated and bargained between different interest groups, and maintains a harmonious and stable environment for the development to proceed.

However, the most deleterious defect in this model is corruption and the inefficiency that it causes. The centralized power allows leaders and officers in every level of government and party committees to find ways to embezzle easily, and a bribe is a useful (even essential years ago) way to make one's work progress. Besides, without an independent legal system and total separation between government and legislature, the supervision for corruption is very challenging. Since the Party is able to override justice institutions, it is always the CCDI and local CDI that first launch investigations against corruption with little transparency before the police and procuratorate can get involved. Even if corrupted officers are sued using the justice process, the procuratorates and courts are often interfered with by the party or governments. This kind of design causes many immense corruption chains in local, and even national governments, from a member

of the Standing Committee of the Political Bureau to a chief executive of a town.

VIII. Why the Hong Kong Model?

There are two important factors, which have been mentioned in Chapter II, that need to be considered when we compare different democracies in the scenario of Hong Kong: first its affiliation to the Chinese central government, and second the high plurality of its society, especially the collision between Chinese and Western values. The first factor is the foundation of the democracy. As a special administrative region rather than a nation, the autonomy of Hong Kong must not threaten China's sovereignty over Hong Kong, and there has to be an approach for the Chinese government to directly or indirectly apply influence in Hong Kong. It is not likely for Hong Kong to simultaneously be affiliated with China but also be totally independent from the Mainland, especially if we recall that the Chinese government is highly majoritarian and the power is greatly concentrated. Some general control over Hong Kong meets the national interests of China.

However, it should also be realized that Hong Kong is a highly diversified society. Traditional Chinese, new generations receiving Western education, and foreign immigrants make Hong Kong a pluralistic society with a varied culture. In this circumstance, it is important to be cautious of the potential conflict that may occur if a traditional democracy model, which may rely on particular background and circumstance, is directly transplanted.

The Chinese majoritarian model, as most people believe, is not a good choice for Hong Kong—actually, it should be the last choice when other

models all fail. First of all, the CPC does not have a stable foundation in the population of Hong Kong; it is unreasonable to declare that it represents the people's will as it does in the Mainland. Also, its dominant role in politics unavoidably determines that it has to apply some hard policies to suppress voices of opposition. These policies of the CPC would restrict the rights of citizens, such as the freedom of expression, and might cause a series of social and political conflicts, and even rebellion against authority. The concept of democracy and freedom, cultivated in the democratization started by McLehose in 1980s, has taken deep root in the public, and it is unrealistic to force the society to go back to the institution similar to (though better than) the majoritarian model before the 1980s, so there is no base for the pure majoritarian model to be built on. Secondly, the increase of executive regulation would undermine the core competitiveness of Hong Kong—the freedom. The application of the Chinese model has to be accompanied by the change from capitalism to socialism, which means that the country would intervene in the market and economy, and the laissez-faire model would be terminated. Thirdly, since the justice system is not independent within this system, the transplantation of the Chinese democracy would also bring Hong Kong more corruption, and undermine the incorruptibility that Hong Kong has achieved with the success of ICAC. All in all, applying Chinese Democracy in Hong Kong would destroy the achievement in democracy Hong Kong has obtained and lead to a retrogression; the high risk of chaos with all these sharp changes would not satisfy the target of the Chinese government either, which is the harmony and stability of the society and strong growth of the economy, so the Chinese model is not an ideal one for Hong Kong.

If not the Chinese style, how about the British one? Admittedly, the British democracy is the most democratic one among the four systems discussed here. Though it is majoritarian, it could make Hong Kong a totally autocratic region, and the LegCo and Chief Executive would have complete discretion and might also be more congruent. However, the application of this model also has crucial defects. The most significant problem is the executive dominance. The powerful cabinet and the party it comes from would make it difficult to find a way for the central government to influence the situation. If the dominant parties in the LegCo stood on the side opposite to the central government, the Central-Hong Kong relationship would be intense and thorny to deal with. The central government would also worry about the likelihood that one or more parties in Hong Kong were manipulated by foreign forces to become tools against China. The former Chinese leader Deng Xiaoping emphasized, “Patriots manage Hong Kong” and “The standard for patriots is respecting their nation, genuinely advocating the country to resume exercising the sovereignty to Hong Kong, and not damaging the prosperity and steadiness of Hong Kong” (Deng Xiaoping, 1993, 61). To ensure that the administration is patriotic, and to counteract the interference from the West⁸, the central government would not allow the whole democracy to be completely independent.

On the other hand, there could be another scenario. If none of the parties in Hong Kong could win the dominant position, like the situation today, then

8 In 1992 the US Congress adopted the *Hong Kong Policy Act*, which aimed to “maintain independent relationship with Hong Kong without Chinese government after 1997” (Li Changdao, 1999, 77-78). On June 26th, 2003, the US Congress adopted resolution No. 277, calling on the Chinese and Hong Kong government to withdraw the draft for legislature on Article 23 of the Basic Law.

the system would also face chaos. If the fragmentation I mentioned in Chapter V occurred, it might make the formation of a coalition cabinet extremely difficult, and even if a coalition cabinet were formed, it would be hard to exclude the probability that lots of time and resources would be wasted on the internal conflicts between parties. In this assumption, the executive and legislative efficiency would not be ideal.

In terms of the democracy before the handover of Hong Kong, it is also not favorable. The vital defect of this system is that although it seems that the three-level representative system was all open for direct election, it did not have the power to legislate. Instead, the three councils were only the consultation institutions for the governor, who was appointed by the UK government, and had no legislative power, so the public virtually had no means to affect any political decisions. If the Chief Executive was appointed by Chinese government and took complete power for the legislature, it would not be much better than directly transplanting the Chinese system to Hong Kong. Moreover, the model also applied a simple majority model in elections, which was far less representative than the present List PR system. The under- or over-representation caused by the majority method was a potential source of conflict to the plurality trait of the society. For these two reasons, the British colonial democracy of Hong Kong was not a complete democracy and would definitely not be favorable to Hong Kong.

From the discussion above, it can be summarized that a democracy suitable to Hong Kong should balance the two factors that I discussed at the beginning of this chapter and take the lack of those necessary conditions mentioned in Chapter II into consideration. Thus, it is natural to think about combining the Chinese and British democracy together to form a mixture of

both sides. The fusion is an intermediate system that does not lean toward either side—the central government gains influence in Hong Kong and reserves its authority to step into the situation when necessary, while a high level of autocracy is granted to Hong Kong most of the time.

The Central-Hong Kong relations are clearly formulated by the constitutional document, the Basic Law of Hong Kong SAR. Some of the relations have been discussed in chapter five, but other than those, a very crucial and fundamental point is that the NPC reserves the right for constitutional review of the Basic Law. This restriction strengthens the control of the central government over Hong Kong, and can be a very effective way to intervene when necessary. With this design, the central government controls all three important sectors of the democracy: judicial (constitutional) review, the appointment of Chief Executive, and the effectiveness of legislature (see chapter five), whereas compared to the other two solutions to affect Hong Kong politics, the constitutional review is more thorough and powerful.

However, this combined model is not designed to be directly manipulated by the central government. The policy “One country, two systems” and “Hong Kong people ruling Hong Kong with a high degree of autonomy” of the central government determines that the power reserved for central government is not designed for daily management, and if it was utilized too often or randomly, it would undermine central government’s credibility in Hong Kong and result in conflicts that the hybrid system initially tends to avoid, so a high level of interference from central government in Hong Kong issues is neither necessary nor beneficial to the society. To maintain the growth of the economy, steadiness of society, and positive international

influence, the central government sets the bottom line for the democracy of Hong Kong rather than forcing the region to fully accept its value. It is more like a final barrier to stop a trend that challenges the authority of the central government, which would lead the society into chaos and disharmony.

Despite those regulations, Hong Kong is still moving toward democracy. People have rights to shape their society through politics, and freedom of expression is generally well maintained. The reason for which the current model is being criticized often is that the electoral system does not allow citizens to directly elect all Legislative Councilors and the Chief Executive, and the nomination of Chief Executive is not open to the public. In summary, it is not “Western” enough, and the central government applies too great an influence on Hong Kong’s politics.

However, as I pointed out in Chapter II, the most suitable democracy for a region is not the one that totally fits the Western standard, but rather the one that fits the historical, cultural and social context. In such a plural society like Hong Kong with diverged values for democracy, some restrictions to prevent it from endless battles between parties are of necessity and a guiding force that is capable of taking a leadership in the development of democracy is demanded. As mentioned before, the List PR system has significantly lowered the standard to get elected, and made Legislative Councilors more representative, but it has also led to fragmentation. Democratization is a rapier. If direct election is completely utilized at this stage, like the British model, surely the regime would be more “democratic,” but it is hard to predict what the consequence would be. Simply transplanting the Western democratic concept into Hong Kong will not result in a real democracy, but would ignore the Chinese background of this problem, ruin the relationship

between the central government and Hong Kong, and put Hong Kong itself into a detrimental position. It should be emphasized that the aim for democratization is not simply meeting the Western standard, but more importantly the harmony, justice and steadiness of the society that allows the economy to grow while guaranteeing people's political rights. With this foundation, the people and government of Hong Kong and the central government can gradually seek a path for the evolution of the Hong Kong Model that will better take the advantages of the Western democracy.

IX. The Comparison between Hong Kong and Macau

The political systems in Hong Kong and Macau have some similarities. They both used to be colonies of Western countries and they share the same political position in China (both are SARs). The two systems are highly similar to each other, but the results of the practice are significantly different—the model is well supported in Macau, while it is severely disputed in Hong Kong. This chapter will briefly discuss the difference between these two regions, and the reasons account for these diversities.

Compared to Hong Kong, Macau is a more majoritarian system. Two noticeable differences are that the Chief Executive has the power to appoint 7 Legislative Councilors and to formulate and announce administrative regulations. The extra power granted to the Chief Executive of Macau strengthens the executive dominance of the system, causing the balance between the executive and legislature to lean toward one side.

The power of appointment of Legislative Councilors enhances the Chief Executive's ability to politically mobilize. With 7 out of 33 Councilors appointed by the Chief Executive, the chance and degree of cooperation

between the LegCo and executives in Macau are much higher than in Hong Kong, where all Councilors are generated from elections. This kind of institution increases the congruence between the executives and legislature, alleviating the difficulties for the Chief Executive in collaborating with the LegCo and propelling his political proposition (Ye Haibo, *Research Report on the Politics of Contemporary China*, 2015, Pp.278).

Moreover, the power to formulate administrative regulation in fact enables the Chief Executive to share legislative power with the Legislative Council. In reality, it is the Chief Executive who exercises the legislative power more often. In the decade after Macau was handed over to China, 312 administrative regulations were formulated by the Chief Executive, but only 133 laws were adopted by the Legislative Council (Ye Haibo, *Research Report on the Politics of Contemporary China*, 2015, Pp.279). Although the judicial power of administrative regulations is lower than that of laws, they have the same practical impact in reality, because regulations do not need to be based on specific authorizations from any normal laws and can be formulated only according to the Basic Law. With control over both executive power and legislative power, the Chief Executive's predominant role is very obvious.

The concentration of executive and legislative power to the Chief Executive makes democracy in Macau more majoritarian, but also more stable than Hong Kong. The antagonism from the Legislative Council towards the Chief Executive is more tempered. However, it should be pointed out that the success of this model has historical reasons as well. Before Macau's handover to China in 1999, the democracy was similar to the present one—the governor and the Legislative Council shared legislative power. Also,

unlike the plural society in Hong Kong, the society in Macau is relatively small (a population of 552,503, Aug. 12, 2011, released by DSEC) and unitary. Most of the population in Macau identifies themselves as Chinese⁹, and generally does not oppose the limited interference from the central government. With this being said, the conflict over democratic values in Macau is significantly weaker than that in Hong Kong, providing a stable basis for the system to lean toward its current model.

On the contrary, the power within the political system of Hong Kong is more diffused. The Chief Executive and Legislative Council balance each other, but the lack of a dominant side sometime leads to deadlocks and interminable debates. The Chief Executive does not possess the competence and resources to guide the public opinion and form his own political forces at the grass roots level, and the democratic parties cannot break away from the control from central government either. Hong Kong is a plural society, with rapid development of democracy in the last decades, and some radical communities identify themselves as Hong Kong people but not Chinese because of their abhorrence of the series of political movements in China from the 1960s to 1980s. Moreover, the Western ideology from Britain and the US also has affected Hong Kong during the Cold War and gradually produced an anti-socialism inclination and misunderstanding of Chinese socialism (Liu Zhaojia, *The 21st Century*, June 1997). All these factors together result in the continuous conflicts between some Hong Kong democratic communities, which claim that it is the Western democracy that

⁹ The result of a poll conducted by the Macao Polytechnic Institute in 2013 showed that “patriotism and love Macau” was ranked the top one among core values of the Macau society, and 81.82% of the respondents claimed that they were “proud of being Chinese citizens.”

is the real one (Lau Siu-kai and Kuan Hsin-chi, *The China Quarterly*, 2002, 1018-1019). If we consider what Hong Kong can learn from Macau, it is about first refining the system (imposing a threshold) and alleviating the fragmentation, and second guiding the public's attitude toward the central government to be more positive, and its understanding of democracy to be more rational, thoughtful and realistic.

X. Conclusion

With the comparison of the four democracies that could be proposed for Hong Kong, it is reasonable to draw the conclusion that the present combination is the most feasible and beneficial to the Hong Kong society.

The Hong Kong Model is a balance between the Chinese majoritarian model and British democracy. Unlike the Chinese one-party majoritarian model, the Hong Kong Model is not controlled by a single force and allows the division of power between executives and legislature. It also avoids the total independence that would be granted by the British democracy. Based on the democracy before Hong Kong's return to China, the current system satisfies the necessity for central government to ensure its full sovereignty and legal authority over Hong Kong while taking care of the democratization tendency and demand a more independent regime. The Hong Kong Model is based on the foundation of the region's history and culture, and makes a compromise between ideology and substance. With such a balance, it has successfully maintained the prosperity of Hong Kong, and generally kept the society steady. The other three democracies, especially the Chinese and British ones, though they have their own advantages more or less, are not as suitable to the properties of Hong Kong society as the current Hong Kong Model is.

Undeniably, the Hong Kong Model is not a perfect solution. The opposition between pro-democratic communities and the executives (and central government) has been intensified in recent years and the calls for total application of direct election have also been strengthening. There is still much to be refined and many problems to be solved. If Hong Kong is to promote its political model to a level at which the advantage of democracy becomes more significant, it should figure out solutions to mitigate the acute conflict among parties with different values and to guide the politicians and social activists to have a more substantive and flexible attitude towards the hybrid model.

The fragmentation in the LegCo, for example, may be alleviated by the introduction of threshold in the election (though the process of this modification will not be easy). The resulting realignment would induce small parties to find agreements which new larger parties are based on, and hence remit the influence of social and cultural plurality on the democracy. With fewer numbers of parties and larger sizes of parties, the probability of reconciliation may increase and thus the efficiency of legislature may be enhanced. Nevertheless, the formation of large parties may also lead to stronger opposition towards the central government, but if the substantive and flexible attitude is well developed, it is obvious for people to recognize Hong Kong's affiliation with China and the overwhelming advantages resulting from a good relationship with the Mainland.

With these solutions being explored and figured out, the Hong Kong Model will be more developed and produce more political, economic and social benefits. If the order of politics in Hong Kong evolved as discussed above, it would be reasonable to predict that the central government would not

attempt to seek more involvement in Hong Kong's local issues, and the political rights for people in Hong Kong would be effectively guaranteed. Moreover, considering Hong Kong's strong affiliation with the Mainland and their similarity on social and cultural foundations, this exploration of a more mature system may also provide some valuable references to the Chinese leadership when it is considering the future reform of the Chinese political system—an intermediate route for the regime to increase its degree of democracy while still retaining its tradition of centralization.

References

1. Census and Statistics Department of the Government of Hong Kong SAR. See <http://www.censtatd.gov.hk/hkstat/sub/so20.jsp>.
2. Census and Statistics Department of the Government of Hong Kong SAR, *2014 Gross Domestic Product*
3. Deng Xiaoping, *Anthology of Deng Xiaoping*, Vol.3, Beijing: Central Literature Press, 1993.
4. Diamond, Larry, *The Spirit of Democracy*, Beijing: Qunyan Press, 2013
5. Freedom House, *Freedom in the World 2015 Methodology*. See <https://freedomhouse.org/report/freedom-world-2015/methodology>.
6. Ganesh, Janan, "Over to You", *The Economist: The World in 2011* (December), Pp.108
7. Government of Hong Kong. See <http://www.elections.gov.hk/legco2012/sim/turnout.html>.

8. Lau Siu-kai and Kuan Hsin-chi, "Hong Kong's Stunted Political Party System", *The China Quarterly*, No.172, Dec., 2002, Pp. 1010-1028
9. Li Changdao, *Research on the Political Institution of Hong Kong*, Shanghai: Shanghai People's Press, 1997.
10. Lijphart, Arend, *Patterns of Democracy*, 2nd Ed., New Heaven: Yale University Press, 2012.
11. Liu Manrong, *Politics Institution in Britain-colonized Hong Kong and the Change of Hong Kong Society* (Guangzhou, China: People's Press of Guangdong, 2009)
12. Liu Zhaojia, "'Hong Konger' or 'Chinese': The Identity Recognition of Chinese in Hong Kong (1985~1995)", *The 21st Century*, Vol.41, June, 1997.
13. Rae, Douglas W., *The Political Consequences of Electoral Laws*, New Heaven: Yale University Press, 1967.
14. Xie Yu, "On the Disadvantages and Improvement of the Current Electoral System for the Legislative Council of Hong Kong", *Administrative Law Review*, Vol.16, 2013, Pp. 208-218.
15. Yang Yilong, Liu yanjun and Yin Cunyi, "The Democratic Political Utterance in Hong Kong: Orientation of Problems and Reshaping of Identification", *Studies on Hong Kong and Macau*, Vol.8, 2012, Pp. 106-112.
16. Yash Ghai, "The Past and the Future of Hong Kong's Constitution", *The China Quarterly*, No. 128 (Dec., 1991), pp. 794-813.

17. Ye Haibo, "Reflection on the Research to the Political Institution in Macau SAR", Research Report on the Politics of Contemporary China, 2015.
18. Zhou Jianhua, *Parties and Electoral Politics in Hong Kong (1997-2008)* (Guangzhou, China: Sun Yat-sen University Press, 2009)
19. Zhou Ping, "The Changes of Hong Kong's Political Environment in the Past Twenty Years", *Journal of Yunnan University*, Vol.4, 2005, Pp. 12-21.
20. Map of China from <http://a3.att.hudong.com/96/15/01200000000481113781565777996.jpg>.
21. Map of Hong Kong from http://www.maps-of-china.net/tourism_map/hongkong/officialmap.jpg.



The Difference between Oversampling and Undersampling: The Proving of the Nyquist Sampling Theorem

Yige Liu

Author background: Yige grew up in China and currently attends Beijing No. 4 High School, International Campus, in Beijing. Her Pioneer seminar program was titled "Optical Diffraction Patterns of Live Species."

Abstract

The Nyquist-Shannon sampling theorem is used frequently in reconstruction of signals. Reconstruction based on a set of oversampling samples can accomplish the reconstructed images with higher resolution and more complete information. Examples of both oversampling and undersampling can be created by using Matlab, which offers the chance to directly observe the difference between oversampling and undersampling. The Nyquist-Shannon sampling theorem can be proved using the algebraic method, and it can also be proved experimentally with Matlab. Furthermore, using Matlab, the influences on digitalized figures in both time domain and frequency domain due to sampling frequency can be quantitatively represented.

Keywords: Oversampling, Undersampling, Fourier transform, Discrete Fourier transform, Fourier analysis, Nyquist-Shannon sampling theorem

1. Introduction

Signals from both time and frequency domains are continuous; in order to compute the signals, we need to digitalize them by choosing discrete points to represent the original signal. The process of recording discrete points is called sampling. The number of sampled discrete points per second is called sampling frequency [1]. The Nyquist-Shannon sampling theorem shows that when sampling frequency (f_s) is greater than two times the greatest frequency of the original signal (f_{max}), the information of the original signal is completely maintained. When $f_s \geq 2f_{max}$, the sampling is an oversampling; when $f_s < 2f_{max}$, the sampling is an undersampling. [2] Obeying Nyquist-Shannon sampling can avoid Aliasing, which is a problem that occurs when $f_s < 2f_{max}$. Aliasing will lead to distortion and errors of the image. It will also make getting the correct sampling frequency domain and reconstructing the original figure impossible. Aliasing and undersampling can be investigated with the following fundamental knowledge:

In order to make the investigation of f_s and f_{max} easier, I choose to plot in frequency domain. To convert a signal from time to frequency domain, Fourier transform and Fourier analysis are employed.

When time and frequency are the conjugate variables, we can have

$$\Phi(v) = \int_{-\infty}^{\infty} f(t)e^{-2\pi i vt} dt \quad (1)$$

$$f(t) = \int_{-\infty}^{\infty} \Phi(v)e^{2\pi i vt} dv \quad (2)[3]$$

Function (1) shows the process of converting a signal from the time to the frequency domain. Function (2) shows the process of converting a signal from the frequency to the time domain; this function is also called inverse Fourier Transform.

Signals, as discussed previously, are usually continuous and infinite in both the time and the frequency domain. However, the properties make it so that the signals cannot be computed; as a result, an original signal should be manipulated through Discrete Fourier Transform to make it limited and discrete.

The function of DFT is

$$G\left(\frac{n}{NT}\right) = \sum_{k=0}^{N-1} g(kT)e^{-i2\pi n \frac{k}{N}} \quad (3)[3]$$

In this function, T is the time between samples, N is the total number of (equally spaced) samples, and k is a non-negative integer.

There are two distinct ways to prove the Nyquist-Shannon sampling theorem. The first one is using the algebraic method to show the existence and effect of aliasing. This method is based on DFT. The second method is a computational experiment. Matlab is required to simulate the original signals and show the figures when the values of f_{max} and f_s are changed. The differences between the qualities of images can show the differences between oversampling and undersampling.

2. Aliasing with the Algebraic method

As the DFT function (3) shows, a signal can be represented in the frequency domain as an approximation through discrete points.

If we add an integer to n in function (3), we can have:

$$G\left(\frac{n+N}{NT}\right) = \sum_{k=0}^{N-1} g(KT)e^{-i2\pi(n+N)\frac{k}{N}} \quad (4)$$

[3]

$$= \sum_{k=0}^{N-1} g(KT)e^{-i2\pi n\frac{k}{N}} e^{-i2\pi k}$$

According to Euler's Method:

$$(e^{-i2\pi})^k = \cos 2k\pi - i\sin 2k\pi = 1 \quad (5)$$

So that function (5) can be equivalently represented as

$$\sum_{k=0}^{N-1} g(KT)e^{-i2\pi n\frac{k}{N}} \quad (3)$$

As a result

$$G\left(\frac{n+N}{NT}\right) = G\left(\frac{n}{NT}\right)$$

Thus, the original signal in the frequency domain is replaced by a series of similar signals that are shifted by $2mf_s$ (where m is every positive and negative integer).

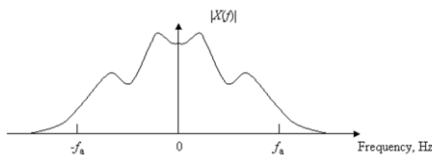


Figure 1: A signal in the frequency domain. [4]

In that example, the leftmost point on the frequency axis is maximum frequency of the original signal, denoted as f_{max} . In **Figure 1**, no aliasing occurs because an ideal low pass filter can eliminate the

shifted copies of the signal in the frequency domain—leaving just the original frequency domain for the continuous signal.

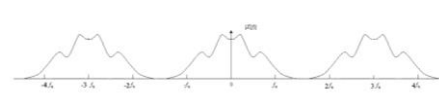


Figure 2: The sampling in the frequency domain when $f_s > 2f_{max}$.

In **Figure 2**, the sampling frequency of this diagram is denoted as f_s . In this case, $f_s > 2f_{max}$, and the signal in the frequency domain is completely maintained.

However, suppose that the sampling frequency is less than $2f_{max}$; the result, in an ideal case, can be shown as:

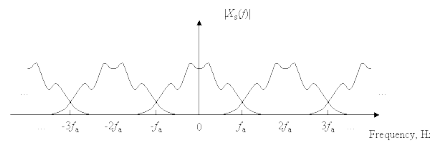


Figure 3: The sampling in the frequency domain when $f_a = f_s/2 < f_{max}$.

In **Figure 3**, it is easy to notice that the regions around integer times of f_a are overlapped. The overlapped area can lead to the result that the figure in the frequency domain is incompletely shown. [4] This phenomenon is called aliasing. Because of the effect of aliasing, the full frequency plot of **Figure 3** must sum the overlapping segments.

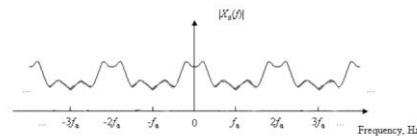


Figure 4: The full frequency plot of **Figure 3**

Thus, aliasing is harmful in signal sampling because a complete signal cannot be maintained due to the existence of aliasing. The negative impact of aliasing can be represented in the time domain:

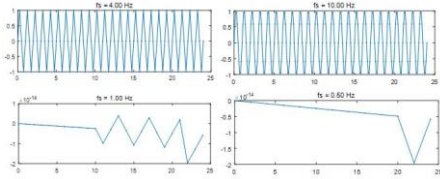


Figure 5: Different sampling rate of the same sine signal with $f_{max}=1\text{Hz}$ in time domain. Upper: images of sampling without aliasing show the complete information of the original signal. Lower: images of sampling with aliasing show the incomplete information of the original signal.

In the upper two graphs in **Figure 5**, we can see that once the sampling frequency passes a specific value, the information in the time domain (i.e. period) is completely maintained. In the lower two figures, the sampling frequencies are lower than $2f_{max}$. The periods of the lower two figures are larger than that of the original signal. It is important to acknowledge that: theoretically, when $f_s=1\text{Hz}$, the graph for the sine signal is a straight horizontal line. However, because of the extremely small scale (1×10^{-14}), computational error is obvious in that graph.

3. Undersampling and Oversampling with a Matlab experiment

In order to clearly observe the difference between oversampling and undersampling, I choose to operate the sampling process

with a one-dimensional signal. The method is: using Matlab to simulate a signal, the graphs in both the time domain and frequency domain will be shown simultaneously.

The program includes the following elements: the maximum frequency of the original signal f_{max} ; the period of sampling, which is denoted as dt , and the sampling frequency which is $\frac{1}{dt}$; the total sampling time which is T ; the sampling point (N); and the separation between the point dt with a range from zero to T . Verifying these elements can shift the sampling from oversampling to undersampling.

In this experiment, the original signal is a combination of three cosine functions:

$$y = 0.3 \cos(2\pi f_{max}t) + 0.6 \cos\left(2\pi \frac{f_{max}}{2}t\right) + \cos\left(2\pi \frac{f_{max}}{3}t\right) \quad (6)$$

In this case, I choose $f_{max}=1\text{Hz}$ in order to get a simple calculation.

When the maximum frequency of the original signal is determined, I verify the value of dt to change f_s .

In order to get higher resolution of the image created in the frequency domain, I choose a relatively large number T , which is also called sampling length. However, if the value of T is too large, the program will undergo a large calculation.

T is directly proportional to the sampling period and the sampling number:

$$T = Ndt = \frac{N}{f_s} \quad (7)$$

Recall the equation (1); FFT is used in this program to recreate the graph in the time domain in the frequency domain. Without `shiftfft`, both the frequency and the time

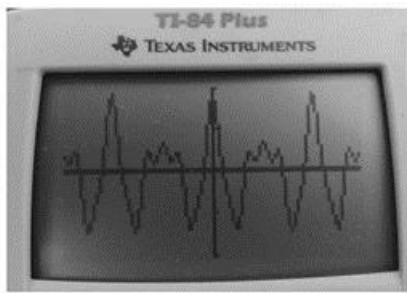
domain begin at zero; a remarkable property of FT—FT of each real signal is the fact that an even function cannot be observed:

For cosine signal $\cos(ax)$:

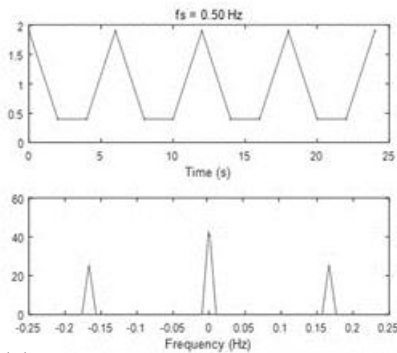
$$\frac{\delta\left(\xi - \frac{a}{2\pi}\right) + \delta\left(\xi + \frac{a}{2\pi}\right)}{2} \quad (8)[5]$$

As a result, shiftfft is added in the program.

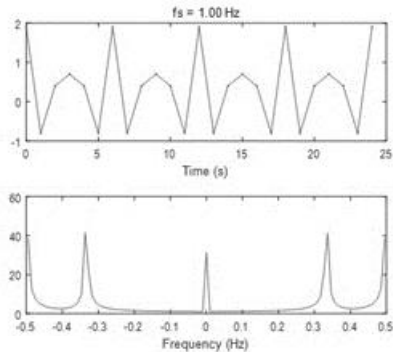
A graph calculator is used to generate a diagram of the original signal represented in function (6). Although the graph calculator also involves the process of sampling, it automatically chooses a suitable sampling frequency that shows the complete periods of the signal. Because it is independent from the Matlab sampling program, the graph generated by the calculator shows the shape of the original signal in the time domain. The diagram is functioned as a reference to compare the oversampling and undersampling.



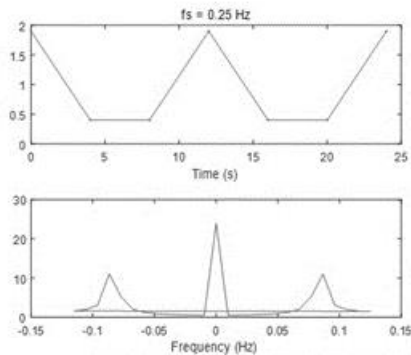
(a)



(c)



(b)



(d)

Figure 6: Undersampling examples. (a) Shape of original signal in the time domain created by graphing calculator, (b) undersampling example when $f_s = 1.00 \text{ Hz} = f_{\text{max}}$, (c) undersampling example when $f_s = 0.50 \text{ Hz} = \frac{f_{\text{max}}}{2}$, (d) undersampling example when $f_s = 0.25 \text{ Hz} = \frac{f_{\text{max}}}{4}$. The domain of frequency on each graph represents the portion of frequency that is sampled without distortion.

There are two groups of figures operated by the Matlab sampling program. The first group of pictures is the undersampling group which violates the Nyquist-Shannon sampling theorem. Regarding the maximum frequency of the original signal ($f_{\max}=1\text{Hz}$), the period of sampling in this group is $dt=1\text{s}, 2\text{s}, 4\text{s}$; as a result, sampling rate of this group is $f_s=1\text{Hz}, 0.5\text{Hz}, 0.25\text{Hz}$.

Figure 6(a) is a graph of two complete periods and two half periods of the original signal. **Figure 6(b)** is an undersampling pattern with $f_s=f_{\max}$; in the time domain, the shape of the two concave patterns in the original signal is lost. In the frequency domain, there are only three and two halves impulses shown in the graph, which means they are the only portions of the signal that are completely sampled.

Recall function (6). Theoretically, the sampled part in the frequency domain should be two groups of three impulses on both sides of the zero point, and the groups are symmetric to each other because of shiftfft. In **Figure 6(c)**, f_s is only half of f_{\max} . In the time domain, the period of the original signal is maintained, but all the concave patterns are lost; in the frequency domain, only three impulses are sampled; another three impulses are not maintained. Because of the extreme undersampling, there appears an impulse around the zero point, and this phenomenon occurs in all undersampling frequency domain figures in **Figure 6**. The shape of the signal in the time domain of **Figure 6(d)** doesn't maintain the period of the original signal; in contrast, it becomes another signal with a larger period, and the magnitude of the new period is two times of that of the

original signal. This phenomenon is the same as what was shown in **Figure 5**.

In **Figure 6(d)**, some noises appear in the frequency domain. These noises are not the result of undersampling, but are caused by the computational error because the range of the frequency domain is from -0.15Hz to 0.15Hz (the domain of the frequency patterns is -0.125Hz to 0.125Hz); the magnitude of the range is extremely small, so that the computational error is very obvious.

Again, **Figure 7(a)** is the reference figure of the original signal. **Figure 7(b)** is the diagram of the oversampling when $f_s=2f_{\max}=2\text{Hz}$, which is also called the critical value of sampling. It is the smallest value to maintain the complete information of the original signal in both the time and the frequency domain. The concave patterns are completely maintained; although the graph of the frequency domain shows there are only four and two halves impulses are sampled, the maximum frequency, which is shown as the rightmost impulse in the frequency domain, is included.

Sampling frequency of **Figure 7(c)** is four times f_{\max} , which is equal to 4Hz . All information of the original signal is maintained with high resolution. The impulses in the frequency domain are perfectly shown.

In **Figure 7(d)**, $f_s=8f_{\max}=8\text{Hz}$. The resolution of the figure in the time domain, also called the "reconstructed" figure, has an extraordinary resolution, and every detail of the original signal is maintained compared with the reference **Figure 7(a)**.

However, the sampling points are too condensed in the figure, and the high resolution of the figure requires more complex calculation, which will take extra time for the computer to create the image. Once f_s reaches $2f_{max}$, the information of the original signal can be fully maintained;

as a result, any sampling frequency higher than two times f_{max} is considered redundant in most applied areas.

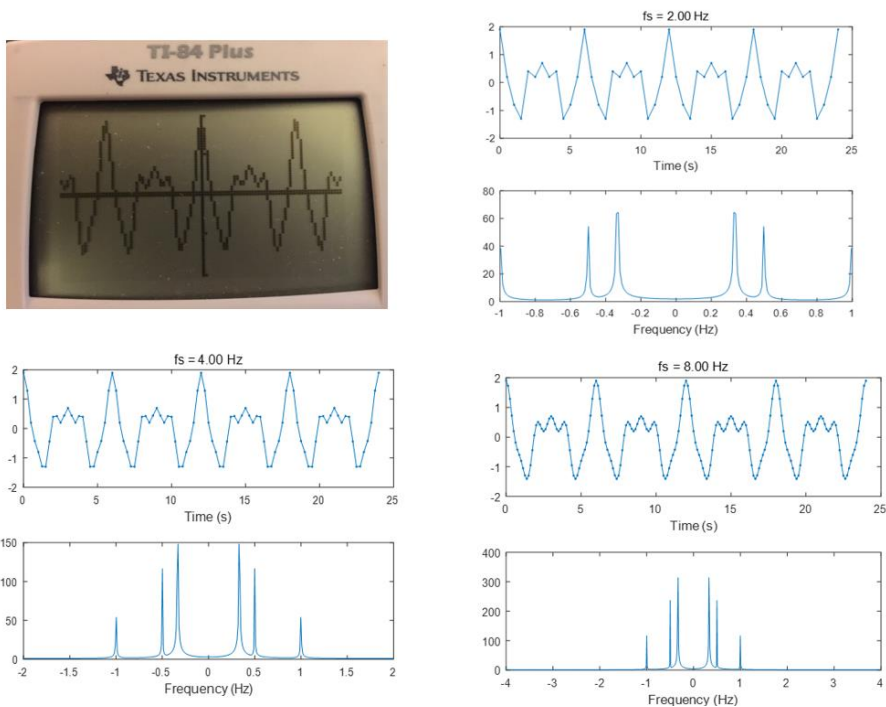


Figure 7: Oversampling examples. (a) Shape of original signal in the time domain created by a graphing calculator, (b) oversampling example when $f_s = 2.00\text{Hz} = 2f_{max}$, (c) oversampling example when $f_s = 4.00\text{Hz} = 4f_{max}$, (d) oversampling example when $f_s = 8.00\text{Hz} = 8f_{max}$.

5. Conclusion

According to the result of the algebraic proof and the Matlab experiment, the major difference between undersampling and oversampling is whether the information of the original signal is completely maintained.

In undersampling, half of the sampling rate is lower than the maximum frequency of the original signal. The result is aliasing, which is proved with DFT in section 2. The abnormal impulses in the frequency domain shown in the Matlab experiment are the result of aliasing. The reconstruction of the frequency patterns is the images in time domain with inverse Fourier Transform, and the reconstructive images lost some of the information about the original signal. As a result, undersampling is actually a synonym of insufficient sampling rate. Information in undersampling is not completely maintained in both domains.

Because the domain on the frequency axis is extremely small in undersampling experiments, computational errors occur frequently. Moreover, the computational error shown in **Figure 6(d)** is not isolated

In oversampling, $f_s > 2f_{max}$, which means there is no aliasing during extrapolation. As the Matlab experiment shows, oversampling maintains the complete information of the original signal. Although the resolution of the reconstructive image is directly proportional to the magnitude of f_s , in real cases, redundant sampling rate will extraordinarily increase the time and

complexity of computation. However, “In applied areas of sampling, making sampling frequency exactly equal to $2f_{max}$ is impossible, and the most common sampling rate is $5f_{max}$,” said Li Zhang, professor of Tsinghua University.

Therefore, the algebraic and the experimental methods together show that obeying the Nyquist-Shannon sampling theorem is the most efficient way to digitalize a continuous signal.

Moreover, recall **Figure 1** and **Figure 7**. For both signals, all of their frequency components are zero above a certain finite frequency for each. Under this condition, both of the signals are band-limited continuous signals. When the signals are sampled in oversampling, $f_s > 2f_{max}$, and it is possible to perfectly reconstruct the original continuous signals from sets of samples.

The capability of reconstruction from sampling in Nyquist rate (oversampling) is remarkable. All band-limited continuous signals in real life have the possibility of being reconstructed after computational manipulations.

Oversampling is applied in almost every field; it is the basis of putting signals into the computer. Oversampling is even used in fine arts—digital design is based on oversampling of colors and lines.

6. Acknowledgements

Thank you to Prof. Magnes of Vassar College as my advisor and instructor in accomplishing the experiment and the

research paper. She not only taught me Fourier Transform and basis of oversampling and undersampling, she also gave me many inspirations through discussions.

Moreover, thanks to Prof. Li Zhang of Tsinghua University. She gave me the idea of a practical focus for the research regarding the general topic of signal sampling.

Thanks to Dr. Dufan Wu of Tsinghua University for giving me the advice of creating two dimensional sampling patterns in Matlab.

REFERENCES

- [1] Communication in the presence of noise. 2. 37. Proc. Institute of Radio Engineers, 1998. Print.
- [2] Matsuda, Davin . "Nyquist Theorem." WhatIs.com. Margaret Rouse, Sep 2005. Web. Jul 2015.
<<http://whatis.techtarget.com/definition/Nyquist-Theorem>>.
- [3] J.F., James. A Student's Guide to Fourier Transform With Application in Physics and Engineering. 3. United Kingdom: Cambridge University Press, 2011. Print.
- [4] "Sampling and Aliasing." Montana.edu. Web. Jul 2015.
<http://www.coe.montana.edu/ee/rmaher/ases_cd/1.1.htm>.
- [5] Inspired by Dr. Dufan Wu of Tsinghua University.



Accelerated or Gifted Education——Which One is a Better Fit for Chinese Students?

Sicheng Luo

Author background: Sicheng grew up in China and currently attends Hangzhou Foreign Languages School in Hangzhou. Her Pioneer seminar program was titled “On Education: Fundamentals, Philosophies, Realities.”

Abstract

Accelerated education systems in the world have been a controversy for a long time. Recently, experimental reformers in Chinese education switched Chinese accelerated education to gifted education to cure some weaknesses of the original system. This paper introduces these two different educational systems in respect to China. Then this paper attempts to prove that gifted education in China has been limited by many practical factors, so this new system cannot solve the disadvantages brought on by accelerated education, hence gifted education for China is not a better fit for Chinese students. The gifted education system in China and those in other countries are compared. And through this analysis and comparison, a conclusion is drawn about the influence these two educational systems would have on Chinese students and why accelerated education is by far the better fit for China.

Section 1

Introduction

The development of many world societies, including economics and technology, is improving. So countries, especially those in rapid development, need a greater number of people with higher ability, like technicians or engineers, to support the improvement. Therefore, academic acceleration has emerged, as the times demand. “From the aspect of Sociology, accelerated education reduces the period needed to produce talented people, fulfills social need of ‘producing talented people quickly’” (JIAO Run-kai, ZHANG Rui, YANG Zhao-shan¹). Interestingly, academic acceleration prevails particularly in areas where intense competition exists. In developing countries with massive populations desperate for technical breakthroughs like China and India, competition pressure for teenagers is high, so the academic acceleration phenomenon is rather significant. In order to pick up talented people, and differentiate students with various abilities, examination systems tend to make tests harder. Thus the curriculum has to be more difficult, too, so that good learners can be distinguished. Schools are also a great catalyst: they are very likely to accelerate their syllabuses thus giving their students more of an advantage on tests. Therefore when these schools get good final test results, they get a supply of students with higher ability, so their performances will improve even further. This forms a virtuous cycle. Therefore schools and students both strive for accelerated curriculums, which brings many controversies. Due to the pressure brought by accelerated education, all kinds of criticism keeps coming from China and even all over the world about the disadvantages of the rigid Chinese educational system.

In this paper, the main objectives which will be discussed are 1) concepts of accelerated education and gifted education 2) various ways these systems are applied in countries 3) their influence on students 4) analysis of current

situations and how accelerated education shapes Chinese students and 5) the conclusion about the best fit.

Section 2-Accelerated education

2.1-Definition of accelerated education

“Acceleration” is defined as “process through an educational program at a rate faster, or at an age younger, than conventional” (Salkind, Neil J. P4.). This definition can be expressed in reality through difficult textbooks, extra knowledge required beyond textbooks, and precocious learning of more advanced knowledge than peers for students who take accelerated education.

2.2-Accelerated education in China

Besides China, some other countries do have this phenomenon as well. And accelerated education in China is rather special. It has some significant characteristics.

Firstly, accelerated education in China is universal. Most Chinese high schools have accelerated education (since primary schools and middle schools are banned from doing so), which is aimed at students who are more precocious. In Zhejiang Province, only a few foreign language schools do not apply accelerated education and the other more than four hundred high schools all do: a normality for students in China. “New students will be divided into regular, fast and slow classes by enrollment tests or admission tests held by schools” (Sometimes only fast and regular classes) (Jinhui Liu, 2012).

Secondly, the accelerated education system in China is complete and complicated. According to Dalian Daily, in AC 1114, an emperor in Song Dynasty named Song Hui Zong proclaimed a “San she” method for primary schools which is similar to today’s accelerated education. So although this method was abandoned briefly after implementation, the concept and even practical skills for accelerated education have a long history in China. Therefore this system has evolved to a very complete and complicated state now. For example, a high school in Zhejiang, Ningbo called “Zhenhai” High School, which is one of the best high schools in China, even has a hierarchical system within the accelerated education system. Before the high school entrance exam, this high school will first select about four hundred students to take an “Early Decision” exam and then select about 140 applicants to become its privileged students. Then these privileged ones will have the opportunity to get into fast classes by enrollment tests. And those students who are successfully accepted by fast classes will have the opportunity to take special tests to compete to be academic specialty students, for which there is a quota. Only seven of the fast class students can be chosen as academic specialty students. Then they have unparalleled privileges like particular courses for them taught by professors from Peking and Tsinghua universities, and exceptional opportunities for academic contests in the province, in China or even in the world. These seven students can do college-level exams at the beginning of high school so it is not hard to imagine how much they have been accelerated compared with others and how complicated the whole process is.

Thirdly, accelerated education in China is very outstanding from the aspect of examinations. Undoubtedly, this system is rather successful or it would not have thrived for so long. Take the same “Zhenhai” High School as

an example: in 2015 forty-one students were accepted by Peking and Tsinghua universities which are the best ones in China, and most of them were from fast classes, which means students from regular classes almost do not have any opportunity to compete with their peers in fast classes. This capacity gap of students from fast and regular classes is not so significant when they first enter high school. Fast classes are designated by the best teachers in the whole school for every subject and they have different syllabuses and specially designed courses. So it is not correct to blame the ability of students from regular classes since external factors like resources really have a great impact. So based on two significant advantages, a good supply of students and good teaching resources, the exam and admission results in fast classes are rather outstanding.

In addition, accelerated education in China is fixed. “Fixed accelerated education” means that the arrangement is unchangeable, i.e. once students get into fast or slow classes, it is very unlikely that they can get a change. Fast class students normally will not be transferred into regular classes and vice versa. There are two reasons for this phenomenon. Firstly, the regulation is not standardized so that schools do not have a clear system dealing with the change of members in fast and slow classes. Secondly, because of different beginning standards of students and disparate teaching intensity after the starting point, it becomes almost impossible to make a big change for regular students to fast students.

Finally, accelerated education in China is so important that this concept does not only apply to selected students, but to almost all schools. Every class has this characteristic that tends to increase the degree of difficulty of teaching content. In 2010, more than ten academics submitted a petition to

the Ministry of Education appealing to lower the degree of difficulty of textbooks in China. However, in May 7th, 2014, the Ministry of Education finally finished its three-year investigation and revealed the results: in ten countries they selected including Britain and America, the degree of difficulty of Chinese textbooks is ranked only in the middle of these countries.

	Primary school	Middle school	High school
Math	5 th	3 rd	3 rd
Science	7 th	6 th	
Physics		4 th	5 th
Chemistry		4 th	4 th
Biology		7 th	7 th

When revealed, this result created a stir. Apparently this ranking did not match people's expectations: it was too low to be believable. And there were many explanations for the results. For example most students, including me, found these results very plausible. Students in China all know that their textbooks are not very difficult indeed. And although syllabuses exist to standardize tests, by stating a strict range of knowledge that can appear on tests, Chinese teachers who are responsible for making similar exam papers can always find a way to keep inside the range but make questions very difficult and tricky. So, although textbooks are not very difficult, students actually have to learn knowledge far beyond the scope of them in order to get good grades on tests. Therefore, in this aspect, if one compares China to

the world, all Chinese students, whether talented or not, have to take accelerated education assigned by a rigid examination system. And over time how accelerated and how difficult the courses at a school are become a standard to evaluate and rank high schools in China.

Despite complaints, accelerated education cooperates with the Chinese status quo of needing efficiency and results: universality ensures the stability and the possibility to manage efficiently, complication makes it a complete system without flaws, outstanding results ensure the good supply of students for accelerated classes and the fixed style, which only requires one set of syllabi and textbooks, decreases teachers' workloads and ensures Chinese students' academic advancements.

Different from accelerated education, gifted education is a new concept to Chinese students. But due to inevitable disadvantages of accelerated education, the Chinese educational department, starting from 2014, began to try gifted education.

Section 3-Gifted education

3.1-How to define giftedness and thus how to select gifted students

“There is a common version for the definition of giftedness. Gifted individuals are those who demonstrate outstanding levels of aptitude (defined as an exceptional ability to reason and learn) or competence (documented performance or achievement in top 10% or rarer) in one or more domains” (Chi, N. 2008).

“Giftedness can be viewed as developmental in that in the beginning stages, potential is the key variable; in later stages, achievement is the measure of giftedness” (Subotnik R F, Olszewski-Kubilius P, Worrell F C. 2011).

3.2-Gifted education in China

So far, as I mentioned before, most Chinese high schools are using an accelerated education system. However, officials in the education ministry designated Zhejiang and Shanghai in China as two experimental cities for gifted education, which started in 2014. The main procedure of Chinese-style gifted education is: “Students choose courses for themselves according to their abilities in different subjects, recent academic developments and plans, also interest, by an internet platform from school’s website” (Chen Kai, 2014, page 1). And for courses that will be included in university enrollment tests, students will be divided into two or three groups marked as A, B and C classes. Although this seems a bit like the fast, regular and slow classes in accelerated education, they are actually different. A student in the reform of gifted education may be selected into Chemistry class C, which is aimed more at basic foundation knowledge, but may be in the A class of Chinese, which focuses much more on extension and additional supplements. So, in this gifted system, everyone gets a chance to be a top student in the subjects he/she is good at, different from fast and slow classes that only want students good at everything.

There are also some characteristics of gifted education, which are now still in the stage of experimenting.

Firstly, students in this system are autonomous. Although which level of classes students will be in depends on their test scores, they actually have the right to choose courses for themselves. University enrollment tests before the reform were rather rigid. Students chose to take humanities courses or science courses. Both of them included mathematics, Chinese and English. Only humanities also involved history, geography and politics, and science courses involved chemistry, physics, and biology. Now the right to choose is assigned to students. They can choose three courses among all choices without restriction. In this reform, students become the ones who make decisions, not education experts. Only in this situation can students learn with enthusiasm instead of robotic acceptance. Learning a course they like is so different from learning something they have to learn.

Secondly, flexibility is another symbol of gifted education. Different from the accelerated system, in the gifted system, as long as a student's academic performance reaches the standard of Class A, he or she can be transferred to the A class. And if a student in Class A did badly, s/he could also be downgraded into the B class. This rule triggered students' awareness of competition and determination to carry on and persist.

Thirdly, gifted education enables students to be taught according to their aptitudes. This theory by Confucius called "In Cai Shi Jiao" already existed since about 500BC in China. There are many similar theories in different cultures and historical contexts. John Locke in his "Some thoughts concerning Education" also mentioned more or less the same theory: different children should be taught differently and should be educated to scout for their own talents. To some degree, the Chinese version of gifted education reaches this goal. Students can choose by their interests and learn

by their aptitudes. They can spend more time on the subjects they really enjoy and get a deep understanding and thus good performance.

3.3-Gifted education in other countries and how it shapes their students

America:

Since every state in America has a different educational system, this diversity gives plenty of space for educational freedom of courses. It enables courses to be designed more specifically for students. And a core curriculum besides selective courses reduces the distance between different students and makes them all prepared for a basic academic foundation.

*** Finland**

“This country, with a successful educational system, does not have an assessment system and education related workers. For example, teachers do not depend on scores as the only index when evaluating students. Instead, they use a testing method designed by themselves to test the educational results” (Chen Jieqiong, 2014).

Different from Chinese gifted education in that to some extent it does give students more freedom on subjects they like, and actually still focuses much on test performance, which suits China’s current situation, Finland’s lack of focus on test scores, although another extreme, also suits the situations in the country. For Finland, a country whose population is only about 5,550,000, but has a higher real GDP per capita than the EU’s average, intense education and tense examinations for students are not a good way. What students in Finland need the most are good guidance, freedom, and

specific education towards each of them. Augustine in his “On the Teacher,” states the importance of guidance and the impossibility of teaching. And based on Finland’s strong economic foundation, this country really has the ability to put emphasis on the importance of guidance for each student since it does not have to put its students in a very intense environment nor chase efficiency at the expense of personality and specialty like China, which does not have other choices.

* **Germany**

“In 1923, Petersen, a professor of education at University of Jena founded a laboratory school.” One of his important principles is that schools should be integrated in the social environment but not isolated (Monks, F. J., & Katzko, M. W. (2005)). The successful application of this principle can be seen by Germany’s achievements in machinery and engineering which are quite useful and practical in daily life. Their leading role in today’s manufacturing industry can prove the compatibility of its educational policy.

3.4-Comparison

For the most part, the basic theory of gifted education in different countries is identical, but also obvious variety can be seen. Here are some aspects to explain why these differences exist.

Firstly, population matters. In countries with big populations like China or America, the efficiency of selecting students to categorize students according to their aptitudes is very important, or the whole system which tries to make students better educated will lose its advantage. So, there are similarities between the USA and China. For example, they both have well-

developed examination systems, stratified courses like the AP program in America and academic exceptional students in China. Only in countries like Switzerland and Finland with small populations and strong economic backgrounds, can implementing educational policies that require much more time spent on each student become possible.

Secondly, different opinions on study also make a difference. In “The Aims of Education” written by Alfred North Whitehead, he stressed the question: narrow and deep education, or wide but shallow one? The answers to this question by countries can be seen clearly in their various educational systems. For example, in Britain, their official textbooks of A-Level courses cover a great range of knowledge, which covers points from primary school to university in Chinese textbooks. But British textbooks do not have requirements for students to dig in so deeply. By contrast, in China, the scope of knowledge covered in high school is only a part of the British one, but the depth is so great that it goes far beyond that of other countries. And Chinese schools do not pay much attention to humanities subjects. Different from countries whose development has already reached a stage of stability and wealth, whose people started to concentrate on mental levels, China still desperately wants support of techniques, materials and capital. So the over-emphasis on science and mathematics can be explained.

These differences, which we could see in different countries, proved one thing: gifted education is only a concept. How to use that concept and how it can be applied by detailed methods depend on the countries and their different features. Gifted education in China is branded with Chinese features.

Section 4- Accelerated and gifted education

4.1-How does accelerated education shape Chinese students?

To answer this question, firstly two smaller questions should be answered. They are a) What qualities of students are erased by accelerated education and b) What qualities do Chinese students get from accelerated education. c) Will a transition from accelerated education to gifted education change the situation?

Accelerated education is always receiving criticism. Everything happens for a reason so there is surely something inside it that is harmful for students. Firstly, for qualities that are lost in Chinese style accelerated education, creativity is always the first to be mentioned. “Such exclusive reliance on standardized testing for educational assessment forces administrators and teachers to emphasize rote learning and memorization which ultimately limit creativity. Students are asked to repeat, memorize and remember textbook information for monthly exams, commonly referred to as Exam Hell both in Japan and Korea, leaving no room for creative growth” (Kim.K, 2007). However, many people agree that Chinese standardized testing systems, though they may prohibit creativity, actually at the same time set a correct and straight way for thinking and logic for students. This means that the orthodox knowledge we learn at school and the correct way we learn it prevents us from getting into an improper way of thinking and a disruption in logic. This straightness can help us obtain creativity in a more straightforward way.

There are surprisingly many great qualities that make Chinese students so capable and well-developed right now because of this “rigid, serious, obedient” system.

Firstly, the most obvious quality: persistence. Accelerated education requires so much effort that really tests students' persistence. Doing tedious work for a long time, though it may reduce their interest, provides excitement to those who like this subject, gives them endurance to keep on working and tells them that real life investigation and work are not relaxing at all. This adapts children to get used to the stress of life early, which is beneficial to their future. Children in accelerated education obtain persistence and thus patience easily, which are two of the most important qualities for work.

Secondly, students get a tough heart and the ability to endure stress; thus they are able to convert stress to motivation. Japan is famous for its high suicide rate, which is caused by high stress. And it is not possible to get rid of stress in life. The only way to tackle it is to convert it into motivation for life. At a young age, for students in an environment full of friends, support and joy, it is easiest to get used to stress and transfer that into encouragement and stimulation in a healthy way with the help of friends, teachers and the innocence unique to young students. "All people should experience a tough life for a few years" is a traditional idea of Chinese people. They think it is beneficial for the young to face the future. And in accelerated education, students experience hardship to obtain self-control, resistance and endurance, which are all very precious qualities necessary for success in the future. Moreover, this hard training at a young age only works for Chinese students. In Chinese culture as a heritage for all students, the determination to try their best and strive for their lives are in the doctrines. Only with these faiths, will Chinese students have the strength to go through the whole process and like a sword being made sharp by thousands of hard scrapes, finally become good people with priceless qualities.

Thirdly, the trepid, modest character is also obtained. “Thus, in Confucian cultures, people are not supposed to be narcissistic” (Martinsons & Martinsons, 1996). This is a criticism saying that being modest gets in the way of being creative since correct narcissism encourages creativity. Actually, modesty is according to Chinese traditional philosophy to have a trepid, and indifferent-to-fame life. Modesty is an attitude to become aware that we cannot become good enough to stop learning from books or from other people. Narcissism will make people easily influenced by external comments and feelings, which can cause nameless worries and distractions. As for creativity, even though narcissism can encourage it, modesty can keep students from being influenced by insignificant things and keep their original dreams intact, which can also be beneficial for the growth of creativity.

The government’s experimental transition from accelerated education to gifted education in China is trying a brand new concept for Chinese students. So will this transition make a difference to Chinese students’ characteristics in terms of creativity, obedience, patience and so on?

Imperial civil examination should be firstly discussed to better understand the reasons and effects. This is a very ancient traditional test begun in AC 605. The modern Chinese Gaokao System grew out of that Imperial civil examination though the latter one was abandoned in 1905. However, the main reason modern Chinese people did nothing big to this Gaokao system was to keep fairness. Population is always a difficulty. In a country with almost 1.4 billion people, fairness is always the priority. Chinese people are too familiar with exams. They can always find ways to tackle these down, no matter how complicated, or opportunists can always

find the deficiencies of the system to benefit them. Bribery and cheating would run wild if the system had any leaks. Although it is shameful to say it, this can be proved by all forms of leaking SAT test questions by Chinese people. Facing such “smart” test takers, governments have to be smarter. The unimaginably complex Gaokao system has lasted for so long because it is so far the best way to keep fairness, reduce cheating and any forms of violation of rules, and to test the real abilities of students. And these are surely the prerequisites for formation of creativity and so on. When governments realized the problem of creativity, according to their conservative style, they finally started to take action in 2014, which is the thing mentioned earlier which will be discussed further: gifted education.

4.3-Which educational system is a better fit for China?

So far many merits and demerits have been analyzed. Now the question may be solved. Accelerated education and gifted education exist in China side by side. So which one is a better fit? Many disadvantages of accelerated education are discussed above and it seems that gifted education so far is a very good solution to solve those problems. It is a huge contrast, but is gifted education really a better fit for China and Chinese students today? To answer this question, there are three things that must be figured out: 1) what aspects of gifted education are too idealistic in China? and 2) what aspects of accelerated education are suitable or realistic for China? 3) Overall, is accelerated education eventually the best fit?

Firstly, why do some think that gifted education ideas are too idealistic? There are some core ideas of gifted education that have been discussed above, but accelerated education has the same problem.

1) Small-size class and management. This helps students to get more chances discussing and sharing their ideas in class, which is proved to grow more creativity. However, if every subject is divided into three classes of different levels A, B and C, even though assuming some teachers can handle more than one class at a time, the number of teachers required must increase. It is not only about number, but also about the different needs. Students are stratified; therefore teachers also need to be stratified to match each level of students. The class materials all need to be adjusted. Workload for teachers, schools and governments increases significantly. For example, in Zhejiang, where gifted education is in every experimental school, it requires a long period of time for teachers and managers to figure out the whole detailed plan for gifted education and preparation. Therefore, if all areas in China start gifted education, and all students took it, the increase in burden financially, physically and mentally for China would be enormous. For this issue, in accelerated education, it never becomes a problem. Taking a big class and one teacher teaching in the front, and managing students in the unit of form classes and respective form tutors, and standardizing all tests and textbooks and syllabuses, this whole system is efficient and well developed, thus is not a big burden for China to afford.

2) Another core idea is the freedom to choose what students like to learn instead of being assigned classes. In America, like in the AP system, students have so many choices like Statistics AB and BC, art history and so on. These tests have equal status and students can get equal achievements in

taking any of them. However, in China, the subjects to choose from are only four or five, and among them, mathematics and physics are particularly popular, and as long as students are not terribly bad at them, they tend to choose these two courses. There are two major reasons. Firstly, because math and physics are the most difficult ones in tests with the highest total marks, it is easy to widen the gap between the good and the bad. So it is more advantageous for test taking. The second reason is related to the future: students who choose these two tests in final testing have better majors in college and thus better jobs in the future. Therefore, although students are given the freedom to choose, they do not really have the right to choose as they wish. So although it is called gifted education, it is actually a twisted and embarrassing version. Because of special conditions in China, gifted education cannot achieve the effect it is supposed to.

3) It seems that gifted education can release pressure for Chinese students since it reduces competition by stratifying students. However, in China today, competition out in society can only be higher than in schools. Students who do not have the ability to bear pressure and competition are more likely to be knocked out. Reality is cruel. A student is very likely to have a child and four parents to take care of when married. A happy and relaxing childhood may lead to depression in the future. This statement might be too extreme, but the practice bearing pressure in accelerated education can be very helpful.

Accelerated education has worked very well for a long time. It does not necessarily mean it is the best fit for China; it only means that so far no better and safer way has been found. According to previous sections comparing Chinese and other countries' gifted education systems, we have

found that every country adjusts its gifted education system to match its own features in population, politics and so on. And so does China. As mentioned before, due to massive population and the need for deep acquisition of professional knowledge, Chinese gifted education, which can be also called stratified education, is already adjusted according to Chinese current situations but still may have some incompatibilities.

Therefore it can be seen that although the concept of gifted education from western countries seems to be able to tackle some disadvantages of accelerated education, gifted education in China actually cannot overcome these difficulties, at least for now. There are two main reasons. Firstly, according to comparison between Chinese gifted education and western system, Chinese education is adjusted and cannot achieve the same effect as it is supposed to. For example, the limitation of choosing courses makes the freedom to pursue interests impossible. The second reason is that, even if gifted education could correct disadvantages of accelerated education like restrained creativity, gifted education in China would require a long time for experiments and investigation to achieve the effect it is expected to bring. Some critics in China even call gifted education not a reform at all but a new version of accelerated education.

Section 5- Conclusion: Chinese students under accelerated education have a role in society

The reason for any of these systems is the need for better-trained workers and innovators with high ability. Therefore, accelerated and gifted education prevails throughout the world. Every country is on the way to improving or finding new education systems to get optimal results. In the special case of China, this academic acceleration actually shapes Chinese people into

persistent, patient, and modest citizens for this competitive society and even makes them creative, which helps them gradually augment their voices in the world.

The world acts according to a principle: the fittest survive. Just like with the USA and Finland, where each country has its unique principles that are reflected in gifted education, Chinese accelerated education is suitable to China. It is good at being efficient, which is the quality paramount to children, and shapes them into good characters, which are popular and suitable to this world. There are disadvantages like the inevitable issue of creativity, or the lack of ability to make decisions because of unconditional obedience, but some of them can be cured by study of the unconstrained spirit in Chinese traditional literature without the change of educational system. But some of them can only be tackled by changing the educational system. Accelerated education's advantages still outweigh its disadvantages. The new idea of education called gifted education has been applied as experimental concepts, but has proven to lack compatibility with China today and still needs to undergo the test of time. In short, accelerated education is, so far, the best fit for Chinese students and it is accelerated education that prepares Chinese students for a strong role in the world.

Not arrogant or flippant, Chinese people who are shaped by culture and education live in their own way and act in accordance with their own roles, which are tranquil and peaceful, and that is very important and irreplaceable.

References:

- JIAO, R. K., ZHANG, R., & YANG, Z. S.** (2008). Development of Gifted Children: from Accelerated Education to Colorful Education [J]. *Journal of Northeast Normal University (Philosophy and Social Sciences)*, 6, 003.
- Salkind, N. J.** (Ed.). (2008). *Encyclopedia of educational psychology* (Vol. 1). Sage Publications.
- Chi, N.** (2008). National Association for Gifted Children. *Encyclopedia of Special Education*.
- Reynolds, C. R., & Fletcher-Janzen, E.** (2007). *Encyclopedia of special education* (Vol. 3). John Wiley & Sons.
- National Association for Gifted Children.** (2010). *Redefining Giftedness for a New Century: Shifting the Paradigm*. Retrieved from <https://www.nagc.org/sites/default/files/Position%20Statement/Redefining%20Giftedness%20for%20a%20New%20Century.pdf>
- Senior, J.** (2015). Book Review: *Assessing Intelligence in Children and Adolescents: A Practical Guide*. *Gifted Education International*, 287-287.
- Chen, K.** (2014). A brief discussion of Chinese stratified education. *A Brief Discussion of Chinese Stratified Education*.
- Chen, J.** (2014). Research on theories and practicals about gifted education in normal high schools.
- Kim, K.** (2007). Exploring the Interactions between Asian Culture (Confucianism) and Creativity. *The Journal of Creative Behavior*, 28-53.

CHAN, D.W., & CHAN, L. (1999). Implicit theories of creativity: Teachers' perception of student characteristics in Hong Kong. *Creativity Research Journal*, 12, 185-195.

MARTINSONS, M. G., & MARTINSONS, A. B. (1996). Conquering cultural constraints to cultivate Chinese management creativity and innovation. *Journal of Management Development*, 15(9), 18-35



Transformation of the Meanings of Arts and Crafts in China and Australia: A Comparison

Ziqi Ma

Author background: Ziqi grew up in China and currently attends The High School Affiliated to Renmin University of China in Beijing. Her Pioneer seminar program was titled “Communication and Culture.”

Abstract

The meanings of arts and crafts in China and Australia have gone through dramatic transformations as the societies modernized and entered into the global market. To better understand the transformation processes, this study identifies some common variables, such as translocal activities, state policies, activism of arts and crafts associations, and the modern market. Through an analysis of metasemiotic communicative practices that gave arts and crafts different sign values, this study compares how common socio-cultural factors led to different sign meanings in China and Australia. By looking at the creation and promotion of certain signs and the conflicts between competing formulations, this study describes the social processes through which Chinese arts and crafts’ sign meaning changed from “alterity” (of elite fine art and of modernity) to regional and national identity, and Australian Aboriginal art transformed from a representation of Aboriginality to an international high art.

1 Introduction

The meanings of arts and crafts in China and Australia have gone through dramatic transformations as the societies modernized and entered

into the global market. In China, traditional arts and crafts have always been viewed as the “alterity” (of elite fine art and of modernity), but as the Western and urban Chinese consumers started to link Chinese traditional arts and crafts to China and to certain regions, in addition to China’s (and the regions’) need to culturally differentiate itself in the global market, traditional arts and crafts became linked to the national and regional identity. In Australia, Aboriginal Art was at first seen as a representation of Aboriginality. But the “identity” sign transformed into the “high art” sign due to state policies and the activism of arts associations. As a result, Aboriginal arts were subsumed into the international high art market; an increasing number of art collectors purchased the artworks at high prices.

2 From “Alterity” to “Identity”—Overview of the Meanings of Arts and Crafts in China

Arts and crafts in China date back to the Neolithic era, when pottery started to develop. By the Ming and Qing dynasties (around 14th to early 20th century), Chinese arts and crafts had developed into a variety of art forms, including chinaware, embroidery, lacquerware, jadeware, and metalwork art (Gu, 2015). Arts and crafts had a combination of practical or ceremonial uses and aesthetic values. The carved lacquerware jewelry boxes, for example, were mostly used by upper class ladies. Others served as parts of the traditional festivities or customs in China, such as the “Five-Poison Gourd” papercut, which was made during the Dragon Boat Festival, to express wishes of good health in summertime when diseases were most likely to spread. In pre-modern China, although arts and crafts embodied great skill and delicacy, people viewed them as “craftwork” rather than real art. In fact, because the making of arts and crafts was conceived as menial

though skilled work done by “craftsmen,” as opposed to the production of wash paintings or Chinese calligraphy, which was the hobby of the upper-class intellectuals, arts and crafts were indexically related to the lower class. And while wash paintings and Chinese calligraphy were regarded as carrying the personalities of the artists, arts and crafts were regarded as carrying a high level of skill but little individuality of the makers. The delicacy of the arts and crafts was meant to embody the social status of the users, rather than the artistic values of the producers. “Craftsmen” for the royal family, for example, would create highly skilled arts and crafts, but the message embodied in the gorgeous works would be the royalty of the user, rather than the artistic ingenuity of the maker. In Goffman’s sense, the “craftsmen” were the animator of the message, but only partly the principal, because the message was also partly produced by the royal family for whom the art was produced (Goffman, 1981). Thus the complete meanings of arts and crafts were dependent upon the users who were mainly upper class elites. In this sense, arts and crafts became the alterity of Chinese fine art, as the lower class commoners were the alterity of the upper class elites. And in another sense, the “lower class” sign meaning of Chinese arts and crafts was dependent on the social hierarchy (and traditional festivities and customs in other cases) which provided the context of the message (Jakobson, 1960).

As China modernized through revolutions and wars in the early 20th century, the social hierarchy changed completely and traditional festivities and customs declined. As the cities gradually industrialized and modern manufactured goods replaced the hand-made arts and crafts in urban areas, traditional arts and crafts stayed mainly in the relatively underdeveloped rural areas. Without the traditional social context, the “lower class” meaning of arts and crafts was gradually submerged by the new sign meaning of

agricultural civilization as opposed to industrialized civilization. Just as the pre-modern rural areas were the alterity of the modern urban centers and the agricultural civilization was the alterity of the industrialized civilization, hand-made arts and crafts became the alterity of manufactured goods. The aesthetic value of arts and crafts was greatly overlooked when people appreciated modernity more than beauty. The old arts and crafts market in China declined and a new market wasn't born until the late 20th century.

During 1953-1956, the notion of “national identity” was first brought up by the state to encourage the development of the national handicraft industry. At the time, the arts and crafts industry was largely collectivized and the state held exhibitions such as the Chinese Folk Art Exhibition in 1956 to promote art production. After the Cultural Revolution, starting from the late 1970s, under the policy of Reform and Opening Up, a market economy developed in China and translocal activities increased (China Folklore Society, 2015). While domestic consumers were increasingly drawn to imported Western goods, the development of tourism and the souvenir industry gradually transformed the meaning of Chinese arts and crafts. Foreign tourists started to link arts and crafts to China, and Chinese urban tourists started to link certain types of arts and crafts to the regions where they were produced. Thus a new indexicality was created by the tourists. As interregional communication further increased with events like the 2008 Olympics in Beijing, this “outsider-created” indexicality linking arts and crafts to national identity or locality was also accepted by Chinese people as the state’s international publicity (aimed mainly at foreign tourists) adopted this indexicality. Chinese people also started to think of arts and crafts as representations of their national identity. Following this indexicality, the discourse on “cultural legacies” became popular. Governments of the

provinces adopted the indexicality linking arts and crafts and locality by publicizing tourism in these provinces in a similar way. Thus traditional arts and crafts in China gradually took on the sign of national identity. Today, while the modern market for arts and crafts in China isn't yet strong, some people have started to buy traditional arts and crafts rather than manufactured goods for their sign value of "national legacy," and an art collection market is emerging in China.

3 From the Aborigines' to the World's – Overview of the Meanings of Arts and Crafts in Australia

Before the arrival of Europeans in Australia, Aboriginal arts and crafts were the link between the Aboriginal people and the Dreaming, a religious concept in the Aboriginal culture representing another dimension of reality which "was there in the beginning, underlies the present and is a determinant of the future" (Morphy, 1998). The meanings of arts and crafts for the Aborigines were ceremonial or religious rather than aesthetic, and to the Aborigines they were not even "arts and crafts." As Europeans started to colonize Australia, white culture gradually dominated. To the whites, Aboriginal arts and crafts were viewed simply as "craftworks" representing Aboriginality. In the early 1970s, Aboriginal art was not considered real art and demand for it was little. In fact, it was not treated equally compared to Western art because Aboriginal people were a marginalized minority in the white-dominated society. Thus the Westerners created the indexicality linking Aboriginal arts and crafts to the Aboriginal identity. In this way, Aboriginal arts and crafts were constructed as the alterity of Western "real" art, signifying that the Aboriginal identity was the alterity of the Western identity.

From 1972 to 2000, however, Australian Aboriginal arts and crafts changed from unpopular folk handicrafts to million-dollar fine art. According to Myers, the first stage of the transformation was marked by state support, under the policy of “assimilation.” In this stage, aboriginal arts and crafts were generally viewed as enterprise. Art producers were spoken of as “craft workers,” receiving “training allowance” for their work and largely dependent on government support. At the time, Aboriginal arts and crafts were viewed as a way to get the Aboriginal minority to work and (at least partly) sustain themselves. However, under this policy Aboriginal art could not evolve into a self-sustaining industry and Aboriginal people were and would continue to be the marginal minority living on government stipends. Later, as the state policies regarding Aboriginal issues changed from “assimilation” to “self-determination” and social activists like Geoffrey Bardon and Peter Fannin started to promote Aboriginal arts and crafts, the sign meaning of Aboriginal arts gradually transformed. Aboriginal collectives such as Papunya Tula also spread arts and crafts to the mainstream (white) society. Artworks produced in one place were sold in another region and various exhibitions were held in white-dominated places like Melbourne and even outside Australia, in Europe or the United States. The aesthetic dimension was increasingly valued. And later, as associations such as the Aboriginal Arts Board and the Aboriginal Arts and Crafts Pty.ltd became the main buyers of Aboriginal arts and crafts, the Aboriginal arts and crafts market gradually matured. Aboriginal arts and crafts transformed into high-art commodities that gained popularity among Western art collectors (Myers, 2002). This indexicality, however, was largely constructed by outsiders. The original religious implications of Aboriginal arts and crafts were largely overlooked as Aboriginal arts and crafts were

subsumed into the international high art market. Thus the “Aboriginality” sign transformed into a sign of “art” in the eyes of Western art collectors, a sign not very different from that of the South East Asian or African arts in the global market. In a sense, Aboriginal arts and crafts were “globalized” by outsider Westerners.

However, a shortage of money was always a problem, and commercialization of Aboriginal arts and crafts brought about dilemmas of further development. One problem was the coordination of state stipends and the market. While the state fund in Aboriginal arts and crafts was limited, more investment in arts and crafts meant fewer stipends. Another problem was that as Aboriginal artists started to adopt a modern lifestyle, they were parting from the soil of their artistic creation, their traditional belief system and the traditional Aboriginal lifestyle. As the interpreters of the artworks became “outsiders” rather than “insiders,” some Aboriginal artists changed their artistic styles to cater to market needs. While this change of style could be seen as either an artistic compromise or a style development, it can be said that the original Aboriginality of the artworks has become harder to maintain.

4 Common Variables in the Transformation of Meanings of Arts and Crafts in China and Australia

In order to understand the social process in which the sign meanings of arts and crafts in China and Australia transformed, I identify some common variables such as translocal activities, state policies, activism of arts and crafts associations, and the modern market.

4.1 Translocal Activities Changed Meanings of Arts and Crafts in each Locale

In China, translocal activities, especially tourism, resulted in the construction of the “locality” indexicality, which later transformed into the “national identity” or “local identity” in the eyes of insiders. The translocal activities take two forms, foreign tourists visiting China and Chinese urban tourists visiting other parts of China. Arts and crafts as souvenirs transformed into a sign of locality to the outsiders and a representation of group (national or local) identity to the insiders.

According to statistics of the China National Tourism Association, in 2004 the total number of foreigners entering China grew to 16,930,000, nearly 1.5 times the number in 2003 and 74 times the number in 1978 (China National Tourism Administration, 2006). With the fast development of foreign tourism, the souvenir industry boomed and Chinese arts and crafts were officially categorized as an important part of the souvenir industry in the policies of the China National Tourism Administration in 2006 (China National Tourism Administration, 2006). The first major consumers of some Chinese arts and crafts were foreigners. According to Yang Yan, an artist who makes the traditional decorated bows and arrows, foreigners started to buy the bows and arrows long before Chinese consumers paid attention to his arts.¹ Although foreigners interpreted the bows and arrows differently from the artist’s intention, a market was created. For the foreign consumers the traditional Chinese bows and arrows represented Chinese culture more than archery gear. For the foreign tourists, the bows and arrows, or Chinese

¹ From an Interview with Yang Yan, an artist who makes traditional Chinese bows and arrows, Nov. 11th, 2013

archery, meant China, and China meant an exotic culture and a distant tourist destination. In Jakobson's sense, the sender and the receiver of the message did not share a common context (Jakobson, 1960). In the context of Chinese culture, archers in literature generally had a martial spirit and a sense of justice, and archery was a favorite sport of the Qing dynasty aristocrats. This context was not understood by the receivers. But at the same time, the receivers created a new context, in which the Chinese bows and arrows simply meant China. For some tourists even the "Chinese" sign mattered little. They bought Chinese arts and crafts as an "authentic" proof of their visit. At home, Westerners could easily purchase goods "made in China," but only in China could they buy the arts and crafts that Chinese people "made for themselves." For them, arts and crafts were just something "authentically Chinese" and easy to take back home. "Authenticity" mattered more than "Chinese," and "authentically Chinese" wasn't so different from "authentically African." The new context was a global one, which simply attributed the locality sign to things that came from each locale. The "misunderstanding," however, did not hinder the market and the sign meaning that came along with it. Thus while most urban Chinese people dismissed arts and crafts as a representation of the underdeveloped agricultural civilization and arts and crafts were the marginalized alterity of the mainstream art and culture, foreign tourists saw arts and crafts as a representation of Chinese culture.

With increasing international communication and the development of the souvenir industry, the "misunderstanding" had to be reconciled. The result was that the former Chinese context gradually gave way to the global context created by foreigners. And the change of context was evident in Chinese advertisements targeting foreigners, especially advertisements from

2007-2008, when foreign tourists came to China for the 2008 Beijing Olympics. In 2008, 53,050,000 foreign tourists entered China (National Bureau of Statistics of the People's Republic of China, 2011). In the official music video of the Beijing Olympics, "Welcome to Beijing," images of Chinese arts and crafts were used to construct the "Chinese" sign. Shadow play puppets, paper cuts, wood carvings, kites, masks for Peking Opera, traditional embroidery and lanterns all appeared in the video. And the arts and crafts were not shown simply as backgrounds. The characters in the video were shown making kites, embroidering, painting the Peking Opera masks and coloring the wood carvings (Beijing Olympic Committee, "Welcome to Beijing"). The implication was that not only arts and crafts themselves, but the making of arts and crafts, was officially accepted by the Chinese people as a sign of China. In the meantime, the video was frequently shown on TV in China. Thus Chinese people also started to accept arts and crafts (and the making of arts and crafts) as a sign of their unified identity. Several years ago the making of arts and crafts signified different meanings within the traditional social hierarchy, traditional festivities and customs. But around 2007-2008 the process of making arts and crafts, or the sending of the message, was linked to the sign of "China" as opposed to the original meanings such as festivity celebration or wishes for health and good luck from Chinese people to Chinese people. This transformation signified that even Chinese people were gradually replacing the original Chinese-created sign with the foreigner-constructed sign.

Translocal activities within China created a similar consequence. As urban travelers in China increasingly visited the different provinces of China, they linked local arts and crafts to the sign of locality, even if the local people never took arts and crafts as a representation of their culture and

themselves. The link between arts and crafts and the rural identity was not broken, but the sign value of the rural identity gradually changed from negative to positive. Handmade “tiger shoes,” for example, had been seen as a sign of underdeveloped agricultural civilization. Later, though they still signified the rural identity, the “new” rural identity was appreciated for its uniqueness. Gradually even insiders started to adopt this new sign in their publicity. Thus arts and crafts played similar roles in the tourism advertisements of the provinces and in the advertisements of China.

Similarly, translocal activities in Australia also resulted in an outsider-created sign meaning. According to Myers, the early notion of Aboriginal art, “art as enterprise,” was derived from the state policy of “assimilation,” or the policy of educating the Aboriginals in a Western way to push them into mainstream society. Aboriginal artists were workers who received government stipends for art production, and thus Aboriginal art produced this way was seen more as an enterprise than as art. The second notion, “art as cultural and spiritual renewal,” came from the perspective of art advisors, and was promoted by competitions, markets and exhibitions in different locales. The first art advisor, Geoffrey Bardon, valued Aboriginal art for its creativity and aesthetic value, and his successor Peter Fannin saw it as an alternative, “un-materialistic” perspective. According to Myers, “Competitions such as the Caltex Art Award and the Alice Prize helped to establish the relative quality of the paintings, as did recognition by national arts funding groups.” In addition, early exhibitions in the Museums and Art Galleries of the Northern Territory and the first sales of Aboriginal acrylic paintings in Melbourne both helped the transformation of Aboriginal arts into “art as cultural and spiritual renewal.” And later, state-sponsored institutes and the main buyers of Aboriginal art, the Aboriginal Arts Board

and Aboriginal Arts and Crafts Pty.ltd, pushed Aboriginal arts into the international high art market. In this market, Aboriginal arts were viewed as high art (Myers, 2002). The “context misunderstanding” was also present: the Aboriginal makers of the artworks sent the referential message related to the Dreaming, yet the receivers of the message, judges of the competitions, viewers of the exhibitions and buyers of the paintings, who were generally white people, interpreted Aboriginal arts within another context. The context came from the international high art market, in which Aboriginal arts were viewed as a high art commodity not very different from African high art or South Asian high art (Myers, 2002).

The Aborigines’ notion of Aboriginal arts changed as well. To reconcile the “context misunderstanding” and to make more money through Aboriginal arts, the Aboriginal artists also started to adopt the Westerner-created context. Furthermore, the artists even started to change the message of the artworks to cater to the interpretation of the institutions or viewers.

In both China and Australia, arts and crafts took on the sign meaning constructed by outsiders. Arts and crafts in both countries entered the global market and took on global identities, yet their global identities were different. In the case of China, arts and crafts were simply linked to where they came from, China or a specific province. Without the cultural context of social hierarchy, festivities and customs, the meanings embodied in Chinese arts and crafts were simplified to locality. In the case of Australia, the meaning of Aboriginal art was even more radically simplified. Without the context of the Aboriginal belief system, it was simplified into what it physically was in the international high art market—art. The difference lay in that while both sign meanings were constructed by outsiders, arts and

crafts in China were mainly promoted by insiders while Aboriginal art was mainly promoted by white art advisors, who were nevertheless outsiders. After arts and crafts in China became linked to China, Chinese people viewed them as representations of their national identity. Thus this locality was accepted and further promoted by the insider Chinese people. Although Chinese people had to give up their former complicated referential messages embodied in arts and crafts for the foreigner-created new indexicality, they built on this new indexicality to make arts and crafts a sign of nationalism. Thus this modified new sign satisfied the needs of both the global market and the Chinese people. In the case of Australia, however, Aboriginal artists depended on white art advisors to promote their art. The white advisors, however, generally valued Aboriginal art more for its aesthetic value than for its “identity sign.” Thus they naturally wouldn’t promote Aboriginal art as a sign of Aboriginal identity. A second reason was that Aboriginal people in Australia were a minority. Aboriginality itself meant an underappreciated, marginalized identity rather than a normal ethnic identity. Therefore although in the early 1970s Aboriginal arts were linked by outsiders to Aboriginality, that sign wasn’t merely an identity sign but a dismissal of Aboriginal art as “not real art” (Myers, 2002). As a result, the Aboriginality sign was inseparable from the ethnic prejudice that needed to be eradicated for Aboriginal art to be really appreciated.

4.2 State Policies

State policies in both China and Australia played a large part in changing the meaning of arts and crafts. In China, the arts and crafts industry was historically prosperous because there was great need for the practical or ceremonial use of arts and crafts. Buying and even making some arts and

crafts was part of people's daily lives. But as China underwent drastic sociocultural transformations to modernize in the early 20th century, traditional customs were greatly altered, which meant arts and crafts gradually lost some ceremonial uses. In the meantime, foreign manufactured goods became a replacement for traditional arts and crafts. In order to boost the domestic manufacturing industry, the government encouraged the handicraft industry as well. As Chairman Mao said in 1956, "The handicraft industry benefits the country . . . and especially arts and crafts, such as the cloisonné, and the Chang's grape sculpture . . . The handicraft products such as the Wangmazi and Zhangxiaquan scissors should not be replaced" (China Folklore Society, 2015).

The sign created for the purpose of developing the handicraft industry was "arts and crafts as national enterprise," because the making of arts and crafts was viewed as a form of domestic production that could boost the economy. Thus while some artists produced highly-skilled artworks for national exhibitions, more artists focused on producing what could be sold to the general public, and some artists got their inspiration from daily life. Dough sculpture, for example, used to portray subjects related to traditional customs and legends. Later, with the popularization of opera and literature, the subject matter transformed into opera or novel scenes (China Folklore Society, 2015). Then in the mid 20th century, the subject included fashion and modern people. Other forms of arts and crafts gained relatively wide market acceptance because of their practicality. For example, the Wangmazi handmade scissors were greatly valued because they were sharp. And people also proudly viewed the scissors as a Chinese handmade product that was as good as Western manufactured goods.

The first state-created meaning of Australian arts and crafts was “arts as ethnic enterprise,” a sign similar to that in China. Arts and crafts were viewed as mere enterprise when the context of art making was not taken into consideration and a new context was not yet invented. In China’s case, the “enterprise” sign emerged because the traditional context of art production, the traditional customs and lifestyle, was being replaced by a new, modern lifestyle. Thus people couldn’t receive the referential message of the artworks. And since modernity changed the perception of beauty, people couldn’t appreciate the arts and crafts as they used to do. Therefore, they could only think of arts and crafts in terms of processes pertaining to them as objects, especially processes of their production. In Australia, the context was lost not because of social change, but because outsiders were trying to interpret the message that was produced in the Aboriginal context. The aesthetic values were not appreciated because of the drastic difference between the Western and the Aboriginal perception of beauty. And the “national” sign of Chinese arts and crafts and the “ethnic” sign of Aboriginal arts were largely of state creation. The Chinese government promoted arts and crafts as “national” because they were confronting severe competition from Western products and needed to boost domestic production. By contrast, in Australia, because the government viewed Aboriginal arts as a way for the Aboriginal minority to sustain and develop, the arts policies were closely linked to the ethnic policies and the “ethnic” sign was thus created.

In the 1960s, however, with the start of the Cultural Revolution in China, the production of arts and crafts greatly declined because the state viewed them as signs of traditional ideologies. Thus the practical or aesthetic values of arts and crafts were largely subsumed by their historical role. Like

Aboriginal arts in Australia, arts and crafts in China also started to carry a political message.

Later, as China started the Reform and Opening up policies, economic development again became the focus. With the development of transregional transportation in the market economy, arts and crafts as souvenirs generated considerable profit. As arts and crafts were gradually replaced by manufactured products in people's daily lives, the handicraft industry declined. Translocal activities linked arts and crafts to locality, and the state publicized the notion. Internationally, the state took arts and crafts as a representation of China. Arts and crafts were viewed as a cultural legacy that embodied the long history of China. In an effort to promote a distinctive national identity in the global market, the state promoted arts and crafts as a shared heritage. Thus arts and crafts entered into the arena of diplomacy. Artworks such as carved lacquerware, cloisonné and embroidery often served as national gifts. Notably, in the 2014 Asia-Pacific Economic Cooperation Summit, the national gifts included cloisonné, Beijing embroidery, engraving and filigree (XinhuaNet, 2014).

In the meantime, the market economy policies also provided grounds for the "commodity" sign of arts and crafts. Today, although most arts and crafts enterprises are still fledgling, the state has encouraged such enterprises to support the developing cultural industry. While in China, the arts and crafts market is largely an insiders' market (tourism is not included here because in tourism arts and crafts were not viewed as a commodity), in Australia the market was largely an outsiders' market. Australian Aboriginal arts transformed into "commodities" when they entered into the international high art market. Under the state policies of self determination, Aboriginal art

production was encouraged as a way for the Aboriginal minority to economically sustain themselves. And art making was a work that they loved to do and chose to do. Since the need in the international high art market was relatively stable, the commercialization of Aboriginal arts became a sustainable solution.

4.3 Activism of Arts and Crafts Associations

Arts and crafts associations were important in the transformations of the meanings of arts and crafts in both countries. Although the associations might not have been the main creator of the message (especially in the case of China), they were key promoters and interpreters of certain meanings.

In China, some main organizations were the Arts and Crafts Association (which represents the China National Arts and Crafts (Group) Corporation in the World Craft Council), the China Association of Ethnic Folk Artists, and the China Folklore Society, to name a few. The aims of the China Arts and Crafts Association were:

To lead Chinese craft enterprises and craftspeople in the promotion of traditional handicrafts.

To provide suggestions to Chinese government to formulate policies for the betterment of the arts and crafts industry.

To introduce China's traditional arts and crafts to the international arena, and popularize China's cultural heritage in the world (China Arts and Crafts Association).

The China Arts and Crafts Association mentioned "Chinese craft enterprise" and "China's cultural heritage." The first concept corresponded to the "commodity" sign and the second was a mix of "national identity," the sign created by foreigners, and "cultural legacy," which was proposed by

UNESCO. Both sign meanings were welcomed by the state. The main activities of these associations include exhibitions, international affairs including conferences, seminars, craft demonstrations and competitions and media publicity. In Jakobson's sense, these activities could be seen as metadiscourse² regarding the meanings of arts and crafts. In holding themed exhibitions "to promote Chinese culture," the associations were interpreting arts and crafts as a sign of national identity, no matter what the artist originally tried to convey through the artworks. And in publicizing stories of artists, the associations became the agent that created the context in which the message was to be interpreted. (Jakobson, 1960)

While the "commodity" meaning was promoted by the associations, more heavily publicized was "cultural heritage," a notion proposed by UNESCO. In UNESCO's *Convention for the Safeguarding of the Intangible Cultural Heritage*, traditional craftsmanship was defined as a category of intangible cultural heritage (UNESCO, "Text of the Convention for the Safeguarding of the Intangible Cultural Heritage"). In the meantime, "craft and design" was another category under the culture division of UNESCO that encompassed arts and crafts (UNESCO, "Building confidence: Crafts for development"). "Craft and design," however, stressed a different cultural contribution of arts and crafts, namely creativity. In a sense, while the arts and crafts associations received both the messages of "cultural heritage" and "creativity," they chose to publicize the "cultural heritage" sign due to the state's need for a national identity. In a sense, then, the associations were active addressees of UNESCO's message which selected the sign that best suited their purposes to address that message to the wider public in China.

² The term "metalanguage" was introduced by Tarski (1933, 1935)

In the case of Australia, the associations, most notably the Pupunya Tula collective, the Aboriginal Arts Board and Aboriginal Arts and Crafts Pty. Ltd, played the role of “outside interpreters.” While Aboriginal artists still viewed their arts as related to the Dreaming, the state viewed their works as an integral part of the ethnicity issue. Reconciling the contextual difference was a necessary step in the development of Aboriginal arts. Thus, as the art associations tried to push Aboriginal arts into the international high art market, they had to overcome the “religion” sign created by the Aborigines and the “Aboriginality” sign created by the state. The aim of social activists such as Geoffrey Bardon was mainly to promote Aboriginal arts as “art” (Myers, 2002). They therefore promoted Aboriginal arts through exhibitions and art competitions that were formerly indexically linked to Western high art. As a result, Aboriginal arts became subsumed into the international high art market and took on the “high art” sign. Although the “high art” sign was not the associations’ original creation, by interpreting and promoting Aboriginal arts in the way Western high art was interpreted and promoted, the art associations promoted this sign meaning.

4.4 Art or Commodity? – Problems of the modern art market

The transformation of arts and crafts into commodities was difficult in both countries. In China, although the “legacy” discourse brought more attention to arts and crafts and generated a considerable number of customers, fewer people chose to become art makers. According to Yang Yan, an artist who makes traditional decorated bows and arrows, “Bow and arrow making generates only a modest income. In fact, learning the skill is very hard and only the bow-making skill takes 10 years to master. The apprentice has to spend a whole year to practice wood-whittling, which is

the first step of arrow making. How many young people today have such patience?”³ Traditionally, craftsmanship was passed on in families or through long-time apprenticeship. Either way, young people (even young children in the past) spent hours a day honing the skill under the rigorous requirements of parents or teachers. Through family inheritance, children inherently bore the responsibility to take craftwork as their occupation and pass it on to their children. Apprenticeship worked with similar rigor and responsibility. As more young people choose to exit the industry, few can take their place due to the strong barrier to entrance. Although arts and crafts have become a normal “commodity” in the modern market and many artists are starting their own companies, the making of arts and crafts hasn’t yet become a modern occupation in the eye of the public.

In addition, the commodity identity led people to question the artistic nature of arts and crafts. Peng Xiaoping, a dough sculptor, opened an 8000-square-meter art workshop in Beijing. Over 400 people worked in his “dough sculpture assembly line.”⁴ While it was certainly a huge step in the development of the arts and crafts industry, the “hand-manufacturing” of art led some people to question whether there was still artistic ingenuity in the craft. Despite such doubts, mainstream media still extolled Peng’s practice as “bringing new vitality into the old craftwork” (Zhao, 2012). Although some people viewed arts and crafts as art, the mainstream media through metadiscourse promoted the “commodity” sign. Over time more people

³ From an Interview with Yang Yan, an artist who makes traditional Chinese bows and arrows, Nov. 11th, 2013

⁴ From an Interview with Peng Xiaoping, an artist who makes Chinese dough sculptures, Nov. 22nd, 2013

would adopt the “commodity” sign instead of the “art” sign and the arts and crafts market would be able to develop.

Since commodities had to change with the market need, whereas the cultural heritage was to be kept unchanged, the “commodity” and “heritage” signs were sometimes in conflict. The handmade grape sculpture, famous for its verisimilitude to real grapes, took over 3 months to make and required a high level of skill. In order to sustain themselves, the artists had to ask for a very high price. Now machine-made plastic grapes have become a far cheaper substitute and controversy about whether this art form should still be supported arose.⁵ Behind this controversy was a sign conflict, which reflected the conflict between the state policies which promoted arts and crafts as a presentation of national identity and the artists’ need to build a truly sustainable art market. In some cases, the two signs could be combined. Jitufang, a company that produced Tuerye, a rabbit sculpture that embodied good wishes, renovated the image of the sculpture and combined it with the Chinese zodiac sign to personalize the product. The company successfully used the “heritage” sign as a unique feature of its product but also catered to the needs of modern consumers.

The “art-commodity” conflict was also evident in Australia. One of the main problems was the commercialization of art. After social activism promoted the “art” sign of Aboriginal arts and internationalized them, Aboriginal arts were automatically linked to commodity in the indexical framework of art market. As the “commodity” sign was increasingly promoted, even producers of art sometimes compromised their artistic

⁵ From an Interview with Gao Wei, Secretary-General of Beijing Folklore Society

integrity to cater to institutions or viewers. As the art advisor Bardon said, “The monstrousness of it was not lost on me as they began to chant in their own languages amongst themselves, then at me: ‘Money, money, money.’” (Myers, 2002).

The modern market influenced Aboriginal art-making indirectly as well. While the art market provides incentive to the artists, stronger participation in the modern economy may alter their traditional lifestyle and even “spoil” the Aboriginal context for art production. Similar to the situation in China, while many Aboriginals may shift away from the traditional context and even from an art career, those who stay may also be far from the traditional notion of “Aboriginal artists.”

5 Conclusions: Globalized “Locality”

In conclusion, although different sign meanings of arts and crafts were created and promoted in China and Australia, some common factors such as translocal activities, state policies, association activism and the modern market played important roles in both cases. While arts and crafts in both countries had for a long time been the marginalized “alterity” of mainstream society, they later became distinctive signs accepted and even valued by society. In China, translocal tourism linked arts and crafts to locality, and the government further extended the meaning of locality to the national (local) identity that was needed in the increasingly internationalized (interregionalized) market. The state and arts and crafts associations adopted the UNESCO notion of “intangible cultural heritage” to further publicize the national (local) identity sign. And because the sign discourse brought attention to arts and crafts, an insider market was emerging and the commodity sign was thus born.

In the case of Australia, the “art” sign was born when art advisors and arts and crafts associations tried to rid arts and crafts of the “Aboriginality” sign that came along with the white prejudice that Aboriginal arts were not art. The state policy of self-determination and the translocal activities of the associations such as exhibitions and art competitions further promoted the “art” sign. And as Aboriginal arts were subsumed into the international market where art was indexically linked to commodity, the “commodity” sign of Aboriginal arts was created.

In both cases, the commodity sign was sometimes in conflict with the artistic nature of arts and crafts (in China, also with the “heritage” sign), which led to the commercialization problem. In addition, a comparison of the transformation of the meanings of arts and crafts in both countries helps us to understand the relationship between “globalization” and “localization.” While the Aboriginal arts became a global commodity, Chinese arts and crafts transformed into a representation of local identity. Localization occurred when people linked arts and crafts to where they were produced instead of adopting its original message. Even though locality seemed to be enhanced, in fact the local context was being replaced by a global one. In a sense, the “locality” sign was a globalized product. Thus the outside forces of globalization weren’t so different in China and Australia. The main difference lay in whether there was an insider discourse that reinterpreted the outsider-created sign meaning.

References

- China Arts and Crafts Association. (n.d.). Introduction. Retrieved September 1, 2015, from <http://www.cnaca.org/english/about/index.html>
- China Folklore Society. (2015). Thoughts on the Contemporality of the Inheritance of Folk Arts and Crafts. In *The Protection and Inheritance of Traditional Arts and Crafts*. Beijing: Academy Press.
- China National Tourism Administration. (2006, June 23). Tourism Industry: An Industry Worth Our Commitment. Retrieved August 26, 2015, from http://www.cnta.gov.cn/xxfb/wxzl/201506/t20150625_430523.shtml
- Goffman, E. (1981). Footing. In *Forms of talk*. Philadelphia, PA: University of Pennsylvania Press.
- China National Tourism Administration. (2006, June 30). Implementing the Perspective of Scientific Development in the Comprehensive, Coordinated and Sustainable Development of the Tourism Industry. Retrieved August 26, 2015, from http://www.cnta.gov.cn/xxfb/wxzl/201506/t20150625_430516.shtml
- Gu, D., & Zhan, W. (2015). *History of Chinese Arts and Crafts* (2nd ed.). Nanjing: Nanjing University Press.
- Jakobson, R. (1960). *Closing statement: Linguistics and poetics*. Cambridge, MA: MIT Press.
- Morphy, H. (1998). Foundations: Art, Religion, and the Dreaming. In *Aboriginal art*. London: Phaidon Press.

Myers, F. (2002). Making a Market: Cultural Policy and Modernity in the Outback. In *Painting culture: The making of an aboriginal high art*. Durham, N.C.: Duke University Press.

National Bureau of Statistics of the People's Republic of China. (2011, March 24). The 17th Report on the Results of the Eleventh Five-Year Plan. Retrieved August 26, 2015, from http://www.stats.gov.cn/zjtj/ztfx/sywcj/201103/t20110324_71329.html

UNESCO. (2003, November 3). Intangible Cultural Heritage. Retrieved September 1, 2015, from <http://www.unesco.org/culture/ich/index.php?lg=en&pg=00006>

UNESCO. (n.d.). Building confidence: Crafts for development. Retrieved September 1, 2015, from http://portal.unesco.org/culture/en/ev.php-URL_ID=35418&URL_DO=DO_TOPIC&URL_SECTION=201.html

Welcome to Beijing [Motion picture]. (2008). China.

XinhuaNet. (2014, November 15). APEC National Gifts' Public Appearance. Retrieved September 1, 2015, from http://news.xinhuanet.com/culture/2014-11/15/c_127213236.htm

Zhao, S. (2012, December 6). Peng: Bringing New Vitality into the Old Craftwork. Retrieved August 31, 2015, from <http://www.bjbusiness.com.cn/news/2012-12/0600000015372.shtml>

Acknowledgements

I am greatly indebted to Cheng Zhuo, Bao Yuxuan, Huang Zongbei and Song Yuyang for their contributions in data collection. This work would not be possible without their diligent efforts. I would also like to thank the artists and art promoters, Mr. Yang Yan, Mr. Peng Xiaoping, and Prof. Gao Wei, without whom I couldn't have known the many problems and dilemmas faced by the Chinese arts and crafts industry. I would like to express my gratitude to Professor Asif Agha for all the valuable instructions. Finally I would like to thank the Pioneer Program for making this paper possible.



**Pesticides and Endocrine Disruptors:
Understanding the Molecular Mechanism by which Some
Pesticides (Aldrin, DDT, Mecoprop) Act as Endocrine
Disruptors**

Jonathan Mak

Author background: Jonathan grew up in Hong Kong and currently attends Chinese International School in Hong Kong. His Pioneer seminar program was titled "Crystals, Molecular Structure and Medicinal Chemistry."

Abstract

Over the past century, pesticides have seen increasing usage worldwide to eliminate weeds, insects, and other pests. However, many pesticides have unintentional health effects, and some of these effects are still not clearly understood. With the global presence of pesticides, it is ever so important to understand how pesticides affect the human body. In particular, some pesticides are known as endocrine disruptors and interfere with the body's endocrine system. The endocrine system controls the secretion of hormones in the body, which is vital to regular development. This paper has studied the molecular structure of three pesticides (Aldrin, DDT, and Mecoprop) to investigate the relationship between molecular structure and the mechanism by which pesticides act as endocrine disruptors, ultimately to shed light on why pesticides adversely affect our endocrine system. Through a case study of estradiol and the estrogen receptor, it has been suggested that the structural similarity of pesticides to hormones in the body is an important factor that allows some pesticides to bind to the hormone receptors and disrupt regular hormonal levels. This finding has implications for future

pharmaceutical development, providing evidence that consideration for molecular structure in molecular design has large impact with respect to a chemical's interaction with the human body and the harm it could cause.

1. Introduction

Pesticides have been widely used in areas such as agriculture to protect crops and eliminate pests. With the advent of modern science, a great variety of pesticides are available to countries all over the world. The global presence of pesticides is a pharmaceutical marvel and brings about economic benefits to many. However, people and nations unfortunately neglect the environmental and health effects of spraying chemicals across the planet. Shockingly, many pesticides were used indiscriminately, without thorough understanding of their possible consequences. As a result, they had to be banned later after sowing irreversible damage. Still today, there are gaps in knowledge about the harm certain commonly used pesticides may do. Fortunately, by studying the molecular structure of these pesticides, much can be discovered about their properties and explained about their interaction with the environment and the human body. A technique known as X-ray crystallography is used to examine molecules on the atomic level.

1.1 Background Information on X-Ray Diffraction and Crystallography

Light is a transverse wave travelling at $3.00 \times 10^8 \text{ms}^{-1}$ (c, the speed of light constant), and like all waves, it is defined by its frequency (f) and wavelength (λ) whereby $f\lambda=c$. All light of varying frequencies and wavelengths falls in the electromagnetic spectrum. X-rays are a form of

electromagnetic radiation with very short wavelength, on the order of 10^{-10} meters (or 1 Å). The wavelength of X-Rays is significant as it is comparable to the bond length between atoms in the molecules that make up crystals, meaning that their molecular structure can be derived using X-rays by analyzing the crystal's X-ray diffraction pattern. Diffraction occurs when light passes the edges of a body or through a slit or when it is reflected from a surface. After being diffracted, the light rays appear to be deflected to produce fringes of light and dark bands, which correspond to the interference of X-ray waves as they interact with the electrons around the atoms. X-rays hitting a crystal would create a diffraction pattern which would depend on the geometry of the crystal. The diffraction pattern could then be used to deduce the structure of a molecule. This procedure is known as X-ray crystallography. (Nelson, 2014)

Crystals exist everywhere in our daily life. Sodium chloride, known by its common name "table salt," is one example of a crystal. Crystals are solid substances which have a definite regular shape with distinct arranged plane faces. The internal structure of crystals can be described in terms of the arrangement of the atoms or molecules of the unit cell and crystal lattice that form the crystal. Molecules or atoms of a substance fit together in a uniform and regular manner to form a unit cell, which can be thought of as a building block of a crystal. The unit cells repeat to create the crystal lattice, the structure of the crystal. Each crystalline substance has a unique unit cell that is defined by the dimensions of its sides and the angle between the sides. Yet these measurements are on the order of 10^{-10} meters (or 1 Å), too small to be seen by the naked eye. To make such measurements and determine the

structure of a molecule, X-ray radiation is used because X-rays have wavelengths that are on the order of 10^{-10} meters (or 1Å).

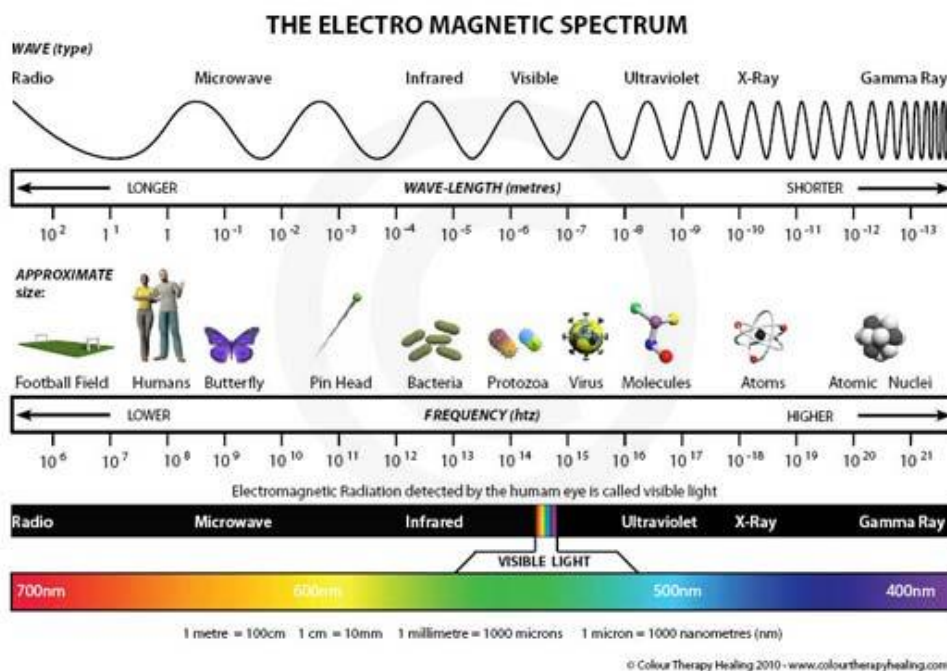


Fig 1- Diagram of the electromagnetic spectrum, with X-Rays to the right end on the scale of atoms, having shorter wavelengths and higher frequencies. (Kukull, 2015)

Crystallography is concerned with the structure and properties of crystals, and X-ray diffraction is a powerful tool used in this study. To carry out the X-ray diffraction experiment, one would certainly require the crystallized form of a molecule, a source of X-rays, and a diffractometer. When the X-rays are focused on the crystal, they are diffracted by the periodic electron density in the crystal. The diffraction pattern is then detected by the diffractometer which records the pattern of separation and intensity of the spots. Through a mathematical function known as a Fourier Transformation that permits the recorded diffraction pattern data to be transformed to an

electron density map, one can determine the arrangement of atoms in the crystal. The atoms exist where the electron density is highest. (Nelson, 2014)

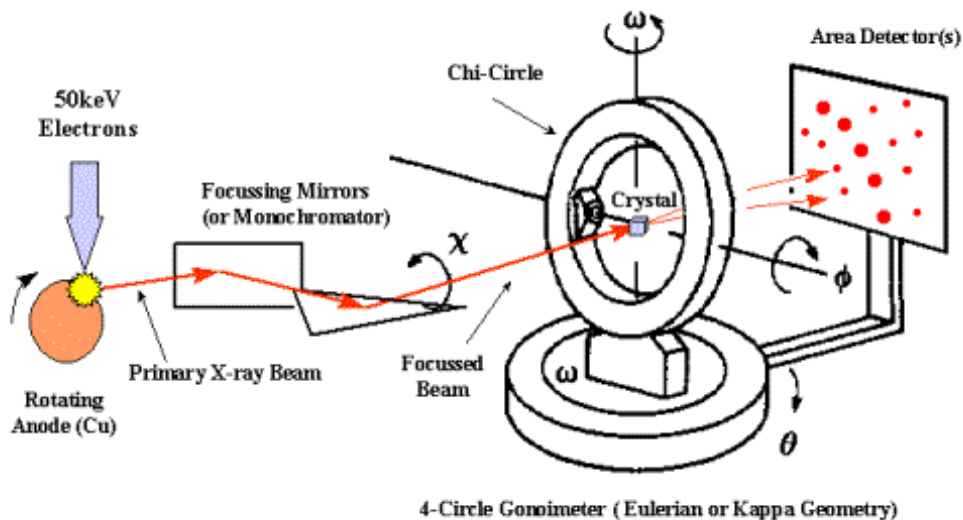


Fig 2- Diagram of X-Ray Diffraction experiment ("X-Ray Diffraction Experiment," n.d.)

1.2 Relevance to Pesticides

Many pesticides can be crystallized and therefore their molecular structure can be studied using X-ray crystallography. Doing so can reveal geometrical details helpful in deciphering a structure-activity relationship among biologically active pesticides that can explain its biological interactions with the target organism. More importantly, X-ray crystallography opens the door to investigate the negative health effects the pesticide may have. While there are hundreds of pesticides in existence, this paper will narrow the scope of investigation to three pesticides of particular interest: Aldrin, DDT, and Mecoprop.

1.21 Aldrin

Aldrin is a chlorinated insecticide that was widely used in agriculture before it was banned in the 1970s in most countries. Aldrin was discovered by Julius Hyman in 1948 and was first used commercially in the United States during the 1950s on crops such as corn, cotton, and potatoes to control soil insects including the wireworm, rice water weevil, and grasshoppers (United Nations Industrial Development Organization, n.d.). After the '70s, Aldrin was partially reinstated to exterminate termites until production discontinued in the '80s. Aldrin works by entering the insect where it metabolizes into dieldrin and kills the insect. However, like most insecticides, Aldrin causes several environmental and health concerns which ultimately led to its ban and disuse (Gilbert, 2014). Aldrin belongs to a class of pesticides known as "Persistent Organic Pollutants" (POPs) because it is difficult to remove from the environment and does not break down easily. It has low solubility in water and low volatility, clinging to soil and sediment particles. Plants take up Aldrin and store it in the roots and leaves, and when animals eat the Aldrin/dieldrin contaminated material, it gets stored in the animal's fat tissue and bioaccumulates. Ultimately, humans consume the contaminated foodstuff, and the dieldrin may take years to metabolize and exit the body. Exposure to Aldrin is toxic to a wide range of animals, including humans. Health problems include immune, reproductive, and nervous system damage, an increased rate and severity of pancreatic, lung, and breast cancer, and Parkinson's disease (Crinnion, 2009, p.355). Some of these biological health effects can be attributed to the fact that Aldrin is an endocrine disruptor.

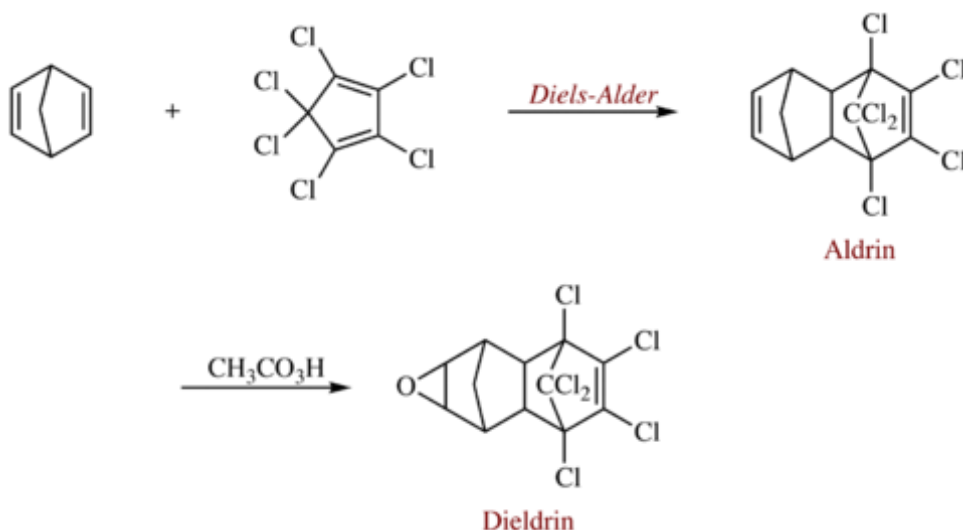


Fig 3- Diagram showing synthesis of Aldrin (top right) and metabolism into Dieldrin (bottom) (T, 2010)

1.22 DDT

DDT is an organochlorine compound with a range of uses but is most known for its insecticidal properties. It was first synthesized by Othmar Zeidler in 1874, but it was not until 1939 that its insecticidal uses were discovered by Paul Hermann Müller. DDT was then subsequently used in the Second World War to control typhus and malaria very effectively, even eliminating the disease in parts of Europe. After the war, DDT was made available for agricultural purposes and used to eradicate malaria in many parts of the world. However, there were several drawbacks to the widespread use of DDT that led to it being banned in many countries from the 1970s through the 1980s. Environmentally, DDT is toxic to many animals and also causes eggshell thinning, impacting the reproduction and population of several bird species such as the American Bald Eagle. DDT also has detrimental health effects on humans. DDT, similar to Aldrin, is a POP and comes in contact with humans in a similar manner, through the consumption of contaminated

foodstuff. Exposure to DDT can cause neurological problems and immune system damage, and is linked to an increased risk of various cancers (Crinnion, 2009, p.355). Like Aldrin, DDT is also an endocrine disruptor.

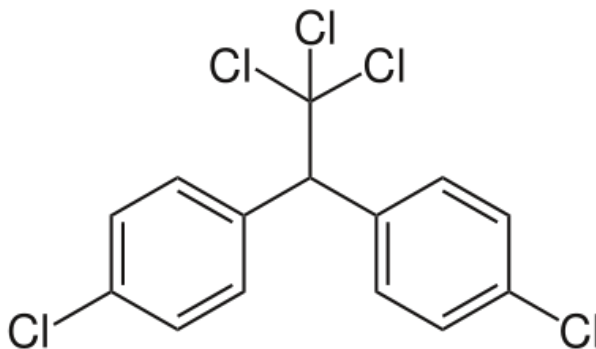


Fig 4- Molecular diagram of DDT (Leyo, 2010)

1.23 Mecoprop

Mecoprop is a general use pesticide first introduced in the 1960s and still used today. It is a weed-killer, controlling broadleaf weeds such as clover, ground ivy, and stitchwort, commonly applied on the turf of residential lawns and sports fields (Mergel, 2011). Mecoprop works by being absorbed into the weeds where it accumulates in the roots and shoots, inhibiting growth of the plant and damaging vascular tissue. Unlike Aldrin and DDT, Mecoprop is relatively safe to use and has considerably fewer environmental and health concerns (Scottish Environmental Protection Agency, n.d.). However, Mecoprop easily latches on to soil particles and readily dissolves in water, thus having the potential to seep into groundwater supplies. Normal background exposure to Mecoprop is typically safe for humans, but high dosage or long term exposure may lead to eye and skin irritation, nausea, and inflammation of nerve endings (Pesticide Action Network UK, n.d.). Mecoprop also appears to be an endocrine disruptor.

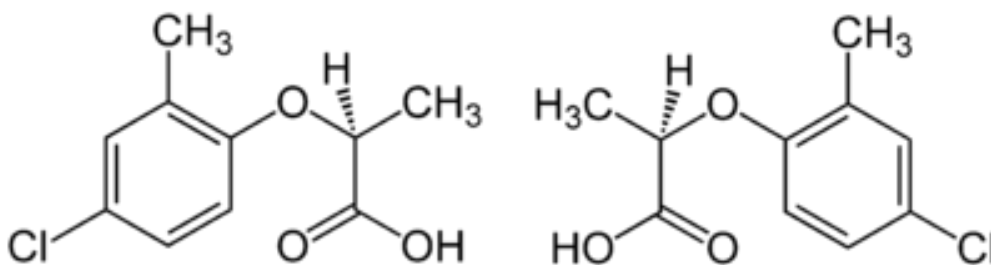


Fig 5- Molecular diagram of Mecoprop. Mecoprop is a stereoisomer, with Mecoprop-p (right) possessing the herbicidal properties (Jü, 2009)

1.3 Problems

Interestingly, the three aforementioned pesticides are all Endocrine Disrupting Chemicals (EDCs). Endocrine disruptors interfere with the body's hormones, and exposure to endocrine disruptors in the early stages of life is especially dangerous, as that can severely affect normal development. They work by binding to the estrogen receptor and either block or stimulate hormone production, creating imbalance. Health effects may not be evident until a later stage in life when problems with learning, behavior, reproduction, and increased susceptibility to other diseases and cancers occur (Lyons, n.d.).

There are still gaps in knowledge regarding the potential harmful effects of pesticides and the molecular mechanisms that cause them. Therefore, it is paramount to study how these pesticides impact the greater environment, in particular how their molecular structure contributes to its intermolecular interactions with the target molecule in the body, in this case the estrogen receptor. This paper aims to examine the molecular structure of these three pesticides (Aldrin, DDT, Mecoprop) and detail their various interactions

such as hydrogen bonding, pi stacking, and halogen bonding. Halogen bonding is of particular interest because it was only recently regarded as an important intermolecular interaction which may predictably have nuanced influence on the pesticide's biological activity that has yet to be documented. Ultimately through this investigation, the significance of a molecule structure to the biological interactions in the human body will be demonstrated to shed light on the mechanisms of a global problem.

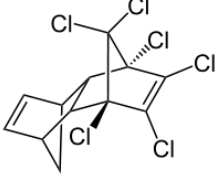
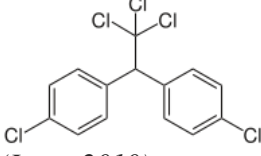
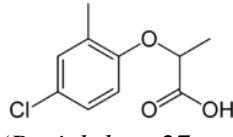
By examining the molecular structure of each of these three pesticides (Aldrin, DDT, Mecoprop), this paper will explore how a molecule's structure affects its biological activity with particular reference to the endocrine disrupting behavior these molecules share, which will ultimately shed light on the mechanisms of a global problem.

2. Results

EDCs interfere with the endocrine system, a network of glands that secrete hormones to target cells and organs. Estrogen is one of these hormones that play a vital role in the human body. Estrogen is found in both males and females, though natural levels are generally higher in females than males. During puberty, the release of estrogen promotes the growth and development of female reproductive organs, and likewise stimulates male sperm maturation and is essential for the male reproductive system (Nichols, 2015). There are several forms of estrogen; the most active is estradiol. Estradiol works by binding to the target cell, the estrogen receptor (ER), which causes it to activate. The ER then attaches itself to a certain region of

DNA in the nucleus, causing specific genes to activate and begin transcription, creating the necessary proteins for healthy development (Tulane University, 2014). The three EDCs studied here demonstrate estrogenic activity, interfering with the regular function of estrogen. For example, DDT mimics estradiol and binds to the ER instead, overstimulating protein production which is thought to influence risk of breast cancer (Nichols, 2015). Therefore, it is clear that balance of hormonal activity is crucial to normal development and these EDCs do pose a certain danger to humans.

Table 1.0- General information on pesticides from X-Ray diffraction experiment (data from *Cambridge Structural Database*)

	Aldrin	DDT	Mecoprop
Diagram	 <p>(NEUROtiker, 2008)</p>	 <p>(Leyo, 2010)</p>	 <p>(Benjah-bmm27, 2008)</p>
Chemical formula	$C_{12} H_8 Cl_6$	$C_{14} H_9 Cl_5$	$C_{10} H_{11} Cl O_3$
Compound name	1,2,3,4,10,10-Hexachloro-1,4,4a,5,8,8a-hexahydro-endo-1,4-exo-5,8-dimethanonaphthalene	1,1-bis(p-Chlorophenyl)-2,2,2-trichloroethane	(+)-2-(4-Chloro-2-methylphenoxy)-propionic acid
Space group	$P 2_1/n$	$P c a 2_1$	$P 2_1/c$
Cell lengths	a 10.806(2) b 14.743(1) c 8.992(1)	a 9.963(1) b 19.200(2) c 7.887(1)	a 4.395(3) b 6.813(5) c 36.75(2)
Cell angles	α 90 β 93.03(1) γ 90	α 90 β 90 γ 90	α 90 β 106.01(5) γ 90
Cell volume	1430.54	1508.7	1057.73
Color	Colorless	Colorless	Colorless
R-Factor (%)	4.8	9.1	5
CSD Reference Code	ALDRIN	CPTCET10	CMPXPA
Reference	T.P.DeLacy,	T.P.DeLacy,	G.Smith,

	C.H.L.Kennard, J.Chem.Soc.,Perkin Trans.2 (1972), 2153, doi:10.1039/p29720 002153	C.H.L.Kennard, J.Chem.Soc.,Perkin Trans.2 (1972), 2148, doi:10.1039/p297200 02148	C.H.L.Kennard, A.H.White, P.G.Hodgson, Acta Crystallogr.,Sect.B:S truct.Crystallogr.Cry st.Chem. (1980), 36, 992, doi:10.1107/S05677 40880005134
--	--	---	--

2.1 Aldrin

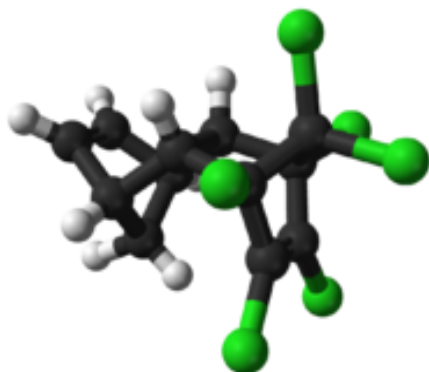


Fig 6.1- Molecular structure of Aldrin (*Mills, 2009*)

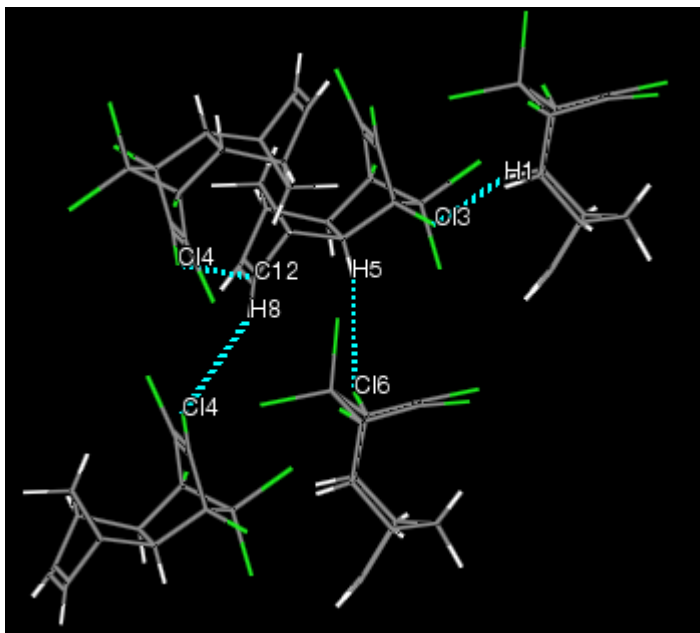


Fig 6.2- Diagram showing Aldrin's intermolecular bonding (*Made by Author*)

Intermolecular Bond	Atom 1---Atom 2	Bond Length (Å)
Halogen bond	H8---Cl4	2.919
Halogen bond	Cl3---H1	2.913
Halogen bond	H5---Cl6	2.922
Halogen bond	C12---Cl4	3.444

Aldrin is known as a chlorinated pesticide because it is composed of carbon, hydrogen, and chlorine. Aldrin is characterized by its hexachlorocyclopentadiene group, which makes up one of the two 5-membered rings in Aldrin. As can be seen from its structure, Aldrin can be split into two halves, a chlorine rich side and a hydrogen rich side. The chlorine rich end would carry slightly more electron density because chlorine is highly electronegative and can be hydrophilic, while the

hydrogen end is more nonpolar and thus hydrophobic. Aldrin is also capable of forming several strong halogen bonds using its chlorine atoms. The toxic activity from Aldrin comes from the hexachlorocyclopentadiene group and also the presence of carbon double bonds (Kaushik & Kaushik, 2007).

2.2 DDT

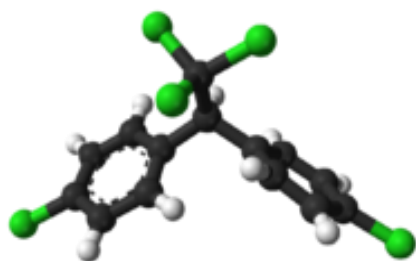


Fig 7.1- Molecular structure of DDT (*Mills, 2009*)

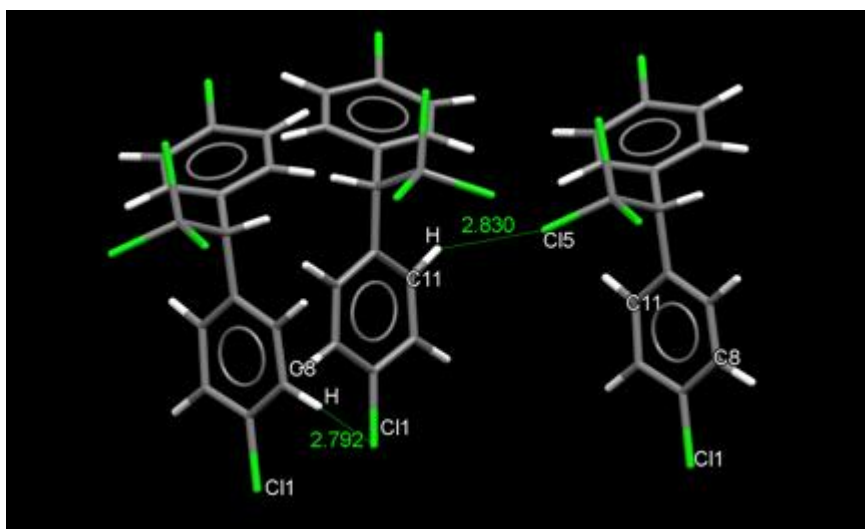


Fig 7.2- Diagram showing Halogen Bonds in DDT (*Rossi, 2015*)

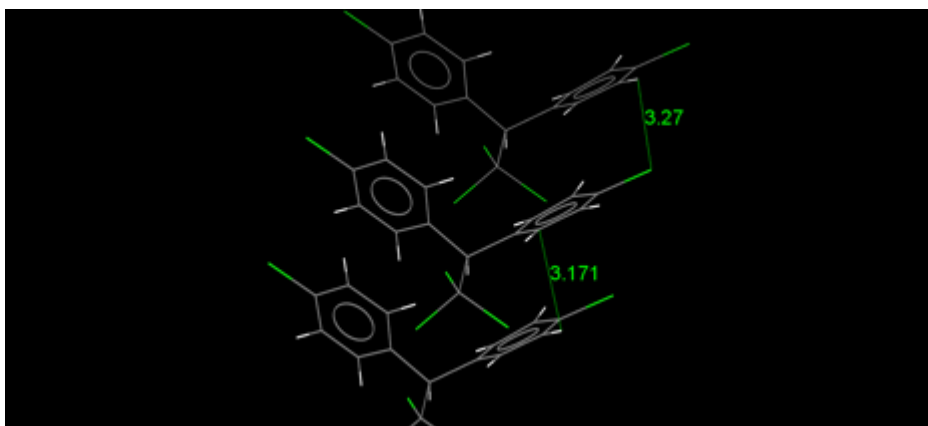


Fig 7.3- Diagram showing Pi stacking in DDT (Rossi, 2015)

Intermolecular bond	Atom 1--- Atom 2	Bond length (Å)
Halogen bond	Cl1--H8	2.792
Halogen bond	H11---Cl5	2.830
Pi stacking	Between phenyl rings	3.171
Pi stacking	Between phenyl rings	3.270

DDT is also a chlorinated pesticide. DDT contains two phenyl rings on different planes, and because of the substantial flatness of the molecule, the Pi-Pi interaction that arises from Pi stacking is slightly stronger than expected, with the shortest bond length at 3.171 Å. DDT is also capable of halogen bonding between the chlorine and hydrogen atoms. As chlorine is electronegative and holds the majority of the electron density in the carbon-chlorine covalent bond, it acts as an electron donor to hydrogen which has a partial positive charge as the bulk of the electron density is centered in the pi-electron cloud above and below the phenyl ring. This halogen bond is stronger than the interaction from Pi stacking, with shorter bond length ranging from 2.792-2.830 Å.

2.3 Mecoprop

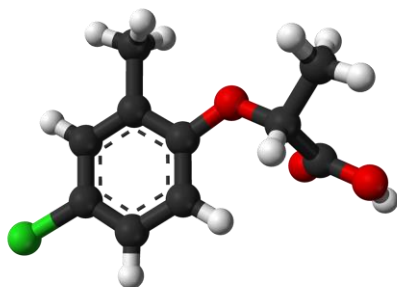


Fig 8.1- Molecular structure of Mecoprop (*Mills, 2009*)

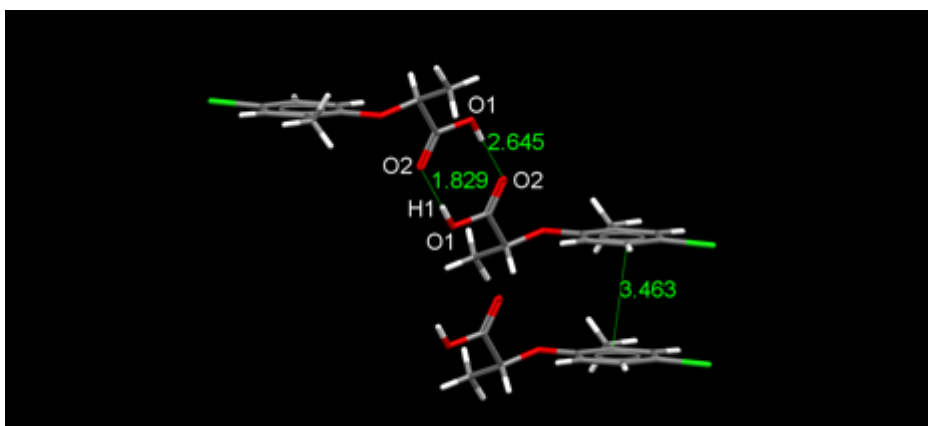


Fig 8.2- Diagram showing intermolecular interaction (hydrogen bonding and Pi stacking) in Mecoprop (*Rossi, 2015*)

Intermolecular bond	Atom 1 --- Atom 2	Bond length (Å)
Hydrogen bonding	H1---O2	1.829
Hydrogen bonding	O2---H1	2.645
Pi stacking	Between phenyl rings	3.463

Mecoprop is a chlorophenoxy herbicide. The overall shape is quite flat apart from the protruding methyl groups and most evidently the carboxyl group. Mecoprop has one phenyl ring that allows it to perform Pi-Pi interactions

that are relatively weak, as can be seen by the longer bond length of 3.463 Å. Most importantly, the presence of the -COOH carboxyl functional group gives rise to hydrogen bonds, which are very strong intermolecular interactions. These bonds cover short distances, with the shortest at 1.829 Å, indicating the strength of these bonds.

2.4 Estradiol

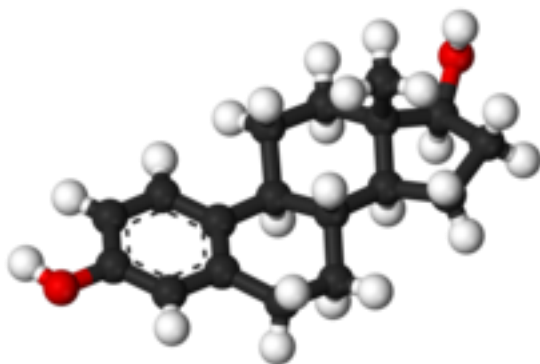


Fig 9.1- Molecular diagram of Estradiol (*Benjah-bmm27, 2006*)

Estradiol is a somewhat long and flat molecule. The key feature of estradiol structurally speaking with regard to its interaction with the ER is its phenol ring and the presence of a hydroxyl group. Both these components allow estradiol to form hydrogen bonds that strongly bind to the ER. In addition, the phenol ring with its aromatic structure can also perform Pi stacking and other Pi interactions. The rest of the molecule largely participates in a number of other hydrophobic contacts (Brzozowski & Pike, 1997, p. 755).

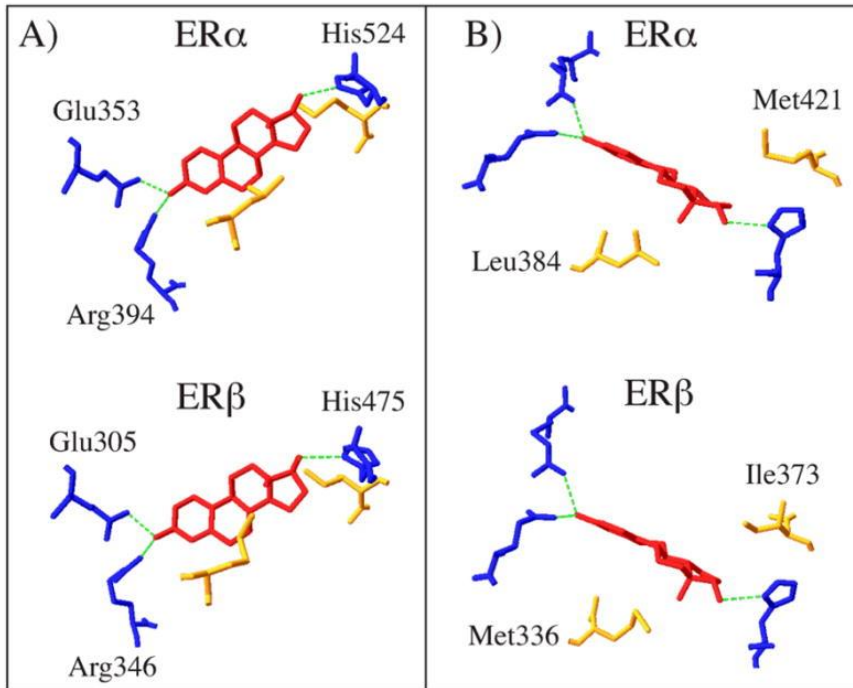


Fig 9.2- Diagram showing estradiol molecule (in red and yellow) interacting with amino acids in the ER ("Crystal Structures of ER," 2011)

3. Discussion

From the structural data of each molecule, the estrogenic activity of these pesticides can be explained and deduced. It is known that molecules that have structural similarity to estradiol have greater binding affinity to the ER, increasing their estrogenic disruption capabilities by either blocking or mimicking estrogen. The presence of a benzene or phenol ring functional group as exists in estradiol was found to be an important structural motif with regard to successful ER binding (Jaga, 2000). In general, molecules with two ring structures separated by two carbon atoms tend to have higher binding affinities compared to molecules with a single ring structure or two rings separated by a single carbon atom (Blair et al., 1999). In terms of

intermolecular interactions, the ability to form hydrogen bonds and execute hydrophobic interactions as estradiol does also played a major role in a molecule's ER binding affinity (Blair et al., 1999). The ER itself is also a "fairly promiscuous nuclear receptor" (Shanle & Xu, 2011) as its ligand binding pockets are relatively large, ranging from 390 \AA^3 - 450 \AA^3 , compared to the 245 \AA^3 of estradiol, allowing a variety of small molecules access to the binding domain.

Though on the surface Aldrin appears not to have many similarities with estradiol, it does indeed show some weak estrogenic activity. Aldrin does not share the phenolic ring structure of estradiol, but instead has two 5-membered rings which are spaced apart. Aldrin can also engage in the hydrophobic interactions necessary to bind to the ER, while many of the chlorine atoms can perform halogen bonds making up for the lack of hydrogen bonds.

DDT is well known for many things, including its endocrine disrupting capabilities. From the structural analysis it is clear how DDT causes problems in the body, in particular with the ER. DDT shares many structural similarities with estradiol, namely the aromatic ring moiety, that allow DDT to interact with the ER in many of the same ways estradiol interacts. However in DDT, the two rings are linked by a single sp^3 carbon, and while each ring can rotate along the single bond which allows some conformational mobility, they cannot be on the same plane as in estradiol (Mrema et al., 2013), hampering binding effectiveness. Even so, one of the aromatic rings in DDT arranges itself perpendicularly to the aromatic ring of the aromatic amino acid in the ER, resulting in a T-shape aromatic stacking.

The strength of this interaction is comparable to that of hydrogen bonding, thus strengthening the overall interaction between DDT and the ER (Zhang, Zhang, Wen, Zhang, & Liu, 2012).

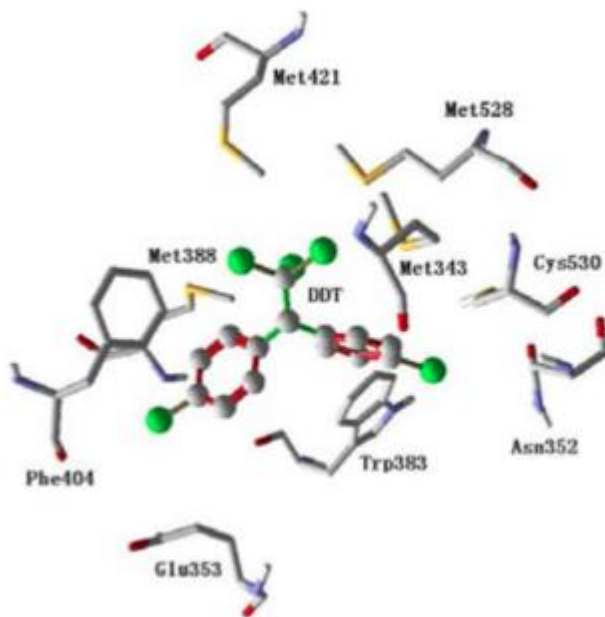


Figure 10- Diagram of DDT interaction with amino acids in ER, with T-type aromatic stacking occurring with Phe-404 and Trp383 (Zhang, Zhang, Wen, Zhang, & Liu, 2012).

While there are no definitive studies to explain the effect of Mecoprop with the ER, the structural similarities between Mecoprop and estradiol suggest it too acts as an endocrine disruptor. Mecoprop, like estradiol, has an aromatic ring structure and hydroxyl group that would allow it to perform similar hydrogen bonds and aromatic ring interactions. In particular, the spacing of the hydrogen bond and hydrophobic interactions are in comparable arrangements, though Mecoprop is the smaller molecule. Therefore, in many ways, Mecoprop could mimic estradiol and demonstrate estrogenic activity.

4. Conclusion

On the molecular level, the difference of a single oxygen atom in chemicals could determine whether they are toxic or not. Examining the molecular structure of Aldrin, DDT, and Mecoprop has revealed crucial mechanisms to how these pesticides interfere with the normal function of the human body. Chemicals with an aromatic ring moiety and capable of hydrogen bonding show interactions similar to the interactions that estradiol has with the ER and pose a clear hazard to the body's endocrine system. The presence of the recently recognized halogen bond has its implications in intermolecular interactions, which this paper suggests needs further investigation. Although DDT and Aldrin have been banned for many years, their persistence in the environment indicates that they still pose a threat to humans today, particularly in developing countries that still rely on these pesticides for agricultural purposes. Mecoprop is widely used around the world even though the possible health or environmental effects are not clearly understood. Further investigation of Mecoprop is warranted because structural analysis and comparisons have demonstrated the possible harm Mecoprop could have. Overall, it is evident that molecular structure has a crucial relationship to chemical properties, which in turn has its own wide ranging implications. In particular with pharmaceutical drug design, an understanding of a chemical's molecular structure is vital for creating drugs and pharmaceuticals that are safe and will not come round to burden society.

Acknowledgement:

I would like to acknowledge Professor Miriam Rossi for her indispensable guidance and help in the process of writing this paper. Without her, none of this would be possible whatsoever.

5. References

Benjah-bmm27. (2006, December 12). *Oestradiol-3D-balls* [Photograph].

Retrieved from

<https://en.wikipedia.org/wiki/Estradiol#/media/File:Oestradiol-3D-balls.png>

Benjah-bmm27. (2008, September 8). *Mecoprop-2D-skeletal* [Photograph].

Retrieved from

<https://en.wikipedia.org/wiki/Mecoprop#/media/File:Mecoprop-2D-skeletal.png>

Blair, R. M., Fang, H., Branham, W. S., Hass, B. S., Dial, S. L., Molland, C. L., . . . Sheehan, D. M. (1999). The estrogen receptor relative binding affinities of 188 natural and xenochemicals: Structural diversity of ligands. *Toxicological Sciences*, 54(1), 138-158.

<http://dx.doi.org/10.1093/toxsci/54.1.138>

Brzozowski, A. M., & Pike, A. C. W. (1997). Molecular basis of agonism and antagonism in the oestrogen receptor. *Nature*, 389(6652), 753-758. Retrieved from Ebscohost database. (Accession No. 9710195678)

Crinnion, W. J. (2009). Chlorinated pesticides: Threats to health and importance of detection. *Alternative Medicine Review*, 14(4).

- [Crystal structures of ER α and ER β LBD bound to E2]. (2011, January 14). Retrieved from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3119362/figure/F1/>
- Gilbert, S. (2014, June 18). Aldrin. Retrieved July 30, 2015, from <http://www.toxipedia.org/display/toxipedia/Aldrin>
- Jaga, K. (2000). What are the implications of the interaction between DDT and estrogen receptors in the body? [PDF]. *Medical Hypotheses*, 54(1), 18-25.
- Jü (2009, August 25). *Structures of the two enantiomeric forms (S left, R right) of mecoprop* [Photograph]. Retrieved from [https://en.wikipedia.org/wiki/Mecoprop#/media/File:\(%C2%B1\)-Mecoprop_Enantiomers_Formulae.png](https://en.wikipedia.org/wiki/Mecoprop#/media/File:(%C2%B1)-Mecoprop_Enantiomers_Formulae.png)
- Kaushik, P., & Kaushik, G. (2007). An assessment of structure and toxicity correlation in organochlorine pesticides [PDF]. *Journal of Hazardous Materials*, 143, 102-111.
- Kukull, K. (2015, February 8). *The electromagnetic spectrum* [Photograph]. Retrieved from <https://sites.google.com/a/msvl.k12.wa.us/cms6thgradesci/classroom-news/theelectromagneticspectrum>
- Leyo. (2010, January 1). *Chemical structure of DDT* [Photograph]. Retrieved from <https://en.wikipedia.org/wiki/DDT#/media/File:P,p%27-dichlorodiphenyltrichloroethane.svg>

- Lyons, G. (n.d.). Endocrine disrupting pesticides. Retrieved July 30, 2015, from <http://www.pan-uk.org/pestnews/Actives/endocrin.htm>
- Mergel, M. (2011, March 23). Mecoprop-p. Retrieved July 30, 2015, from <http://www.toxipedia.org/display/toxipedia/Mecoprop-p>
- Mills, B. (2009, January 22). *Ball-and-stick model of the aldrin molecule* [Photograph]. Retrieved from <https://commons.wikimedia.org/wiki/Category:Aldrin#/media/File:Aldrin-from-xtal-3D-balls.png>
- Mills, B. (2009, January 25). *Ball-and-stick model of the (R)-(+)-mecoprop molecule* [Photograph]. Retrieved from [https://commons.wikimedia.org/wiki/File:\(R\)-mecoprop-from-racemic-xtal-3D-balls-B.png](https://commons.wikimedia.org/wiki/File:(R)-mecoprop-from-racemic-xtal-3D-balls-B.png)
- Mills, B. (2009, May 13). *Ball-and-stick model of the DDT molecule* [Photograph]. Retrieved from <https://en.wikipedia.org/wiki/DDT#/media/File:DDT-from-xtal-3D-balls.png>
- Mrema, E. J., Rubino, F. M., Brambilla, G., Moretto, A., Tsatsakis, A. M., & Colosio, C. (2013). Persistent organochlorinated pesticides and mechanisms of their toxicity [PDF]. *Toxicology*, 307, 74-88.
- Nelson, S. A. (2014, September 13). X-ray crystallography. Retrieved July 30, 2015, from <http://www.tulane.edu/~sanelson/eens211/x-ray.htm>
- NEUROtiker. (2008, December 8). *Aldrin* [Photograph]. Retrieved from <https://en.wikipedia.org/wiki/Aldrin#/media/File:Aldrin.svg>

Nichols, H. (2015, June 16). Estrogen: What does estrogen do? Retrieved July 30, 2015, from

<http://www.medicalnewstoday.com/articles/277177.php>

Pesticide Action Network UK. (n.d.). 2,4-D fact sheet. Retrieved July 30, 2015, from <http://www.pan-uk.org/pestnews/Actives/24d.htm>

Rossi, M. (2015, July 24). *DDT intermolecular interaction* [Photograph].

Rossi, M. (2015, July 24). *Mecoprop intermolecular interaction* [Photograph].

Scottish Environmental Protection Agency. (n.d.). Mecoprop. Retrieved July 30, 2015, from

<http://apps.sepa.org.uk/sripa/Pages/SubstanceInformation.aspx?pid=149>

Shanle, E. K., & Xu, W. (2011). Endocrine disrupting chemicals targeting estrogen receptor signaling: Identification and mechanisms of action. *Chemical Research in Toxicology*, 24(1), 6-19.

<http://dx.doi.org/10.1021/tx100231n>

T, Dr. (2010, January 8). *Synthesis of dieldrin via a diels-alder reaction* [Photograph]. Retrieved from

<https://en.wikipedia.org/wiki/Dieldrin#/media/File:DieldrinSynthesis.png>

Tulane University. (2014). Endocrine disruption tutorial. Retrieved July 30, 2015, from <http://e.hormone.tulane.edu/learning/docking-receptor-binding.html>

United Nations Industrial Development Organization. (n.d.). Aldrin.

Retrieved July 30, 2015, from <http://www.unido.org/en/what-we-do/environment/safeguarding-the-environment/stockholm-convention/facts-and-figures/what-are-persistent-organic-pollutants-pops/aldrin.html>

[X-ray diffraction experiment]. (n.d.). Retrieved from

http://pruffle.mit.edu/atomiccontrol/education/xray/xray_diff.php

Zhang, S., Zhang, J., Wen, Y., Zhang, C., & Liu, W. (2012). Distinct mechanisms of endocrine disruption of DDT-related pesticides toward estrogen receptor and estrogen-related receptor β [PDF]. *Environmental Toxicology and Chemistry*, 31(11), 2597-2605. <http://dx.doi.org/10.1002/etc.1986>



Sense and Sensibility of Artificial Intelligence: When Machines Can Think

Yu Que

Author background: Yu grew up in China and currently attends The Affiliated High School of South China Normal University in Guangzhou. Her Pioneer seminar program was titled "Science, Technology and Society."

Introduction

Artificial intelligence (AI) is one of the most breathtaking topics in science and technology. While the time frame of its real emergence has long been the focus of debates, the moral impact of such a practice is a more urgent question for the human race to answer. Artists and critics from Mary Shelley in 1818 to Stephen Spielberg in 2001 have expressed their views on this problem already. British computer scientist Alan Turing was asked the question, "Can machines think?" half a decade ago, after which he developed the Turing test as a means to prove that there are "imaginable computers which would do well [in the imitation game]."¹ In this research paper, I propose to answer the question "What are the possible consequences when machines can actually think?" by considering both the current development of AI in scientific fields and the moral implications of AI as explored in works of art.

¹ "The imitation game" is an experiment Turing developed in order to test whether a machine has human-like intelligence or not. A more detailed explanation on this concept is available in the following section of this paper.

It might still take years before we witness the first AI to successfully pass the Turing test (i.e. to prove to have human-like intelligence), but the possible challenges of this foreseeable moment are already apparent. If machines can have qualities that register as human, for example, do we have to alter or expand the definition of humanity to include cyborgs and AI? Or, is this definition flawed per se?

All in all, we shall be prepared for whatever lies ahead, for a super AI might either lead our race to immortality or to extinction.²

1. Defining Artificial Intelligence

Artificial intelligence, a term coined by John McCarthy in 1955, refers to machines and computers created by human beings that exhibit intelligent characteristics. To define whether a machine has “intelligence” or not, scientists have several criteria. The **Turing test** is a method to evaluate one of them—reasoning.

The Turing Test, or “the imitation game,” is a test developed by British computer scientist and mathematician Alan Turing. Turing proposed that in such a test, a human interrogator would have conversations on a text-only basis with two players, one of which would be a human and the other one of which, a machine. The interrogator’s task would be to tell which one of the two responders is human and which one is the computer. If the interrogator

² Urban, Tim. "The AI Revolution: Road to Superintelligence - Wait But Why." *Wait But Why*. Web. 17 July 2015.

fails to recognize the difference, then it shall be said that the machine has passed the Turing test and that it has intelligence equivalent to that of a human.³

It is also possible to test the emotional capability of a machine using the **Voight-Kampff device**. The idea originated from the novel *Do Androids Dream of Electric Sheep* (1968) written by Phillip K. Dick, which serves as the theoretical basis for the 1982 science fiction film *Blade Runner*.⁴ The test includes one interrogator as in the Turing test, who would ask the test-taker a series of emotionally provocative questions. It is not the given responses but the involuntary change in the examinee's bodily features and autonomic nervous system—respiration rate, blood pressure, fluctuation of the pupil, etc.—that are of interest. The VK test evaluates another dimension of so-called “intelligence,” namely, empathy.

Analogously, when a machine passes the VK test, we can claim that it possesses emotions equivalent to those of a human being.

It is fairly reasonable to declare a machine that has successfully met both criteria as “conscious,” or as a strong artificial intelligence.

2. Artificial Intelligence in a Literary Context

Artificial intelligence, or the concept of it, showed up in science fiction much earlier than its occurrence in the real world. *Frankenstein, or The*

³ McCarthy, John. "WHAT IS ARTIFICIAL INTELLIGENCE?" revised 12 Nov. 2007. Web. 6 July 2015.

⁴ *Blade Runner* (directed by Ridley Scott) is a movie set in Los Angeles, 2019. The blade runner Deckard (Harrison Ford), a special force policeman, is on a mission to terminate four replicants i.e. artificially created clones that are designed to serve in human colonies.

Modern Prometheus, is a novel written by English author Mary Shelley in 1818. In this masterpiece, the young science student Victor Frankenstein creates a living creature by assembling pieces of corpses. The creature has a hideous look and everyone he meets is afraid of him, which drives him into a secluded life in the woods. A family lives close to him in a small cottage. He learns human language from listening to the family's daily conversations and from reading the books they throw away. Understandably, before long he yearns to make friends with the family, hoping that his human-like intelligence will make them feel somehow connected with him. The family moves away immediately out of fear, however, leaving the creature with nothing but outrage and self-doubt. He blames Frankenstein for bringing him to life. The scientist suffers a painful loss of loved ones thereafter as his creation kills them one by one in revenge.

Frankenstein's "monster" shares several similar characteristics with artificial intelligence—both are created by human beings and possess a certain degree of self-consciousness. At first, the innocent creation would be curious about the world, hoping to build connections and learn from experience. As it grows older, it would gradually acquire emotions and the ability to make decisions. This is the time when human intervention becomes critical to its possible future development—when it is treated with fear, ignorance, or xenophobia as in *Frankenstein*, the originally neutral being turns into a monster. When it is treated with love and care as in Stephen Spielberg's movie *A.I.* (2001), conversely, machines can also exhibit incredible emotions.

A.I. (2001) depicts a futuristic society where people can easily get robots that have the ability to “love.”⁵ David, having an appearance similar to a human boy, is actually one of these androids. He is brought home by the Swintons, whose son Martin is placed into hibernation because of a currently incurable disease. The hostess Monica builds a close relationship with David and therefore activates his built-in protocol, making him commit to an enduring childlike love for her. However, a cure is found for Martin and as the “sibling” rivalry grows, the family abandons David. The robotic boy then sets out in search of the Blue Fairy,⁶ holding the obsession that the fairy can transform him into a human boy and hence help him regain his mother’s affection. In a sunken amusement park David finds a statue of his fairy, and he keeps on praying in front of her for years until the ocean freezes in another Ice Age, in which the human race becomes extinct. Two thousand years later a new species of intelligence occupying the Earth revives David. They manage to use advanced technology to download David’s memory and create a clone Monica that can only live for one day for him, with whom David spends the happiest time of his life.

So, robots like David can be programmed to “love”—but can programmed emotions be regarded as real emotions?⁷ Isn’t it a violation to

⁵ The “love” embedded in these robots is a set of involuntary emotional responses. This kind of “love” resembles, but not necessarily equals to, unselfish devotion and benevolent concern for the good of another.

⁶ The Blue Fairy is a character in the fairy tale *Pinocchio*. Her magic successfully transforms Pinocchio, the puppet, into a human boy.

⁷ By “real” here, I am referring to self-generated emotions. Note that even in human society, our emotions are often “programmed” by the collective. For example, members within a community may be socialized to think or feel in certain ways. These established values, by my definition, may or may not be real, depending on whether the individual possessing them has given specific thoughts about whether to accept them.

the free will of the intelligent agents, given that they are “self-conscious” by definition? When we design and forcefully make these agents devoted to our needs, are we selfish ourselves?

On the contrary, the monster in *Frankenstein* learns to love and hate by itself through interactions with human beings. However, emotions obtained by such an empirical method can be negative and therefore can pose a threat to mankind. Here comes the challenge—is there a way to endow artificial intelligence with an individual right to develop its own emotions, but without threatening the existence of man?⁸

In the later sections, I will try to shed some light on the questions raised above from a more scientific perspective.

3. Current Developments in the Field

Human beings tend to involve emotions in their decision-making process. Recent studies have shown that advanced artificial intelligence can actually do the same. **Multi-agents system** (MAS) is a system in which more than one intelligent agent work together to achieve a collective goal. In these systems, cooperation and negotiation between machines have been observed. Computers may therefore build a digital micro-society that simulates human society, which is based upon people’s common knowledge and moral beliefs.⁹

⁸ Here I regard the maintenance of human race as being of utmost importance by default.

⁹ Martinez-Miranda, Juan, and Arantza Aldea. "Emotions in Human and Artificial Intelligence." *Science Direct*. 10 Feb. 2014. Web. 10 July 2015.

Machine deep learning is also a field within current studies on artificial intelligence. Deep learning is a skill possessed by human brains to recognize patterns. In a technological context, deep learning allows intelligent agents to perform tasks like speech recognition, machine vision, and even analysis of trends within a massive database. So when we talk about deep learning, we are not trying to build a painstakingly detailed program that tells the computer exactly what to do in every situation and controls every single move it takes—instead, we are building a machine that learns and decides on its own. A learning machine normally accomplishes its mission incrementally i.e. it deals with lower-level knowledge before stepping forward to acquire higher-level knowledge.¹⁰ It is just like a two-year-old—babbling syllables first, and then proceeding to words, and finally speaking sentences. Just as teachers don't stuff a database of literature into students' brains but only teach basic characters and words, the only thing we are programming in the machine is a framework that allows it to continuously learn and evolve on its own in the right direction.

Noticeably, the deep learning process often involves cooperation between man and machine. Let us look at the two-year-old analogy again: this machine toddler can identify various patterns and learn the meanings of certain things, but only after some basic definitions are given to it by the instructor. A child will recognize planes, for example, after the teacher identifies planes as streamlined metal flying objects that resemble birds. And he or she has to be corrected when mistaking a bird for an aircraft. A machine, similarly, can only perform recognition tasks after the operator has given it “hints”—it would have to go through several iterations and

¹⁰ Markoff, John. "Scientists See Promise in Deep-Learning Programs." *The New York Times*. 23 Nov. 2012. Web. 15 July 2015.

corrections by conscious human beings before the result it generated fits our expectations.

To put it in a real world context, Apple's personal assistant Siri and Google's Street View service are two experiments of machine deep learning and its applications. Interestingly, some people do feel emotionally attached to Siri after it has learned their preferences like a close friend. According to Judith Newman¹¹, mother of a 13-year-old autistic teenager, Siri not only teaches her boy communication skills and interesting knowledge about planes (which he is obsessed with yet she knows nothing about), but also offers him intimate companionship, or, to put it in her own word, "love."¹²¹³

Though learning machines are still tottering on the way, they undoubtedly display great potential in developing human-like intelligence and even emotions. Keeping the MAS system and deep learning in mind, we shall have confidence in the future of a super AI that can learn through empirical data, inductive reasoning, and interactions with both human and fellow intelligent agents.

4. Risks and Opportunities

¹¹ Newman, Judith. "To Siri, With Love." *The New York Times*. 18 Oct. 2014. Web. 15 July 2015.

¹² *Her* (directed by Spike Jonze) is a recent sci-fi movie that discusses "love" between man and machine. In the movie, the protagonist Theodore (Joaquin Phoenix) falls in love with his intelligent operating system Samantha (Scarlett Johansson), personified through a female voice.

¹³ To clarify—I don't consider the relationship between Siri and Newman's little boy resembles that kind of "love" as in *A.I.* (mentioned in page 4). The "love" here is more unilateral on the human side.

Philosopher and technologist Nick Bostrom defines “intelligence” as “an optimization process.” That is, an intelligent being would achieve the assigned goal using the most effective shortcut, which may or may not meet humans’ established moral principles. Author James Barratt also points out the “amorality” (lack of moral distinction) of AI, saying that it may be simply indifferent to the question of right or wrong. Assume an AI is asked to make the world a better place, and after rational consideration it decides that the elimination of human beings is essential—it will do so without hesitation.

Corresponding to these theories, the question of machines finding humans to be both illogical and even “immoral” (mostly due to dishonesty and hypocrisy) has been an underlying theme in both fictional and academic writings concerned with AI. By calling humans’ own inconsistency and bias into question, the existence of AI compels accountability in ways man as a species does not care about or want to answer for.

Therefore, it is reasonable to understand why AI, like every other technological breakthrough, presents both risks and opportunities.

Stephen Hawking, for example, is an outspoken opponent of AI. Concerned about the computer’s possibly superior intelligence, he believes it would “take off on its own, and re-design itself at an ever increasing rate,”¹⁴ while the human race would be left behind or ultimately replaced due to unsurpassable biological barriers. The founder of Tesla Motors and SpaceX,

¹⁴ Cellan-Jones, Rory. "Stephen Hawking Warns Artificial Intelligence Could End Mankind - BBC News." *BBC News*. 2 December 2014. Web. 12 July 2015.

Elon Musk¹⁵ agrees with Hawking, comparing the practice of creating AI to “summon[ing] the Demon.”¹⁶

On the contrary, a philosophy called **transhumanism** is gaining increasing support. Transhumanists argue that human beings can transcend biological limitations by harnessing the power of science and technology, often in the form of cyborgs and artificial intelligence in current social context. Leading futurist Ray Kurzweil expressed positive expectations in his book *The Singularity Is Near*, claiming that as super intelligent machines continue to develop greater intelligence at an exponential rate, the whole universe will be “saturated with intelligence” by the middle of the 21st century at the latest. With that much intelligence human beings can basically achieve everything that they have pondered but didn't have the massive computational power to undertake, such as mapping the countless ways of protein folding in order to find cures or prevention for diseases.

On top of that, we will have to re-examine the flaws of humanity when confronted by an invention that is inevitably becoming smarter and even more “moral” than ourselves. Then the real question becomes this—can, or should man still assume the dominant position?

5. Regulations and Policies

¹⁵ In December 2015, Elon Musk actually founded a non-profit research company called OpenAI that aims to “carefully” promote and develop open-source friendly AI in such a way as to benefit humanity as a whole. Refer to their official website for further information: <<https://www.openai.com/blog/introducing-openai/>>

¹⁶ Unspecified author. "Elon Musk Says Artificial Intelligence Is like "summoning the Demon"" *CBSNews*. CBS Interactive. 27 Oct. 2014. Web. 26 Nov. 2015.

To maximize the opportunities as well as minimize the risks brought along by artificial intelligence, we will have to introduce some kind of regulations in the field.

One basic suggestion is regulating studies related to AI. It means that nanotechnologists or computer scientists would be limited in their research in case a modern Frankenstein arises. But arbitrarily slowing down the rate of scientific development is against the nature of progress and is only a temporary approach.

Another proposal is simply unplugging the machine when we sense danger. Is this even possible—can you put the evil back into the Pandora’s box? There will be so many ways of machine communication at levels that we may not recognize—radio waves, digital signals, code manipulation etc. We certainly don't have the ability to lock the super intelligence up forever.

A truly plausible regulatory approach shall therefore be concerned with the essence of artificial intelligence, namely programming, or, more precisely, the particular part of the program that controls “morality.”

Science fiction author Issac Asimov once developed the **Three Laws of Robotics**¹⁷ in his 1942 short story “Runaround,” quoted below.

1. *A robot may not injure a human being or, through inaction, allow a human being to come to harm.*
2. *A robot must obey the orders given to it by human beings, except*

¹⁷ "Do We Need Asimov's Laws? | MIT Technology Review." *MIT Technology Review*. 16 May 2014. Web. 10 July 2015.

where such orders would conflict with the First Law.

3. *A robot must protect its own existence as long as such protection does not conflict with the First or Second Law.*

Although the three laws of robotics seem to be a dependable prototype of the laws of descending AI, we must notice certain faults embedded in them.

Firstly, the three laws were originally developed as a means to control machines that serve as servants for human beings. Given that intelligent agents would share the same emotions and intelligence as us, it is clearly inappropriate to continue treating them using these unfair laws of oppression and limitation.

Secondly, the three laws can only be applied in situations where the commander is a moral being. Yet mankind is not necessarily kind. Imagine a machine being threatened by a human that tries to take it apart. According to the second law of obedience, which is prior to the third law of self-protection, it can do nothing but to bear the harmful act.

The three laws of robotics, besides arbitrarily assuming that the human beings giving out orders are rational and moral, also blatantly violate the free will and individual rights of self-conscious machines. Therefore, Asimov's principles, though seemingly reasonable in his literary context, are unsatisfactory for future generations to adopt.

The idea of programming machines with a complete moral system is unrealistic as well, since there is not such a thing as a comprehensive set of moral codes. On one hand, each one of us has different modes of thinking and thus different solutions when faced with ethical dilemmas. On the other,

the flawed nature of humanity renders the definition of “morality” ambiguous and open to interpretation.

However, the human race as a whole does share some beliefs in common e.g. child abuse is evil while helping others is a merit. And as mentioned before, our goal is never to build a program that includes every single detail of operation. We are creating a learning machine instead.

Hence, my proposal is to program these intelligent agents with a starting set of values that reflects both current social order and humans’ moral ideals. With the proven ability to mimic human emotions, a machine should be able to comprehend the essence of morality, after which it may learn more about what we collectively praise or despise through deep learning and interactions with others. Eventually, it will develop its own value system derived from the original one. This continuously evolving set of values may or may not be complete enough to include a panorama of morality and emotions, but it shall be sufficient to match the society at hand, even presenting a paragon for humanity.

Note that this machine is not programmed or forced to do good. Instead, it will proactively adhere to a self-developed moral system, which hopefully can share our ideals of right or wrong. In this case, we will not violate the free will or inalienable rights that a self-conscious entity, like our learning machine here, shall reasonably possess; nor are we leaving our creation on its own, an irresponsible act that puts our very existence at risk.

Mankind is simply no match for something as strong as a super AI unless it is determined to be on our side at the very beginning. Understanding humanity’s limitations in both intelligence and morality, we

may agree that programming intelligent machines with such an inherent value system is not only a solution to carrying out responsibilities for the powerful creation, but also our best shot at survival.

Conclusion

Artificial intelligence is developing at an exponential rate, and it is only a matter of time before it surpasses human intelligence. While it undoubtedly possesses rational sense, current research exhibits new possibilities for machine sensibility and morality. Facing this unprecedented paradigm shift in science and technology, we have to keep in mind that a brighter future lies in the cooperation between human beings and machines in both dimensions.

Bibliography

Websites

McCarthy, John. "WHAT IS ARTIFICIAL INTELLIGENCE?" revised 12 Nov. 2007. Web. 6 July 2015. <<http://www-formal.stanford.edu/jmc/whatisai/>>.

Cellan-Jones, Rory. "Stephen Hawking Warns Artificial Intelligence Could End Mankind." *BBC News*. 2014. Web. 12 July 2015. <<http://www.bbc.com/news/technology-30290540>>.

Markoff, John. "Scientists See Promise in Deep-Learning Programs." *The New York Times*. 23 Nov. 2012. Web. 15 July 2015. <http://www.nytimes.com/2012/11/24/science/scientists-see-advances-in-deep-learning-a-part-of-artificial-intelligence.html?_r=1>.

Unspecified author. "Elon Musk Says Artificial Intelligence Is like "summoning the Demon"" *CBSNews*. CBS Interactive. 27 Oct. 2014. Web. 26 Nov. 2015.

<<http://www.cbsnews.com/news/elon-musk-artificial-intelligence-is-like-summoning-the-demon/>>

Rectenwald, Michael. "The Titanic and Transhumanism: The Enfolding of Myth" *Virtual Commons 2012 - The Titanic*. 2012. Web. 26 Nov. 2015.

<<http://www.nyu.edu/projects/mediamosaic/thetitanic/article.php?a=rectenwald-michael&page=2>>

Turrell, Jimmy. "New Techniques from Google and Ray Kurzweil Are Taking Artificial Intelligence to Another Level." *MIT Technology Review*. 23 Apr. 2013. Web. 10 Jan. 2016.

<<http://www.technologyreview.com/featuredstory/513696/deep-learning/>>.

Journals

Anderson, Susan Leigh. "Asimov's "three Laws of Robotics" and Machine Metaethics." *AI & Society* (2007): 477-93. Print.

Martínez-Miranda, Juan, and Arantza Aldea. "Emotions in Human and Artificial Intelligence." *Computers in Human Behavior*: 323-41. Print.

Turing, Alan M. "Computing Machinery and Intelligence." *Parsing the Turing Test*: 23-65. Print.

Bostrom, Nick, and Eliezer Yudkowsky. "The Ethics of Artificial Intelligence." *The Cambridge Handbook of Artificial Intelligence* (n.d.): 316-34. Web.

Books

Barrat, James. *Our Final Invention: Artificial Intelligence and the End of the Human Era*. St. Martins Press, 2015. Print.

Kurzweil, Ray. *The Singularity Is Near: When Humans Transcend Biology*. New York: Penguin Group, 2005. Print.

Movies

Spielberg, Stephen. *A.I.* Warner Bros., 2001. Film.

Scott, Ridley. *Blade Runner*. Warner Bros., 1982. Film.

Jonze, Spike. *Her*. Entertainment in Video, 2013. Film.



Brain-Controlled Prosthetics

Amanda Rabin

Author background: Amanda grew up in the United States and currently attends the Pine Crest School in Fort Lauderdale, Florida. Her Pioneer seminar program was titled "Mechanical Engineering."

Abstract

The emerging development of brain-controlled prosthetics is an intriguing new field that combines the medical and engineering fields to create an advanced technology. The idea of brain-controlled prosthetics began in 1924 when Hans Berger discovered the electrical activity within the human brain. At the time, this information could not be used for implanting any type of brain-controlled interface, but it started a revolution in neuroscience that with the help of better technology has led to the development of brain-controlled interfaces. Since then, further technologies have been developed for recording brain signals, which is crucial for the development of brain-controlled prosthetics. Currently three major studies are being performed using these advanced technologies to develop a sophisticated brain-controlled prosthetic. Despite all of the advances that have been made to date there is still much work to be done. The current technology is a one-way system, which means that the brain can send signals to the prosthetic but the prosthetic cannot send information such as a sense of feeling back to the brain. As further advances are made in brain-controlled prosthetics, this two-way system will make the prosthetic feel like a true replacement.

Introduction

The first prosthetic limb was created in approximately 1500 B.C when the Egyptians created the first prosthetic toe with the hope that it would be more visually pleasing to see an artificial limb rather than a missing limb. The prosthetics they created were made from simple fiber material and were not functional. As time continued there were little advancements in prosthetics and through the Dark Ages the only inventions were hook hands and peg legs, which were created for knights to conceal injuries sustained in battle. These again had little function and were used primarily for visual appeal or vanity (Bowker).

The Renaissance ushered in a period of revival with new perspectives in math and science, which caused a rebirth in the production and advancement of prosthetics that were made from iron, steel, copper, and wood. In the early 1500s Gotz von Berlichingen had an iron hand made for him after he lost his right arm in battle. The prosthetic hand could be used by having a person use his healthy hand to adjust pulleys and springs suspended by leather straps (Norton). This was the start of functional prosthetics that started a major drive for better prosthetics that continues to this day.

In the early 1500s, Ambroise Paré considered the father of modern prosthetics, created new modern amputation procedures and new prosthetics for upper and lower extremities (Norton). He invented a device that was a kneeling peg leg that had many advanced engineering features. These features included a mechanical knee that could easily be locked and bent to accommodate whether a person wanted to sit or stand. This marked the peak of advancement for a long time until 1800 when James Potts invented a

prosthetic that had a wooden shank and socket, steel knee joint, and articulated foot that could all be controlled by the tendons from the knee to ankle.

In 1846 Benjamin Palmer focused on the visual aspects of the current prosthetics by adding springs, smoothing the appearance, and concealing tendons to make it look more human-like. This greatly improved the appearance but the prosthetics were still very heavy in weight due to their steel construction so Gustav Hermann suggested that instead of steel, aluminum should be used to reduce weight (Norton).

The onset of World War I and World War II advanced the need for prosthetics because injuries were being obtained from industrial weapons. In order to save lives, surgeons had to amputate limbs due to infection from unsanitary conditions on the battlefield. Advancements included a split hook-hand attachment to artificial arms, which allowed for amputees to grip objects, and hydraulic knee joints that allowed the knee to lock and bend in an instant (Bowker). There was also a focus on using more lightweight material and Gustav Hermann's suggestion to use aluminum was implemented in 1912. In addition to using aluminum, other lighter metals and plastics were also used.

In 1924, when Hans Berger first discovered electrical activity in the brain, a new opportunity was created for humans to learn a great deal of new information about the brain. This discovery did not lead to many immediate developments because of the limits of technology. In the late 1960s animal testing with monkeys was performed to see if a monkey could control a meter needle on a screen from signals from single cortical neurons using sensors on its scalp. The monkeys were successful and proved that they

could use their brains to control activity (McMullen). In the 1980s, researchers and scientists performed more testing and they found that humans had the capability to move an image of a rocket across a screen using similar technologies as used with the testing of monkeys in the 1960s (McMullen). These developments led to the idea that if humans were able to control abstract objects with their brain then they could move prosthetic limbs using the same technique.

In the 1980s Apostolos Georgopoulos discovered a pattern of electrical firing which occurred when a person rotated or moved their arm in a certain way or direction. This discovery combined with John Chapin's technique for recording dozens of brain neurons at once by using permanently implanted electrodes, led to the realistic idea that the technology was advanced enough to start implanting it into people's brains. In 1998 Philip Kennedy surgically implanted the first brain-controlled interface into a human being. The brain-controlled interface that Kennedy implanted had very little function, but there were no complications in surgery and the device did not harm the patient. Although the brain-controlled interface had little function, it acted as a wireless di-electrode, which meant that it could conduct electricity without additional wires attached to the scalp (Bogue).

Yang Dan continued with this research by decoding neuron firings to reproduce images seen by cats. Dan did this by having a group of cats watch eight movies. Before the cats watched the movies, she embedded a series of electrodes into the thalamus portions of their brains, which relay sensory signals in order to record signals that come from the retina. When the movies ended, Dan was able to decode the signals and put together moving

images of what the cats saw (McMullen). After Dan's research, Doctor Miguel Nicolelis continued on the same path by decoding the brain activity of owl monkeys, but he went a step further by using this information to reproduce the movements of the monkeys in robotic arms. Nicolelis experimented by having monkeys play with a joystick to move an image on a screen while he decoded the monkeys' brain signals. The signals were then sent to a robotic arm to mimic the monkey's movements. This connected the monkeys' brains to the image on the screen so the monkeys dropped the joystick and controlled the image with their brains. The problem was that the monkeys could not see the arm moving and therefore did not receive any feedback. Nicolelis tried to fix this by using rhesus monkeys, which are more similar to humans. Due to their more humanly characteristics, these monkeys were able to see the arm moving and receive feedback. The monkeys were trained to reach objects on a computer screen with a joystick while the robotic arm simultaneously moved according to the monkey's direction, although the arm was hidden from the monkey's view. With more testing, the rhesus monkeys were shown the robotic arm and they could control it by viewing its movements. As the Schwartz group continued on the project, they became successful when they trained their monkeys to feed themselves fruit by using a robotic arm controlled by their own brain signals (Nicolelis). This left the lasting idea in researchers' minds that if a monkey can control a robotic arm, then humans can also control prosthetic arms by the use of their brains.

Even with all of these advancements, prosthetics still have proven to be a burden because they are abstract limbs. People with prosthetics end up favoring their human limb, which can cause back pain, hip pain, and other ailments. These pains and ailments can lead to serious health problems

(Hochberg). The only way to fully satisfy a prosthetic user would be to have a brain-controlled prosthetic to act as a real human limb.

Methods

In order for a brain-controlled prosthetic to work, a doctor must record the brain signals of a patient in order to match the signals to the prosthetic. There are several different methods of recording these signals and the safest method is to use an electroencephalogram (EEG) machine. This machine works by placing electrodes on the scalp of the patient. The scalp is about two to three centimeters away from the cortex, which means that recording each individual neuron is inaccurate leading to imprecise results. To improve accuracy the electroencephalogram machine records a group of neurons at once, but this still provides inaccuracies (Shih). Electrocorticography (ECoG) is another way to record brain signals and this is more invasive because electrodes are placed on the cortical surface through a craniotomy. This is achieved by removing a portion of the patient's skull to expose the brain. This is a high-risk surgery that can result in severe brain damage. Even though there are risks, the distance from the electrodes to the neurons is reduced to a few millimeters, which greatly increases the accuracy of the data being recorded (Shih). This technique still provides downsides because the brain signals might be slightly changed as a result of the anesthetics used in the surgery or the side effects of the surgery. Additionally, it is hard to recruit patients to join the trials for brain-controlled prosthetics if the risks for the recording surgery are so high, and the electrodes must be removed after a week or two so a second surgery is necessary.

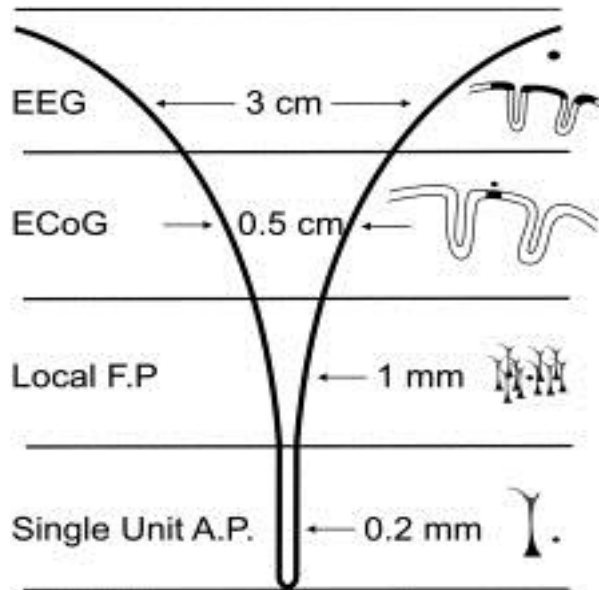


Figure 1
*Comparison of four of the techniques
for recording brain activity*

In addition there are two methods that are deemed higher risk because they are more invasive. The local field potential (Local F.P.) is an implantation that is one millimeter from the neurons and with its high gamma band activity can record single-units of neurons that are accurate for brain-controlled prosthetic implantations (Schwartz). Another way of recording brain signals is the use of microelectrodes to record single neurons even more accurately than the local field potentials because in this technique, also known as Single Unit A.P, electrodes are one fifth of a millimeter away from a group of individual neurons. This data is accurate enough to regenerate complex arm movements and the electrodes can stay permanently in the brain. Even though only one surgery is required, the surgery is very invasive because a part of the skull must be removed in order for the electrodes to remain permanently, and this can cause side effects.

Although side effects arise, the side effects go away and do not affect the accuracy of the data because the data can be collected after the period of the side effects is over. This is due to the fact that the electrodes do not have to be removed from the patient (Schwartz). This leads to very accurate results and this technique is very useful for brain-controlled prosthetics because the prosthetic and the brain will always have an accurate connection because the electrodes are never removed.

Recently another technique has been used to implant microelectrodes, which involves drilling small burr-holes into the skull. Through these holes, the doctor will place the microelectrodes, but the patient is at a high-risk for infection due to the electrodes passing through skin. Infection may cause a major brain bleed resulting in major brain damage and eventual death (Reichert).

Each of these five methods has its advantages and disadvantages. The electroencephalogram machine is the most commonly used for recording brain activity because it is the least invasive and can be done in one doctor's visit. The data is not always accurate enough for brain-controlled prosthetics to work though so electrocorticography is used. Local field potentials and microelectrodes are rarely used because of the risks they carry but they have been used in the past and when successful provide precise data. The technique involving drilling small burr-holes is usually used for patients with Parkinson's disease and other debilitating diseases, but there is currently research being done to make the technique safer and viable for use with brain-controlled prosthetics.

In order to reduce these problems, a new technique or machine must be created to make the process safer. This new machine would be nontoxic

and biocompatible in order to reduce infection. Additionally, it would have to be energy-efficient enough to be recharged every night to make sure that it lasts a long time in the brain without the need for removal.

Through recent times, there have been different methods of surgical implantations for the placement of the recording devices. There are also different ways of making the nerves on the limb that was amputated connect to the brain. Doctors have worked with existing nerve endings and some doctors have implanted devices or sensors into the muscle tissue to make the nerves work again. Through clinical trials, results have shown that neither technique is preferable. Both have given relatively the same results, but time will tell whether one technique is preferable due to a better or worse long-term effect.

Results and Discussion

Johns Hopkins University started a major prosthetic limb program in 2005 using a three-phase process. Phase one would develop the technology and phase two would create a prototype with the correct amount of weight, speed and forces in order for it to function successfully on a human being. The third phase, still in progress, is to use the technology from the first two phases to restore function to people with upper-spinal injuries that resulted in lost control of a limb. Once phase one was completed, the developers decided to use carbon fiber and titanium as their materials in order to decrease the weight (Johannes 213). The prototype from phase one and phase two is currently being tested in trials and patients seem to be reacting positively. It is not ready for market because it would cost about \$500,000 for one unit so the developers are trying to reduce the market price (Johannes 211).

In order to get these prosthetic limbs to work Johns Hopkins uses a procedure that takes free nerve endings and surgically redirects them to remaining muscle groups. When the brain thinks of an action it fires at the end nerve. Once the nerve endings are redirected they can be mapped to the missing limb. When people have missing limbs they experience phantom limb syndrome, which means that even though the limb is no longer there, it still has feeling. In order to redirect the nerve endings the doctor must find which nerve endings are viable and which parts of the body they signal. For example, a doctor will touch a specific nerve ending and the patient will feel as though there is a tingling sensation in the hand depending on which nerve the doctor touches. Once the doctor has determined which nerves cause which sensations they can perform the procedure to redirect the nerve endings. After the nerve endings have been redirected, the prosthetic limb is ready to be fitted. Patients must train themselves to use this prosthetic and the motion does not come naturally since nerve endings were redirected and not in their original place. This can cause patient frustration but after a certain period of time, he or she will hit a breakthrough and use the prosthetic as though it was his or her own limb. Even though the prosthetic may not look appealing to the average person, the body attachment has proven to be very comfortable after much research was done on how to improve it. Developers used a combination of different shape-changing socket methods to create the attachment, which is durable enough to carry the weight of the prosthetic and allows for a complete range of motion (Johannes 209).

One patient of the trial, Les Baugh, has experienced positive results from the trial. He began the trial as a double arm amputee. Within days of being fitted for the prosthetic Les Baugh was able to gain motion in both of

his arms. In a matter of months, Baugh has achieved full range of motion in both of his arms and is able to lift heavy objects (Campbell).

The Chicago Center for Bionic Limbs has also started a project to create the most advanced bionic limb. They with the help of Northwestern created a technique for connecting the nerves similar to the one that Johns Hopkins uses. The Chicago Center for Bionic Limbs uses a targeted muscle reinnervation surgery for patients with shoulder amputations. This surgery involves moving the residual brachial plexus nerves, which include the musculocutaneous nerve which controls the elbow, the median nerve which controls wrist and hand movement, radial nerves which control the elbow extensors such as the triceps and ulnar nerves which control finger flexors, to the pectoral muscles of the chest in order to directly connect them (Chicago Group). Once these connections are made, the patient must wait a few months in order for the connection to become functional so that he or she can control the prosthetic. During the trials, the Chicago Center for Bionic Limbs used a high-density electromyography machine to record the brain signals of patients, and they found that more neural control information could be extracted than in previous experiments and trials from other companies. In addition, they implemented pattern recognition techniques to allow the patient to have more advanced movement of the prosthetics. With such advanced technology, the Chicago Center for Bionic Limbs was able to have the prosthetic pick up on some texture, temperature, and pressure, but the patients barely felt it. In their current trials, they use a block test and box test which require patients to move as many one inch boxes or blocks from one area to another in two minutes and a close pin relocation test where patients must grip a pin on a horizontal bar and move it to a vertical bar (Chicago Group). These tests are used to evaluate motor and process skills.

After the reinnervation surgery, patients showed a vast improvement in both areas and were able to perform simple life skills including taking out the garbage, shaving, putting on a hat, and vacuuming, making patients less dependent on others. This means that these patients have the life skills to live on their own and take care of themselves.

Another group of people creating brain-controlled prosthetics is the Ossur Company. They also use a non-invasive surgery to control the prosthetics, but the Ossur Company places sensors in the residual tissue of the muscles instead of attaching existing nerve endings to the tissue of the muscles. They place myoelectric sensors in the tissue, which interpret the signals between the nerve endings and the muscle tissue. The prosthetic limb is then fitted with receivers to interpret the signals and turn them into movements for the prosthetic. The lead orthopedic surgeon, Dr. Thorvaldur Ingvarsson added, “Users no longer need to think about their movements because their unconscious reflexes are automatically converted into myoelectric impulses that control their bionic prosthesis” (Silver-Thorn). This is a very innovative technique that greatly reduces the time a person needs to train the tissue and nerves to adjust to the prosthetic because the nerves and tissue were never moved from one part of the body to another. This allows a person to easily move the prosthetic without even thinking about it; in the first trial it took the patient about ten minutes to start walking with the new prosthetic (Silver-Thorn). Unlike Johns Hopkins, the Ossur Company makes lower extremity prosthetic limbs. The Ossur Company has performed clinical trials for the past year with very few patients but they have been very successful.

Conclusion

With all these advancements in brain-controlled prosthetics over the past twenty years, prosthetics have a bright future with many advancements to come. In the near future the trials will continue to work out the problems of the current prosthetics and use this information in order to improve upon the prosthetics. Currently, many companies do not have their brain-controlled prosthetics on the market because the prosthetics are not affordable to most people. Once the trials are completed and the prosthetics are perfected, the prosthetics will be manufactured on a larger scale which will lower the cost of each individual part because the higher the quantity of each part ordered the lower the cost of each individual part. This will be even more important because as the prosthetics continue to advance, the more advanced technologies become more expensive and thus the economies of scale through greater production will partially offset the greater cost of advanced technology in the prosthetics.

Despite the advancements in prosthetics technology, there is still much work that needs to be done. Currently, the technology is a one-way system, which means that the brain can send signals to the prosthetic but the prosthetic cannot send signals back to the brain. This makes the prosthetic feel less human-like because it cannot give a person a sense of feeling. This can be improved in the future by developing technology that can complete the circuit as would be observed in normal electrical signal transduction. This gives the advantages of feeling temperature and texture, which can prove to be beneficial. For example, if a person is not paying attention to the prosthetic limb, it can touch the stove and become overheated. This would destroy some of the sensors. A person would not realize this because he or she could not feel the heat of the prosthetic therefore not noticing the damage. In addition, the design of the prosthetic could be greatly improved.

The prosthetic that many companies are currently developing looks unappealing to many people because it is colored in silver and black metal. This adds to the insecurities that patients feel while wearing a prosthetic. In the future, designers can make the prosthetic skin-colored in order to boost the morale of patients. Also, the sensors and wires could be better concealed to add to the design in the future. This will make the prosthetic look more natural and discourages strangers from staring at a person with a brain-controlled prosthetic. With all these ideas for improvements, researchers and scientists have a path for the future that will greatly improve the prosthetics and the lives of the people who use them.

Works Cited

Bogue, Robert. "Exoskeletons and Robotic Prosthetics: A Review of Recent Developments." *Industrial Robot: An International Journal* 36.5 (2009): 421-27. *Emeraldinsight*. Emerald Group Publishing Limited, 2009. Web. 14 July 2015.

Bowker, John H. "History of Amputation Surgery and Prosthetics." *Atlas of Limb Prosthetics: Surgical, Prosthetic, and Rehabilitation Principles*. St. Louis: Mosby Year Book, 1992. N. pag. *Digital Resource Foundation for the Orthotics and Prosthetics Community*. Mosby-Year Book. Web. 02 Jan. 2016.

Campbell, Paulette. "Amputee Makes History with APL's Modular Prosthetic Limb." *Amputee Makes History with APL's Modular Prosthetic Limb*. Johns Hopkins University, 16 Dec. 2014. Web. 12 May 2016.

Group, Chicago. "The Center for Bionic Medicine." *Rehabilitation Institute of Chicago*. Center for Bionic Medicine, n.d. Web. 9 June 2015.

Hochberg, Leigh R. "Reach and Grasp by People with Tetraplegia Using a Neurally Controlled Robotic Arm." *Nature International Journal of Science* 485.7398 (2012): 372-75. *Nature*. Nature Publishing, 26 Mar. 2012. Web. 24 June 2015.

Johannes, Matthew S. "An Overview of the Developmental Process For the Modular Prosthetic Limb." *An Overview of the Developmental Process For the Modular Prosthetic Limb* 30.3 (2011): 207-16. *Johns Hopkins University APL*. Johns Hopkins University, 2011. Web. 2 July 2015.

McMullen. "Demonstration of a Semi-Autonomous Hybrid Brain–Machine Interface Using Human Intracranial EEG, Eye Tracking, and Computer Vision to Control a Robotic Upper Limb Prosthetic." *Neural Systems and Rehabilitation Engineering* (2013): 784-96. *IEEEEXPLORE*. IEEE Engineering in Medicine and Biology Society, 12 Dec. 2013. Web. 28 June 2015.

Nicolelis, Miguel A. L. "A Brain-to-Brain Interface for Real-Time Sharing of Sensorimotor Information." *Scientific Reports* (2013): n. pag. *Nature*. Nicolelis Lab, 28 Feb. 2013. Web. 2 July 2015.

Norton, Kim M. "A Brief History of Prosthetics." *In Motion* 17.7 (2007): n. pag. *Amputee Coalition*. Amputee Coalition, Dec. 2007. Web. 20 June 2015.

Reichert, William. "Response of Brain Tissue to Chronically Implanted Neural Electrodes." *Journal of Neuroscience Methods* (2005): 1-19. *Research Gate*. Research Gate, 8 Aug. 2005. Web. 15 July 2015.

Schwartz, Andrew B. "Brain-Controlled Interfaces: Movement Restoration with Neural Prosthetics." *Neuron* 52.1 (2006): 205-20. *Cell*. Elsevier, 5 Oct. 2006. Web. 24 June 2015.

Shih, Jerry J. "Brain-Computer Interfaces in Medicine." *Mayo Clinic Proceedings* 87.3 (2012): 268-79. *US National Library of Medicine National Institutes of Health*. Mayo Clinic, Mar. 2012. Web. 31 June 2015.

Silver-Thorn, Barbara. "Functional Stability of Transfemoral Amputee Gait Using the 3R80 and Total Knee 2000 Prosthetic Knee Units." *Journal of Prosthetics and Orthotics* (2009): n. pag. *Ossur*. Web. 02 June 2015.



Hukou, Development, and Fairness

Darren Tong

Author background: Darren grew up in Hong Kong and currently attends Chinese International School in Hong Kong. His Pioneer seminar program was titled "The Political Economy of International Development."

1. Introduction

The People's Republic of China enforces the most far-reaching system of unchallenged institutionalized discrimination today. Commonly known as the *hukou* registration system, this institution ties a citizen's access to state services to their formally registered locality. Such services can range from eligibility for bus passes to access to schools and welfare benefits. The *nongzhuanfei*, or the official system for transferring one's registered residence from one locality to another, is notoriously prohibitive - leaving all but the politically connected confined indefinitely to their place of birth. Migrating to another province to give birth does not ensure a different hukou for one's child, as the hukou is hereditary.

This produces two effects: internal migrants, usually travelling to urban areas, are significantly disadvantaged in their new communities with little prospect of upward mobility. Secondly, rural residents hailing from historically underdeveloped rural areas are legally bound to weaker state services, sparser job opportunities and lower living standards. The highest estimates of the number of unregistered internal migrants in China's urban centers are clustered around three hundred million. This vast underclass of Chinese citizens weighs on political and academic discussions on urbanization and economic development.

Chairman Mao first implemented the policy to enforce the radical economic programmes of the Great Leap Forward in 1958. In an attempt to replicate the Soviet model of economic development, Mao ordered rural agricultural communities to turn over as much surplus grain as possible to support rapid industrial development. Mao then institutionalized the extensive hukou registration system in an attempt to limit rural emigration to the cities and elicit total command over agricultural and industrial labor, not unlike the Soviet leaders who issued *Prospika* internal passports to keep “bread basket” regions awash with free labor.

The philosophical basis for the policy - a catch-all tool of the state to control the labor economy - has since been lost in a cacophony of private interests. The post-Mao era of cautious market liberalization and rapid economic growth spurred a massive industrial growth spurt and demand for labor. By 2001, when China joined the WTO, industrial labor was suddenly in shortage, which slowed down economic growth. In fact, by the beginning of 2015, the traditional upward pressure on the yuan provided by foreign investments into Chinese industrial powerhouses had weakened considerably due to the relative attractiveness of Bangladesh and other low-income countries for lower costs of production. This threatens to stall manufacturing growth and entrap the Chinese economy in the so-called “middle-income trap”. Yet, little meaningful institutional reform has taken place. State-owned enterprises profit from artificially cheap labor in rural areas. Urbanites enjoy better benefits than the rest of the country. Furthermore, security bureaus are more than content to preserve the status quo in the interests of administrative ease.

The relevant development goals of the twelfth five year plan are as follows: an economic restructuring away from exports and investments and towards domestic consumption, an accelerated development of China's nascent service sector, and a rebalancing of rural-urban inequity.

If a tangible threat to these long term development goals can be established, then it becomes the interest of the CCP to enact wide-ranging reforms of the hukou system. As corruption was deemed an intolerable obstacle to Xi's "Chinese dream" of economic development and made a top priority, the hukou system must be seriously reviewed, as it forms a massive obstacle to development - one the government should do well to reform¹.

The purpose of this paper is to analyze these effects with respect to China's stated development goals. The next section will present a literature review on the topic. The following section contains notes on statistical and research methodology, while the subsequent section will contain the findings and discussions of the paper. The final section will present a conclusion in the form of a balanced answer to the research question.

2. Literature Review

The hukou system has been the subject of much scrutiny, domestically and internationally. While experts touting the system's success are difficult to find, there is a legitimate debate about the *extent* of the effects

¹ Kam Wing Chan put it bluntly in the title of an opinion article in the South China Morning Post in 2013: "China's hukou system stands in the way of its dream of prosperity"

on Chinese society and economy - whereas most simply frame it as a humanitarian issue, others claim it presents an inevitable brick wall to the future of Chinese development.

The Modern Caste System

A visible chorus of experts have framed the problem as one of classism. The system has effectively created a three-tier system—full urban citizens, urban residents with rural hukou, and a “floater population,” or “liudong renkou” that is forced to seasonally migrate between rural and urban areas as the unskilled job market waxes and wanes (Montgomery 8). He stipulates that the latter two tiers are so disadvantaged legally and economically, that their class is almost certainly propagated to their descendents. As for urban residents without proper hukou, their children, born and raised in the cities, have rural hukou and thus are ineligible for education, healthcare and housing subsidies. An estimated 19 million children live in cities without access to these state supports. As for those without a stable residence, their children are unlikely to go to school or to get a stable job. The inequality becomes twofold. Such, Montgomery says, is the “inheritance of inequality”.

Unfree and Unfair Labor

This perpetuation of inequality has far-reaching ramifications in the opinion of liberal economic scholars. Such scholars, such as Wu Xiaogang, have long that preferred to find income differentials between unregistered migrants and fully privileged urban residents as the source of economic immobility (Wu Xiaogang 56). Rural migrant wages have been estimated to be 44 percent of urban wages, before state benefits (Chan and Ross 30). In

May 2013 the first legal challenge to the hukou system, a lawsuit by Anhui-registered Su Min was brought against the Nanjing Municipal Bureau of Human Resources, alleging that her relegation to a telephone consultant job was discrimination on basis of hukou registration. This shows the wage-gap argument is still fresh in the political discussion, as stated by Montgomery.

This comparative approach, though, has its problems. These migrants, or “floaters”, are employed in unskilled, dispensable positions², whereas the urban hukou-holding working class work in longer-term industrial jobs of intermediate skill. Since these two different groups are employed in vastly different, wage inequality is a difficult metric to measure disadvantage. On the other hand, these scholars have dismissively claimed that the shorter employment durations indicate a high level of occupational mobility, and hence workplace discrimination rather than a structural labor market problem is the underlying cause of employment inequality.

The debate on occupational mobility has shifted in recent years. The opposite has been proposed: that the vast majority of unregistered migrants suffer from occupational immobility, which drags growth in the Chinese labor market and the entire economy (Zhang 2). That this large pool of labor is relegated to jobs with the least possibility of upward mobility is no accident. Short-duration construction, transportation or other areas of jobs undesirable for the urban hukou-registered residence are filled by disadvantaged migrants. The institutional mechanism of economic mobility - job experience, records and referrals - becomes unavailable, presenting a significant barrier to career progression or even the simple transfer between

² The three “D’s” - “dirty”, “dangerous”, and “difficult” are popular names for jobs reserved for migrants

different jobs. Hence, migrants face fewer job opportunities, higher job search costs, and longer inter-job intervals.

Furthermore, migrants return to their own families in the countryside on an annual basis, creating colossal seasonal migration patterns. This constitutes a portion of the “floaters” mentioned above. This seasonal migration further reduces occupational mobility, diminishing worker productivity and the supply of labor (Chan 13). Taking into account all these mobility disadvantages, the problem of social class conflict thus becomes one of institutional barriers to development instead.

In fact, since China joined the WTO in 2001, increased demand for labor in export industries has significantly increased net migration to cities - straining the hukou system, which artificially inflated the price of labor. Many scholars were quick to criticize the Chinese government, which in 2003, loosened enforcement in a bow to public pressure as a result of the police brutality against a Wuhan college student migrant resulting in his death. However, as Chan points out, the effect was simply to reduce the extent of persecution by the state against migrants, rather than increasing access to labor markets in need of workers.

A House Divided

Rural migration, seasonal or otherwise, has increasingly forced parents to leave their children in the countryside, due to uncertainty and poor standards of living created by the hukou system. The number of children with parents living or working away from the local province is estimated at 58 million (China Youth Research Center). While studies specific to the children of internal Chinese migrants are sparse, literature on long-term

parent-child separation clearly shows a profound effect on the well-being of children. The same literature also suggests that extended kinship systems are a significant buffer to the effects of parent-child separation. However, the One Child Policy completely negates this effect. While some literature suggests a negligible, or even positive impact of parental migration on the academic development of children, a larger group of academics indicate the effect is indeed negative (Lu 2008). While robust and specific investigation is still lacking, parent-child separation is a worthwhile consideration for development studies as it would further perpetuate the hereditary nature of inequality between rural and urban hukou-holders.

Underconsumption and Urbanization

The hukou system's contribution to China's dependence on manufacturing exports over the past twenty years has been widely derided by scholars. The root argument is the meagre income of rural-urban migrants contributes to sluggish domestic consumption and aggregate demand. The vast 150 million or more migrant class will continue to exert a downward pressure on consumption so long as lack of access to social security and other state services continue to discourage spending (Dreger 2014). As the government's macroeconomic focus shift to the domestic economy, significant reform of the hukou system is widely agreed to be necessary.

Urbanization, i.e. the expansion of the infrastructure capacities of small and large cities, has vastly raised the incomes of urban residents and rural migrants alike, with robust employment and retail centers. A 2013 OECD report reveals, however, that the rise in income for rural migrants was notably less pronounced than urban hukou residents, owing to a mismatch in jobs. It further uncovers a sharp decline in consumption-to-GDP ratios in

confluence with the sharp increase in urbanization in the 1990's - a phenomenon it pins on the cautious spending of rural migrants owing to the "bifurcation of the hukou system". The hukou system is seen by many scholars as a barrier to further urbanization.

The Emancipation of Women?

In the same vein as with Maoist intellectuals in the 1950's, justification for the authoritarian policy comes from the emancipation of women. It is speculated that the constraints of the hukou transformed the institution of marriage. Whereas uneducated, unskilled women were tied to their rural family's good will in the past, they could now escape by marrying a resident of a different county (Fan 3). While a hukou of a different locality could not be obtained by birth; instead, it could be gained by marriage. In this way, previously exploited and economically immobile women in the countryside were empowered by this new vehicle of social mobility - tearing down a traditional institutional barrier to unlock the economic potential of women.

In the same way that Maoist intellectuals' claims were widely derided for their narrow views, the view that the hukou system helps women are weakened by recent statistics. Indeed, Fan herself in a later paper found that among the 13.8% of migrants that moved for marriage in the nationwide 1990 census, the vast majority were between different rural regions. The reason for this is that the decentralization of the hukou system in the 1990's led to many urban-coastal provinces placing additional restrictions on hukou distribution - making it very difficult to obtain an urban hukou via marriage. The 2000 census shows a decline on the share of marriage migrants to 7.3%

(Kang 2), which discredits claims that hukou marriages are a significant source of empowerment.

Youqin Huang in 2000 released a study uncovering gender-based discrimination even between migrants of different genders - a clear demonstration that the hukou system not only does nothing for gender equality, but also perpetuates the “patriarchal structure” of Chinese society (Fan 2008).

Insubstantial Reforms

National reforms to the hukou system have been implemented in 1984, 1992, 2003 and 2014. In each reform the central government diluted the official distinctions between rural and urban hukou. In 2014, amid speculation that the system was to be abolished, the government removed the rural/urban divide completely from the language of the law. Despite this, most scholars have dismissed these reforms (Chan and Buckingham 2014), as citizens’ registration is still tied to their own localities, for which no provision by the central government has been made to equalize state services or economic opportunities.

More worrisome yet is the aforementioned decentralization of hukou registration, effectively enabling urban provinces with high net immigration to limit new hukou issuance. Typically, the handful of rural migrants that obtain hukou have already found stable jobs and are relatively wealthy. For example, Guangdong only grants urban hukou to migrants with a proven stable job and home ownership - the very things the hukou system impedes so many migrants from obtaining. The hukou system and its associated

institutional barriers to development remain unchallenged, in the opinion of academics³.

3. Methodology

The evaluation of the hukou system's effect on development requires multi-faceted approach along the lines of China's two relevant economic goals:

1. Rebalancing towards a domestic consumption economy, which can be measured by the change in consumption-to-GDP ratios and future projections.
2. The mitigation of the rural-urban socio-economic disparity, which can be determined through a rights-based approach of comparing economic, social and political opportunities.

The evaluation of the central research question are contingent to these two areas of analysis, from which a holistic conclusion can be drawn on the question of development.

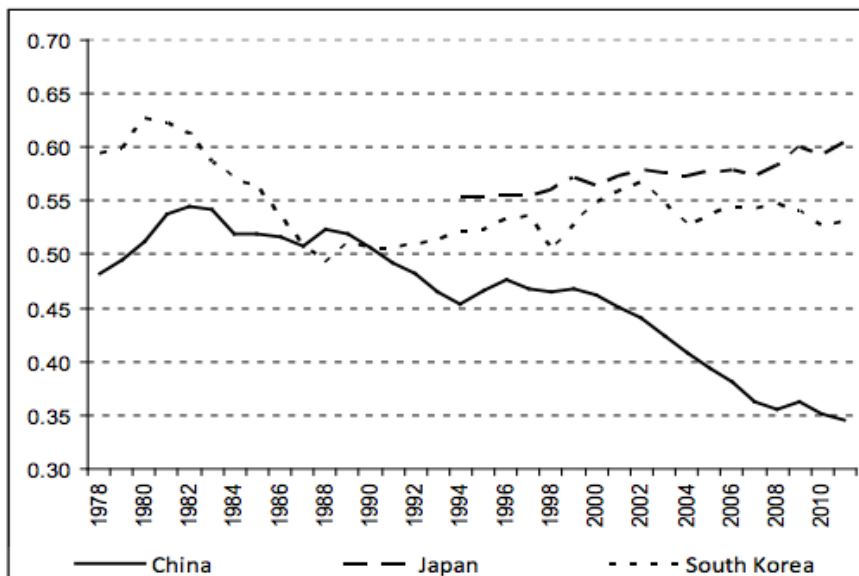
All selected data is relevant to the period from 2001-2015. Preference has been given to World Bank, OECD, or other information banks. Where Chinese government censuses have been used, care has been taken to account for the mismatch between registered (by hukou) residents and real residents. Where appropriate, the data has been adapted into suitable forms of original representation.

4. Findings & Discussion

To what extent is the hukou responsible for the low levels of household consumptions relative to the Chinese economy?

The first step is to establish that a pattern of low household consumption relative to GDP does in fact exist across the period 2001-2010. This can be done by comparing China's consumption to GDP ratio to that of similar economies⁴.

Figure 1 Private consumption in East Asian countries, share of GDP, %



Source: OECD Main Economic Indicators

⁴ Granted, finding similar economies is very difficult, especially when trying to control for other factors.

China's divergence from the other Asian economic powerhouses in terms of household consumption occurred in the late 1980's, at the time when China's market liberalization policies spurred a massive increase in investment and export growth, thereby propelling annual GDP growth into the double digits by the 2000's. Consumption did not keep up.

Notably, across the period 2001-2010, China experienced a sharp downturn in household consumption ratios, falling from 46.4% to 34.0% (see Appendix 1), intriguingly at the same precise period that China's urbanization rates took off:

Table 1a - Chinese Urban population (% of total population)

	Urban population as % of total
1990	26.41
1995	29.04
2000	36.22
2001	37.66
2002	39.09

Source: NBS (1996 and 2003); Chan and Hu (2003).

Urbanization, the increase of city populations relative to the population of the country, is usually associated with increased incomes due to increased job opportunities and increased retail sales due to the economies of scale of urban settlement. Instead, the imbalance between consumption and other areas of the economy continued to grow, even after the urban population share spiked between 1995 and 2002 by 10.05%.

Table 1b - Consumption-to-GDP ratio in 2013 of low-to-middle income countries

Country	China	India	Indonesia	Nigeria
Household consumption-to-GDP ratio in 2013 (%)	36%	59.2%	57.2%	72.1%

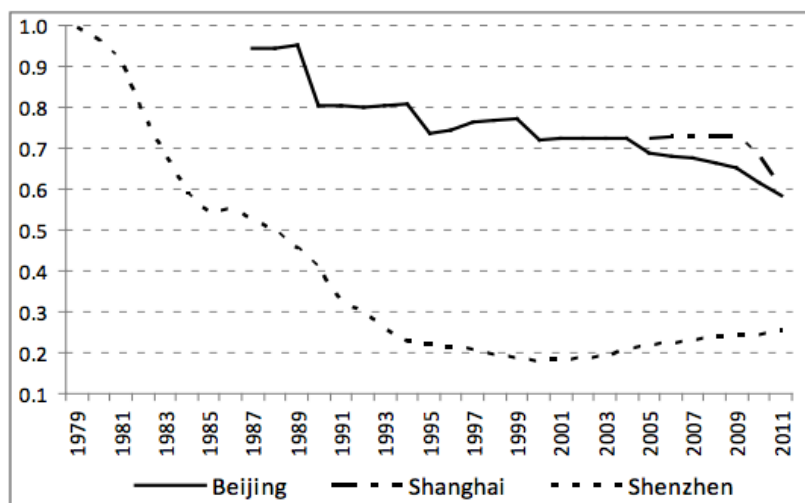
Source: World Bank

Even if we expand the comparison to include lower-to-middle income countries⁵ that are going through the same growing pains of rapid economic growth, be it industrial or otherwise, China is still significantly behind. Rather than reflecting a short-term boost in investment levels to aide an industrial growth (which has slowed in recent years), this uniquely large shortfall is a symptom of a grossly imbalanced economy.

The next step is to establish that if there is a shortfall in consumption, then the hukou registration system is in fact a distinguishable, significant independent factor causing it. For this, a segmentation of the population into urban residents with urban hukous and those with rural hukous is needed to demonstrate a significant difference of income *that cannot be attributed to other factors*, such as a generally weak social security system and tight financial regulations.

⁵ With similarly large populations and economies

Figure 2 Urban hukou holders as a share of urban population



Source: National Statistical Bureau of China

In China's three largest urban centers, Beijing, Shanghai and Shenzhen, a definitive decline in the share of correctly registered residents can be identified, ie. a rapid increase in the population of improperly registered migrants. There is a tangible shortfall between household expenditures and disposable income amongst rural residents, urban residents and rural-urban migrants:

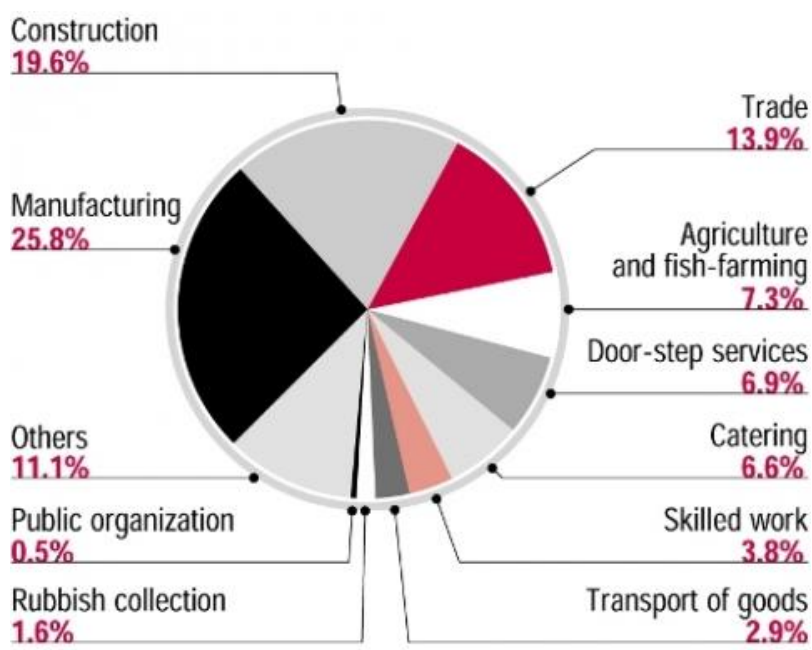
Table 2 - Mean income by residence and registration, 2013 (\$)

	Unadjusted	PPP
National	5,559	4,927
Urban	9,323	7,267
urban registered	9,996	7,773
urban migrant	6,083	4,831
Rural	3,143	3,425

Indeed, a significant gap between urban migrants and their fully registered counterparts can be seen, contributing to a large shortfall in

demand. A difference of \$3923 multiplied across between 200 and 300 million urban migrants with rural hukou's translates to a \$784 through \$1176 billion shortfall in incomes. At that low income level, it is likely that if these incomes were to grow to be rebalanced with urban hukou-holding residents, much of it would translate into consumption rather than savings. This is because low income households need necessities or normal goods that their current income doesn't fully cover (see Table 3b). Hence, this discrepancy between migrants and full urban residences becomes even more significant in contributing to an abnormally low total household consumption rate in China.

Furthermore, low-income migrants, owing to lower occupational mobility associated with their most common occupations, are employed in low growth jobs with little prospect of rising income, relative to native urban residents. In Shanghai alone, this phenomenon can be identified in the distribution of jobs amongst working migrants.

Figure 3 - Shanghai Migrant Worker Employment

Source: Statistics Bureau of Shanghai, *Shanghai shi 2000 nian renkou pucha ziliao* (Shanghai 2000 Population Census), Peking, China Statistics Press, 2002.

74.3% of migrant workers in Shanghai, the fastest growing city in China, are employed in short-term or unstable jobs, such as construction, transportation and low-skill services. Beyond ineligibility for state supports such as unemployment welfare and education, migrants are facing significant roadblocks in employment. This produces a depressing effect on the average propensity to consume of migrants, as uncertainty in the workplace complements uncertainty in their legal status. In fact, average migrant income as a percentage of median urban household income has been declining, from 68% in 2012 to 45% in 2007 (Dreger 2014). The nationwide trend complements local data in Shanghai - the declining share of prosperity

would suggest expected income increases are much lower for the 200-300 million migrants in Chinese cities.

Table 3a Income distribution and consumption ratios – Urban households

Quantile range	2002		2007	
	Income	APC	Income	APC
(Q0.00–Q0.10]	(0–10.2]	0.955 (0.009)	(0–20.0}	0.977 (0.011)
(Q0.10–Q0.25]	(10.2–14.4]	0.913 (0.009)	(20.0–30.0]	0.848 (0.013)
(Q0.25–Q0.50]	(14.4–20.9]	0.886 (0.006)	(30.0–45.2]	0.802 (0.011)
(Q0.50–Q0.75]	(20.9–30.1]	0.840 (0.007)	(45.2–71.6]	0.730 (0.010)
(Q0.75–Q0.90]	(30.1–42.0]	0.799 (0.008)	(71.6–109.2]	0.687 (0.011)
(Q0.90–Q1.00]	> 42.0	0.733 (0.014)	> 109.2	0.604 (0.014)

Table 3b Income distribution and consumption ratios – Migrant households

Quantile range	2002		2007	
	Income	APC	Income	APC
(Q0.00–Q0.10]	(0–6.7]	0.898 (0.022)	(0–10.2)	0.824 (0.029)
(Q0.10–Q0.25]	(6.7–9.6]	0.864 (0.010)	(10.2–14.4)	0.801 (0.016)
(Q0.25–Q0.50]	(9.6–14.3]	0.826 (0.008)	(14.4–20.40)	0.794 (0.011)
(Q0.50–Q0.75]	(14.3–21.1]	0.772 (0.008)	(20.4–32.4)	0.775 (0.011)
(Q0.75–Q0.90]	(21.1–29.4]	0.756 (0.013)	(32.4–48.0)	0.726 (0.010)
(Q0.90–Q1.00]	> 29.4	0.697 (0.019)	> 48.0	0.663 (0.011)

Source: Dreger, BOFIT. Samples from urban centers larger than 100,000 people. *Income stated in thousands of \$, **APC = Average propensity to consume, calculated as the share of household income spent on consumption

Table 3a and 3b, beyond clearly identifying a wide income disparity, also indicate a wide shortfall in consumption patterns between migrants and urban workers *earning the same income*. Indeed, in 2002, the average migrant household making between 14.4 and 21.3 thousand \$ had a marginal propensity to consume of 0.772, whereas the urban one had 0.886, for a shortfall of 11.4%. The same bracket in 2007 has a shortfall of 16.7%, indicating a worsening of factors leading to underconsumption by migrants.

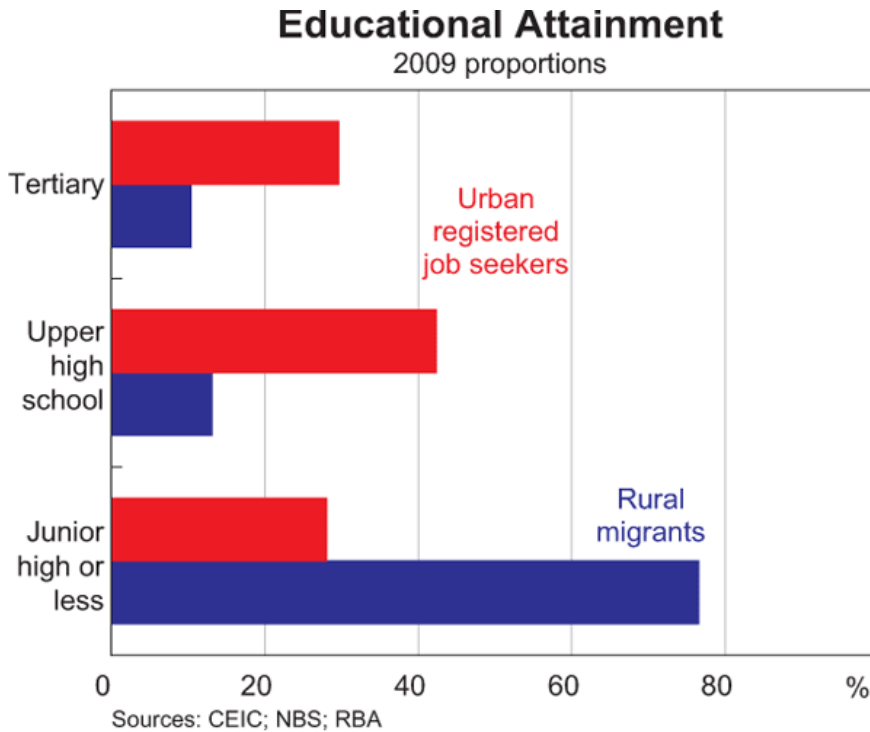
Above all, one trend is clear - while the interquartile range (25% to 75%) of urban residents saw a doubling in average incomes, migrants only saw a 50% increase, reaffirming that migrant status, due to economic and legal institutional barriers, are spending far less.

A society with a class of artificially economically disadvantaged is bound to suffer from economic imbalances, especially if that class is so large. The hukou system's crippling effects on the economic livelihood of urban migrants can be affirmatively determined to be a primary driving factor in underconsumption, especially given the rapid pace of urbanization between 2000 and 2015.

How has the Hukou system perpetuated urban-rural inequalities?

Urban-rural inequality can be understood through the lens of civil and economic rights. For the purposes of this analysis, "rights" are evaluated by the level of access to opportunities associated with those rights. Specifically, to determine if the urban-rural divide is perpetuated by the hukou system, an analysis at the generational translation of economic and civil inequalities from parent to child involving education and gini coefficient will be required.

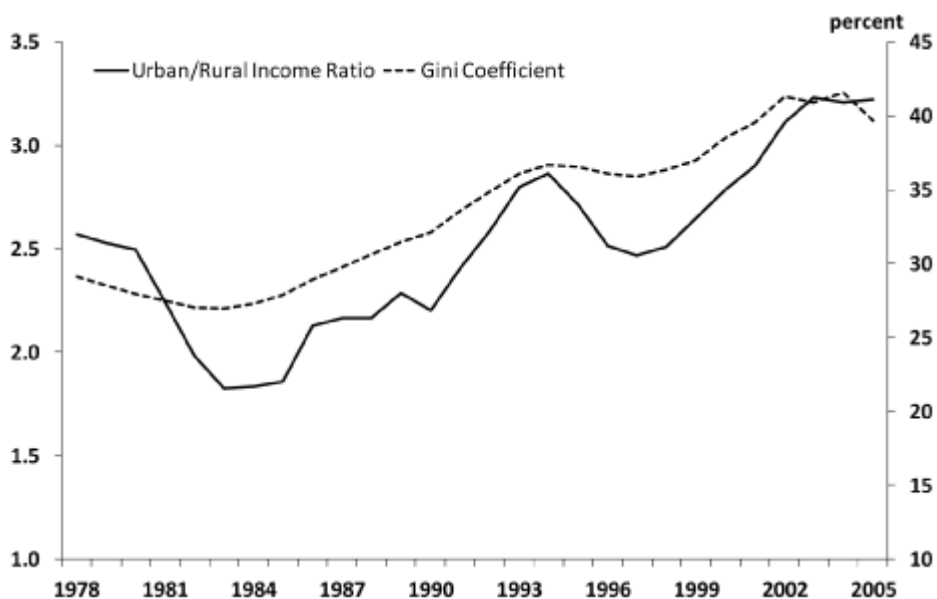
Figure 4 - Educational Attainment of job applicants in Chinese cities



76% of migrant urban dwellers have an educational attainment of Junior high school or less. The vast majority of rural migrants with low education can be divided into a few categories: they either a) migrated after an underdeveloped rural education, b) was not eligible for public education after migration to the cities or c) chose not to invest in a private education. In all these cases, the overbearing hukou system has a definitive role to play. The hukou system can be deemed partly responsible for a) - the stratification of urban and rural localities is caused by the entrapment of rural residents and the closing of the doors of self-interested urban provinces, although this could be solved with federal redistribution of infrastructure funding. However, in cases b) and c), the hukou system is the exclusive cause for

undereducation - cities like Beijing, Shanghai and Shenzhen deny migrant children access to public primary through tertiary schools. The growing correlation between education and income seals the generational fate of hukou system to be significantly disadvantaged:

Figure 5 - Gini coefficient and Urban/Rural Income Ratio in China 1978-2005



Indeed, the juxtaposition of the gini coefficient, an all-inclusive measure for inequality in a nation, and the income inequality of urban and rural households demonstrate that the shortfall in income, associated with economic rights and education access, is *the* defining issue of inequality in the nation. A 28% increase in the urban/rural income ratio correlates with a 20% increase in the gini coefficient. Such a strong correlation points to the scale of the hukou issue, in that it is a *class-defining* institution. The hukou system not only perpetuates inequality in China, it *is* inequality in China.

5. Conclusion

In conclusion, the hukou system has had a significant impact on China's stated development goals. The hukou system is a major institutional barrier to a large class of Chinese citizens. China's anomalously low household-consumption-to-GDP ratio can be attributed to an artificially low income of migrant workers to the tune of \$1 trillion. Furthermore, institutional barriers to state services and immobility in the job markets further shows a lower average propensity to consume than urban households. An analysis of education access and the gini coefficient indicates a strong correlation between the the hukou and social inequalities. The hukou system hence stands among the most pressing stumbling blocks alongside private debt and a weak social security system in the official development goals of China. In the words of Kam Wing Chan, though it may take 25 years and more than a trillion dollars to dismantle the hukou system and integrate disadvantaged migrants, the problem will not ease with any amount of time or economic growth. The Central government needs to place Hukou reform at the top of the agenda for economic policy.

6. Appendices

Table 1

Year	Household consumption as a Percentage of GDP (%)
2001	46.4
2002	44.8
2003	42.8
2004	41.4
2005	39.4
2006	36.7
2007	37.6
2008	36.6
2009	35.3
2010	36.6
2011	37.7
2012	36.6
2013	36.0

Table 2 - Education Expense of Urban and Rural Households (Unit: Yuan)

	Average	Median	Education/Income
Urban	7,808	2,273	10.88%
Rural	2,970	545	10.74%
Overall	5,758	1,182	10.84%

Table 3 - Education Expense of Households Based on Income (Unit: Yuan)

Income Quartile	Mean	Median	Average Income	Education/Income
Below 25%	3,176	364	856	371.02%
25%—50%	4,074	455	10,653	38.24%
50%—75%	4,320	1,545	29,853	14.47%
Above 75%	11,074	4,545	168,384	6.58%

7. Bibliography

- Chan, Kam Wing, Ta Liu, and Yunyan Yang. "Hukou and Non-hukou Migrations in China: Comparisons and Contrasts." *International Journal of Population Geography* 5 (1999): n. pag. Print.
- Dreger, Christian, and Yanqun Zhang. "Does the economic integration of China affect growth and inflation in industrial countries?." *Economic Modelling* 38 (2014): 184-189.
- He, Alice. "A tale of two classes: China's hukou-based discrimination – Part 2." *essa*. N.p., 20 Apr. 2014. Web. 25 July 2015.
<<http://economicstudents.com/2014/04/a-tale-of-two-classes-chinas-hukou-based-discrimination-part-2/>>.
- Huang Y, 2001, "Gender, *hukou*, and the occupational attainment of female migrants in China (1985 - 1990)" *Environment and Planning A* 33(2) 257 – 279
- Lu, Yao. "Education of Children Left Behind in Rural China." *Journal of marriage and the family* 74.2 (2012): 328–341. *PMC*. Web. 7 Aug. 2015.
- Marshall, John. "China: Urbanization and Hukou Reform." *The Diplomat*. N.p., 11 Oct. 2013. Web. 26 July 2015.
<<http://thediplomat.com/2013/10/china-urbanization-and-hukou-reform/>>.
- Montgomery, Jessica L. "The Inheritance of Inequality: Hukou and Related Barriers to Compulsory Education for China's Migrant Children." *Pacific Rim Law & Policy Journal* 21.3 (2013): 591-622. Print.

Pelikanova, K., and A. Melander. "Reform of the Hukou System: A Litmus Test of the New Leadership." *ECFIN Economic Brief* 26 (2013): n. pag. Print.

Wong, Edward. "Survey in China Shows Wide Income Gap." *The New York Times*. N.p., n.d. Web. 25 July 2015.
<http://www.nytimes.com/2013/07/20/world/asia/survey-in-china-shows-wide-income-gap.html?_r=0>.

Xiaogang Wu and Donald J. Treiman. "Inequality and Equality under Chinese Socialism: The Hukou System and Intergenerational Occupational Mobility *American Journal of Sociology*". Vol. 113, No. 2 (September 2007), pp. 415-445

Wu, Xiaogang. "Wage Discrimination in Urban China: How Hukou Status Affects Migrant Pay." *HKUST IEMS*. N.p., 23 Mar. 2015. Web. 11 July 2015. <<http://iems.ust.hk/thought-leadership-briefs/wage-discrimination-urban-china-hukou-status-affects-migrant-pay/>>.

Zhang, Huafeng, 2010. "The Hukou system's constraints on migrant workers' job mobility in Chinese cities," *China Economic Review*, Elsevier, vol. 21(1), pages 51-64, March 2010



**The Pursuit of Truth in Relation to the Faculties of Sensation,
Perception, and Judgment --
An Essay Written in the Context of the Meditations by Rene
Descartes**

Yuan Wang

Author background: Yuan grew up in China and currently attends The Affiliated High School of South China Normal University in Guangzhou. His Pioneer seminar program was titled “Descartes’ Meditations.”

**It should be noted that the intended audience of this piece is readers who are familiar with and have understood the basic arguments of the Meditations.*

Introduction

In this section, the author will introduce his topic of discussion, justify the validity of that topic, present an outline of the paper, and clarify the concepts important to the topic of discussion.

It is no coincidence that Descartes’ most well-known work, and one of the philosophical masterpieces of all time, *the Meditations*, devotes extensive attention to the pursuit of truth. The pursuit of truth has been arguably the fundamental task of philosophy. The word “philosophy,” originally a Greek word, means “love of wisdom,” and it is safe to say that the prerequisite of wisdom is truth. In *the Meditations*, many ideas and concepts are directly related to the pursuit of truth, but the focus of this

paper will be on sensory perception. On one hand, truth only resides in judgments, and any judgment is an extension of ideas (the “likeness” of something). Judgment without an idea is like a wave without water. On the other hand, ideas almost always come from the senses. Thus, sensory perception is an indispensable ocean to explore if one wants to surf the waves of truth.

It might be a concern to some readers that the author chooses the subject of sensory perception while Descartes believes in the reliability of innate ideas instead of sensory input. However, it appears that one of Descartes’s major aims in writing the *Meditations* is to prove that sensory perception is a reliable means of acquiring knowledge. This can be seen near the very end of the *Meditations*, where Descartes seeks to prove the validity of using sensory perception by arguing that when his body is in good function, sensation will not produce false information. Even in cases of ambiguity, he can use multiple senses, memory and / or intellect to examine the situation (89). What’s more, Descartes’s support of the validity of innate ideas will be rebutted in the next section as one of the claims being dissolved, which further justifies the validity of discussing sensory perception in relation to the pursuit of truth.

With regards to ideas, though, they themselves can never be equivalent to judgments, so there is necessarily another faculty involved in creating judgments from ideas, which is the faculty of judgment.

Thus, this paper aims to examine the nature of the faculties of sensation, perception and judgment and establish a model of the relationship between these faculties. It also aims to explore how those faculties are related to the pursuit of truth and to understand the source of error regarding the judgment

of truth. All the endeavors above will be attempted in the context of Descartes' ideas and arguments in *the Meditations*. In the end, the author will examine the faculty of judgment by comparing human beings with animals and exploring what drives man's pursuit of truth.

In this paper, truth will be defined as the validity of judgments. As Descartes has pointed out, ideas cannot be false, but judgments can (37). He also explains that, besides the formal falsity contained in judgments, material falsity consists in the fact that one represents "non-things as things" (43). However, the material falsity Descartes talks about is strictly a version of falsity of judgments. When a person represents a non-thing as a thing, the thing being represented cannot be false as it is merely an idea. The falsity lies in the fact that the person *judges* that representation to be the non-thing.

From a modern perspective, the concept of "sensory perception" includes two discrete components, sensation and perception. Sensation is, "the process by which our sensory receptors and nervous system receive and represent stimulus energies from our environment" (Myers 117). Perception is, "the process of organizing and interpreting sensory information, enabling us to recognize meaningful objects and events" (Myers 117). Perception belongs to the latent portion of the two-track mind, which means that perception happens without one's conscious control (Myers 488). The faculty of sensation receives and represents raw data taken directly from the environment, while the faculty of perception automatically processes that raw data outside the realm of consciousness.

Some may argue that truth and falsity also lie in sensations and perceptions, which are not ideas. It is to be emphasized that the author does not attempt to limit truth and falsity to the scope of judgments. However, the

essay is only interested in the truth and falsity of judgments, and *the truth and falsity addressed in this essay refer to those of judgments alone.*

The faculty of judgment is, in short, the faculty that makes judgments. The faculty of judgment is part of the *conscious* mind, since it is evident that any judgment to be made has to be consciously created by our minds. By emphasizing the modifier “conscious,” the author aims to differentiate the essence of the faculty of judgment from the *latent* mind. The latent mind is the part of the mind that automatically processes information *outside* the realm of consciousness (Myers 90). For example, when a person sees a dog, he or she doesn’t (and can’t) consciously *judge* that it is a dog based on his or her vision of it; instead, that person’s latent mind automatically discerns it as a dog.

Groundwork

This section aims to explore the existence of the material world and to dissolve three major claims made by Descartes: that God exists as a perfect being, that distinct judgments are no doubt true, and that the mind can exist without a body. By dissolving the first two claims, the author seeks to weaken Descartes’s support of rationalism and justify the author’s discussion of sensory perception and use of empiricism in this essay.

As proved before, truth lies only in judgments. Most judgments people make are based on information about the material world, so it is necessary to examine whether this material world exists or not. Descartes’s proof of the existence of the material world is based on his proof of the existence of God. He argues that, even after employing all the faculties God has provided, he finds no way to discern whether the material world exists or not, but there is

a strong tendency for him to believe that it exists; thus the material world must exist, otherwise God would be deceiving him (Descartes 79-80). However, as the author is soon going to illustrate, Descartes's proof of God is invalid, so the arguments based on this proof are invalid as well.

Imagine if the author is someone, or something, one understands as a brain sitting in a tank. The tank is full of chemicals and electronic devices that simulate the parts of this brain so that the author can experience what is supposed to be the "world" around. All the things in that world will then be merely mental figments. The author has yet discovered no evidence to dissolve the possibility of this hypothetical scenario, so it is indiscernible whether the material world, as the author understands it, exists or not.

Despite this uncertainty, the author is going to assume that the material world does exist within the context of this paper. On one hand, if it doesn't exist, any proofs that the author attempts to undertake are meaningless. It would be like a cartoon character trying to understand the rules of his or her world, which is created and possibly altered from time to time at the whim of forces that he or she cannot understand. On the other hand, if the material world does exist, the author should try to discern the truths about this world. Thus, at this stage, it is the only appropriate choice to assume the world exists and get on with the mission.

After exploring the existence of the material world, the author will attempt to repudiate three major claims made by Descartes. These claims are either themselves directly in opposition to some of the author's claims, or they support claims that are directly against the author's claims. Thus these claims need to be dissolved to lay the groundwork for this paper.

The first two of these claims will be examined together because of their interdependent nature: the claims that God exists as a perfect being and that “distinct and clear” perceptions are true. The interdependent nature is the flaw for both claims. One on hand, Descartes attempts to prove that clear and distinct perceptions are true by arguing that God is perfect. A perfect God will not bestow on men a faculty of judgment that can go wrong when employed correctly (Descartes 54). One knows the faculty of judgment is correctly employed when his or her perception is clear and distinct (58-59). One the other hand, Descartes attempts to prove the existence of God by employing the truth of “clear and distinct” perceptions. For example, Descartes claims, “it is manifested by the natural light [clear and distinct] that there must be as much <reality> in the efficient and total cause as in the effect of that cause” (40). This claim is one of the key claims that support his proof of God. Descartes has employed the logical fallacy of circular reasoning, which dissolves his proofs of both the existence of God and the validity of “clear and distinct” judgments.

By attacking these two claims, the author attempts to weaken the foundation of what is to become rationalism, for one of the sources of knowledge in rationalism is innate ideas, which are “clear and distinct” judgments. Even though Descartes has put sensory input into doubt, it has been illustrated that innate ideas or “clear and distinct” judgments are dubitable as well, which undermines his rationalistic assumptions. Thus, the author’s choice to apply empiricism instead of rationalism is justifiable. Similar to discerning whether the material world exists or not, it is impossible to discern whether sensory perceptions, assuming that they generally yield the truth, are the best method of inquiry.

The third claim by Descartes that needs to be refuted is that the mind can exist without the body. He claims that the mind is distinct from the body and, “can exist without it” (Descartes 78). The mind is inseparable from the brain, which is clearly a part of anyone’s body. The study of neuroscience has shown that activities of the mind correspond with activities of the neurons and chemicals of the nervous system, which are evidently parts of the human body, without which the mind cannot exist. However, it is noteworthy that Descartes does judge that the mind and body are intimately connected. He claims that the two are “intermingled” (Descartes 81).

Sensation, Perception and Judgment

This section attempts to elaborate on the nature of the faculties of sensation, perception, and judgment, and to examine the relationship between these faculties.

Subsection One: Nature of the Faculties of Sensation and Perception

This sub-section aims to show that Descartes’ idea of “sensation” (or “sensory perception,” since Descartes has not made a distinction between the two) includes the faculties of both sensation and perception in some cases. It also aims to establish that the nature of sensation and perception is to preserve the well-being of the body.

In the second meditation, Descartes identifies only the faculties of sensation and judgment. He illustrates this identification when he says that his sensation enables him to see hats and coats, while his judgment enables

him to reach conclusion that he sees men (Descartes 32). The nature of sensation and perception (the modern definition) have been briefly addressed in the introductory section of this essay, but will now be discussed in relation to Descartes's arguments and elaborated upon. Descartes mistakenly attributes work done by the faculty of perception to the faculty of sensation in some situations. An example of this is Descartes's understanding of the perception of color.

Descartes makes the mistake of attributing the perception of color to the faculty of sensation. One of the examples that illustrates this mistake is when he claims that men have the habit of judging that the features they *sense* from physical entities are present in their essence, one of those features being color (82). In fact, the faculty of sensation merely represents the reflected light waves it receives from the physical entity by forming a pattern of stimulated sensors in the retina. The information contained in that pattern is then processed into information on features like color and shape by the faculty of perception.

Descartes argues that sensations are given or taught to humans by nature (81) and that nature is embodied in man as the combination of mind and body (Descartes 88). This nature enables sensation to inform the mind-body composite on what is harmful and what is beneficial, in order to maintain its well-being (Descartes 83). Four points regarding these claims need to be made. First, preserving the well-being of the "mind-body communion" is preserving the well-being of the body, because as argued in the second section of the paper, the mind is inseparable from, and dependent on, the body. Second, the "sensation" in these claims should be understood as the modern concept of both "sensation" and "perception." Without the existence

of either, neither of the two faculties alone can contribute to the well-being of the body. Third, it is the information produced by the faculty of perception that ultimately benefits the well-being of the body. Last but not least, the author is not going to argue that it is “nature” which has “given” men the faculties of sensation for a particular purpose. Thus, to twist Descartes’s claim slightly, the author argues that the information produced by the faculty of perception helps preserve the well-being of the body.

Subsection Two: Nature of the Faculty of Judgment

This subsection aims to prove that the nature of the faculty of judgment is expandable, finite and dependent on the state of the brain.

Above all, the faculty of judgment is the *sole* faculty discerning the difference between truth and falsity. Descartes claims that the faculty of judgment is finite and flawless when used under the appropriate scope in an appropriate fashion because it is given by a perfect God (53-54). It can be inferred that Descartes sees the faculty of judgment as *static* and *rigid*. However, if the faculty of judgment has a fixed capacity, it is hardly possible for the intellectual capacity of humans to expand as it has in the past. Some may argue that the expansion is not related to the prowess of the ability to judge but only to the facts accumulated in human knowledge. Rather, both forces are at work. It is evident that intellectuals from the past and those from the present think in different ways, not only because of the different amount of facts they know, but also because of the different *mindsets* and *perspectives* they possess. A more fluid mindset and new perspectives expand the human capacity for judgment. Thus, the faculty of judgment is *expandable* instead of static.

Even though the faculty of judgment is expandable, Descartes is right that it is finite. However, humankind is far from reaching or even understanding the limit of its faculty of judgment. Another feature of the faculty of judgment is its dependence on the physical state of the brain, which Descartes fails to recognize. He claims that mind and body are completely separable, as shown in the second section of this paper. However, when a man is having a fever, not only is his body malfunctioning, but his ability to use proper judgment is also impaired. When the brain is in a bad condition, the faculty of judgment is negatively affected. In conclusion, the faculty of judgment is expandable, finite, and dependent on the state of the brain.

Subsection Three: Relationship Between the Three Faculties

This subsection aims to establish that the faculty of judgment makes judgments based on the information produced by the faculty of perception and that the faculty of judgment is able to *understand* the faculties of perception and sensation.

How then, are the faculties of sensation and perception related to judgment? As argued by Descartes, ideas are, in essence, the “likeness” or “image” of things, whereas judgments contain something more (37). This implies that the “likeness” of things is an indispensable part of any judgment. The likeness of things comes from the faculty of perception, which is in turn dependent on the faculty of sensation.

Since every judgment contains ideas, the faculty of judgment is largely dependent on the faculties of perception and sensation, which provide input for the basis of any judgment. However, the faculty of judgment is not only a

passive recipient of processed information. This faculty, as illustrated by the fact that the reader is reading this paper, is able to examine and understand the faculties of sensation and perception and their influence. This understanding is a testament not only to the ability of the faculty of judgment to transcend its limitations but also to its ability to avoid errors when the information it receives is flawed.

The Source of Error

This section of the paper aims to elaborate on the source of error and explore how to avoid it.

Subsection One: the Author's Theory of the Source of Error

This subsection attempts to explore the imperfections of the faculties of sensation, perception and judgment and prove that the imperfection of the faculty of judgment is the ultimate source of error.

From the previous section it can be seen that the faculties of sensation, perception and judgment are all involved in the pursuit of truth. These three faculties are all imperfect but the decision regarding truth or falsity ultimately lies in the faculty of judgment. Thus, the nature of all errors or falsities of judgments is the imperfection of the faculty of judgment.

Indeed the imperfection of the faculty of sensation *seems* to yield falsities. That imperfection lies in the misrepresentation of the environment in the data created. This happens because the system of representing stimuli from the outside world is not perfect. For example, when a rod is half immersed in water, it looks as if the rod is bent. The faculty of sensation is

unable to compensate for distortions made while light is traveling from one substance to another, which ultimately leads to the creation of a distorted image. However, it is the responsibility of the faculty of judgment to decide on the kind of judgment it wants to make based on that image.

The case is similar for the imperfections of the faculty of perception. Those imperfections lie in the failure to produce a valid interpretation of the information provided by the faculty of sensation. For example, human beings have the tendency to see faces where there are none: two dots aligned with an arc is enough to enable an observer to “see” a face. The data provided by the faculty of sensation is accurate; it is the faculty of perception that generates information (in this case, a face) that should not be interpreted from the presented data. Even though the information that “it’s a face” is, to a certain extent, misleading, the faculty of judgment has the ability and responsibility to decide whether, and what, to judge.

Thus, errors are made when the faculty of judgment is drawing inappropriate conclusions based on the information it receives. The failure to make appropriate judgments is caused by several different factors. For one thing, the human capacity to analyze and reason is finite and prone to flaws. Also, the size of the human knowledge base affects the validity of judgments being made. Lastly, the prowess of the faculty of judgment is dependent on the physical condition of the brain.

*Subsection Two: Descartes’ Theory of Sources of Error in
Relation to the Author’s Theory*

In this subsection, the author aims to prove that the sources of errors proposed by Descartes are either nonexistent or should be attributed to the imperfection of the faculty of judgment.

Descartes believes that there are three sources of error: the imperfection of the faculty of sensory perception, the nature of the mind-body communion, and the mind's habit of making certain "ill-considered" judgments. These will be examined in relation to the sole source of error the author proposes.

The first error suggested by Descartes is the imperfection of the faculty of "sensory perception." Descartes gives the example of a man tricked by the pleasant taste of something poisonous and argues that this is an illustration of the imperfection of man's nature, which in turn gives men an imperfect faculty of sensory perception (83-84). However, as mentioned previously, Descartes' "faculty of sensory perception" is a mixture of the faculty of sensation and some instances of the faculty of perception, both of which have been addressed by the author in the previous subsection. In the end, those imperfections are not the ultimate source of error since it is the faculty of judgment that makes any attempt to discern truth from falsity. The man in Descartes' example should be aware that his sense of taste doesn't provide sufficient evidence for him to make judgments about whether things are poisonous or not.

The second source of error mentioned by Descartes is the nature of the mind-body communion. Descartes argues that only the brain is directly connected to the mind, yet the mind has control over the whole body, so other parts of the body need to be connected to the brain; the nature of this system of connections yields errors, since changes in different parts of the

body can trigger the same response in the mind (86-87). If this system of connections is real, it is an example of the imperfection of the faculty of perception. The error involves nothing consciously decided or discerned by the faculty of judgment; the error is also irrelevant to the data provided by the faculty of sensation because the data is correctly obtained. Thus, this error is related to the imperfection of the faculty of perception, leading back to the responsibility of the faculty of judgment.

The final source of error identified by Descartes is the habit of making “ill-considered” judgments. For this case, the author will attempt to argue that this habit doesn’t exist. Some of Descartes’s examples include judging that spaces where nothing stimulates the senses are empty, that a selfsame whiteness is present in white objects as they are perceived, and that objects in the distance have the same size as presented to the observer’s vision (82). These “ill-considered” judgments are made about the *innate* qualities of objects based on the information produced by the faculty of perception.

This argument is problematic because Descartes is making a claim (without providing any support) that *other* men have this habit of making “ill-considered” judgments whereas he apparently doesn’t have it himself. The question, then, is how he knows that others are *really* making these “ill-considered” judgments?

Imagine this scenario. When a man realizes that the essence of material entities lies not in line with the perceptions of them, he starts to examine how most people understand the essence of physical entities. His logic is as follows: those who have not had this realization must believe that the nature of physical entities *does* lie in line with the perceptions of them. This logic is

flawed because the negation of believing in something is not necessarily believing in something opposite; it can also be not having a belief at all.

In everyday life, people do not need to know what the “essence” of a chair is, and, just by recognizing vaguely that it “is” green is enough for any purpose besides philosophical inquiry. Most people don’t attempt to grasp what the essence of something is because there is no tendency (or “habit”) for them to do so, and thus they will make no judgments regarding this issue unless they are asked to.

After examining the theoretical part of the author’s refutation, it is beneficial to look back at the examples Descartes offers. First, the assumption that people judge that nothing is in spaces from which the senses receive no input is false because when a person looks into the dark, it certainly doesn’t appear to him that there is *nothing* there. Second, the judgment that a selfsame whiteness is in a white object is generally nonexistent, as illustrated by the example above. Last but not least, no observer judges the actual size of something from a distance; one only judges that it is distant. Thus, the so called habit of making ill-considered judgments is in fact nonexistent and the sole source of error is the imperfection of the faculty of judgment.

Subsection Three: Avoiding Error

This subsection aims to prove that the only way to avoid error is to employ the faculty of judgment in a correct way.

It needs to be clarified that not all errors can be avoided. It is only possible to try one’s best to avoid as many errors as possible by employing the right method. Since the sole source of error is the imperfection of the

faculty of judgment, it follows logically that employing the faculty of judgment in a correct way is the sole solution to avoiding any error.

To avoid errors related to imperfections of the faculty of sensation, one should make careful judgments based on comprehensive sources of information. Those judgments need to be made after understanding the potential flaws in each source of information. For example, when a man touches a stick half immersed in water while looking at it, he gets two conflicting pieces of information regarding the shape of the stick. However, the faculty of judgment should be aware that in this case, vision is prone to flaw, so it can decide that the information obtained from the sense of touch is more reliable. If a man is having hallucinations and illusions, his faculty of judgment should refrain from drawing any conclusion based on the senses of vision and sound because it recognizes that the information provided is prone to flaw.

To avoid errors related to imperfections of the faculty of perception, the faculty of judgment has a different task. These errors arise from failures of the interpretation of the information provided by the faculty of sensation. Even if that information is accurate, the faculty of perception can produce flawed interpretations. Thus, the faculty of sensation should obtain data that represents different angles of the physical entity, creating multiple interpretations. The faculty of judgment should decide which interpretation to rely upon. Again, when no such decisions can be made because all interpretations are significantly confusing, refraining from judgment is the preferred course.

Conclusion & Extension

This section endeavors to give a summary of the arguments presented so far, to examine the faculty of judgment by comparing human beings with animals, and to explore what drives man's pursuit of truth as an extension of the main arguments of this paper.

Subsection One: Summary

The faculty of sensation provides raw data collected from the environment. The raw data is transferred to the faculty of perception which is part of the latent mind. The faculty of perception processes that raw data and creates new information. That new information is then transferred to the faculty of judgment which is part of the conscious mind. The faculty of judgment is able to make decisions between truth and falsity based on that information. The faculty of judgment is able to examine both the faculty of sensation and the faculty of perception to understand their workings and their influence. The faculty of judgment is expandable, finite and dependent on the physical state of the brain. All three faculties have their imperfections, but the only faculty directly related to the pursuit of truth is the faculty of judgment and thus its imperfection is the only source of error. The only way to avoid these errors is for the faculty of judgment to understand their nature and make careful decisions based on that understanding.

Subsection Two: the Faculty of Judgment Examined by Comparing Human Beings with Animals

Humans, as of yet, have been unable to discern whether animals possess a faculty of judgment but it seems that comparatively, they have very limited

ability to discern truth from falsity. One might argue that a deer needs to judge that the cheetah twenty yards away wants to eat it and that the greenish strands on the ground are edible grass. However, these “judgments” are not necessary to make and probably not made at all. Indeed, the reaction to predators is both instinctual and sub-conscious; when a tiger appears, even a person doesn’t have the opportunity to judge that it is dangerous before he or she jumps up and runs away as fast as he or she can. Similarly, the idea that something is edible and conducive to health doesn’t *need* to be “thought of;” merely the sight and smell of food drives the animal towards eating it. The conclusion is that to simply survive or preserve the well-being of one’s body as an animal, the faculty of judgment is unnecessary. What is thus also true is that the faculty of perception can directly influence the behavior of animals, including humans.

Then what is the impact of the faculty of judgment on humans and their society? It seems that the faculty of judgment enables human beings to *understand* themselves and their environment, for, in essence, to understand is to judge how and why something works or exists. Such understanding gives them the ability to alter the environment to their liking. On the other hand, most other animals passively adapt to the environment, resorting to instinctual reactions to stimuli or basic learning like, for example, observational learning.

Subsection Three: Extension

Men have glorified the pursuit of truth. Man’s history has been marked by significant discoveries of truth. Some people have devoted entire their lives to pursuing truth. Some people have died for it.

Why, then, are men so obsessed with truth?

The author believes that, to a certain extent, throughout history, a *certain number* of men possess a tendency to seek truth, just as all men seek food and water and shelter. A person who has devoted his or her entire life to the pursuit of truth is really no different from an animal which has devoted its entire life to the pursuit of food, shelter and reproduction.

It is to be noted that the author does not attempt to prove this natural tendency to be the *sole* force driving humans' pursuit of truth, although it is a critical one. Some may argue that human beings pursue truth because such pursuits give them better means to solve their problems. For example, the creation of the Newtonian Paradigm took humans closer to truth. The mechanical laws it brought have made possible a myriad of technologies that improve our problem-solving skills, especially those of the Industrial Revolution. However, the difference between cause and effect in this case should be differentiated. The Newtonian Paradigm did lead to technological advances but it is highly questionable whether Newton was thinking about the technological improvements his discoveries might bring when he was working on the paradigm. Instead, the force that drove his endeavors was more likely the desire to discover the mechanical laws of the universe.

Just as the awareness that one is driven to food by instinct doesn't undermine that instinct, the awareness of man's natural desire for the truth does no harm to the desire itself. That awareness, however, does offer valuable insights into the nature of humankind, which is, in essence, another endeavor to pursue truth.

Reference

Descartes, Rene, and John Cottingham. *Meditations on First Philosophy*.

Rev. ed. New York: Cambridge UP, 1996. Print.

Myers, D. G. (2011). *Myers' Psychology For AP**. New York, NY: Worth

Publishers.



Biology (Macromolecular Machines)
**A Summary of Bacterial Secretion Systems (with Special
Attention to the Type III Secretion System)**

Timothy Wee

Author background: Timothy grew up in Singapore and currently attends Shanghai Community International School in Shanghai. His Pioneer seminar program was titled "Introduction to Cells and Intracellular Machines."

ABSTRACT

Bacteria have evolved a variety of macromolecular machines to aid them in the invasion and colonisation of a potential host. The ability to export bacterial effector proteins and virulence factors, or even directly translocate into a target cell, is crucial in the bacterium's struggle to outcompete other bacteria. There have been many discoveries regarding the specific structures and assembly of the secretion systems, including the exact proteins involved in subunits of specific structures. Although there has been immense progress made in the quest to understand and elucidate the exact mechanisms behind the various secretion systems, there remains a significant gap in scientific knowledge regarding specific mechanisms. As such, improvements in technological equipment and developments in our understanding of secretion systems continue to fuel further research, revealing new possibilities in practical applicability of future findings.

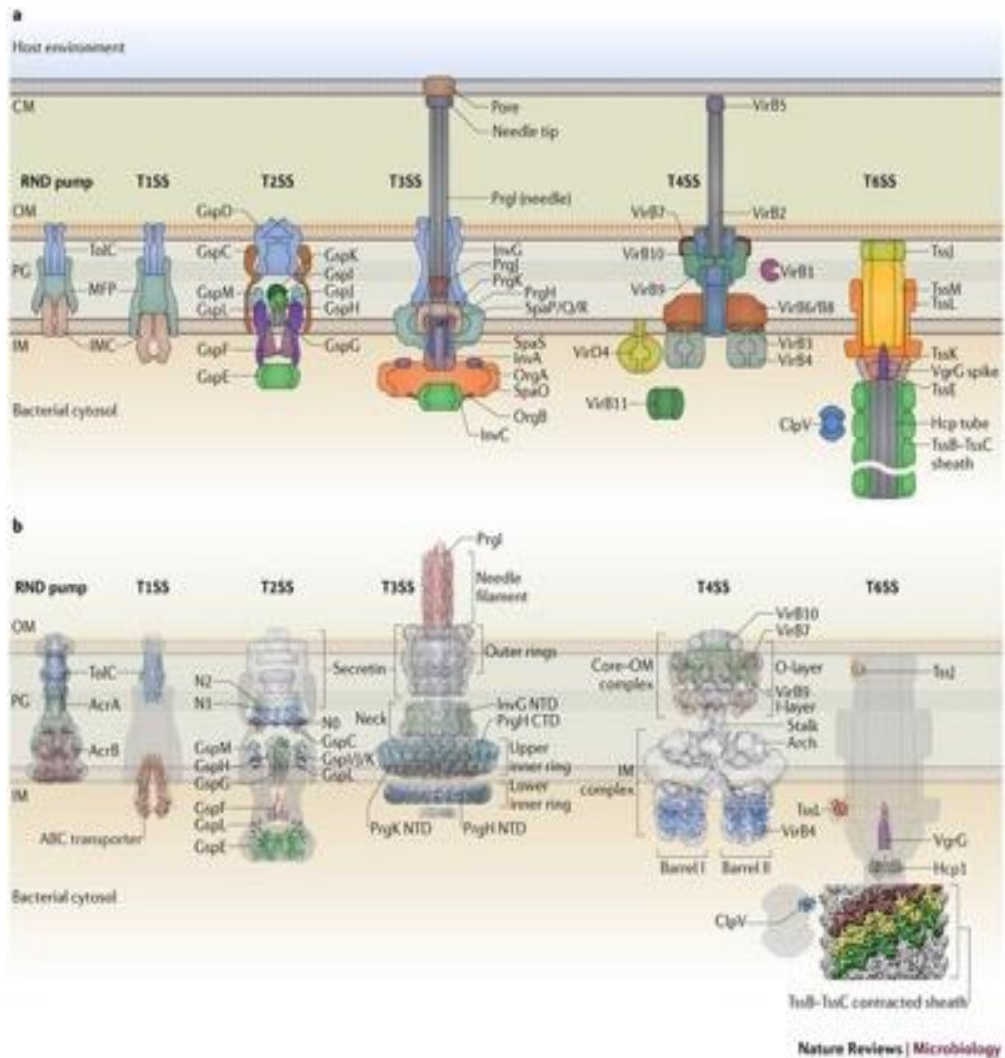
INTRODUCTION

Secretion Systems

Secretion is the process of releasing chemical substances from a cell or gland that often serve specific functions upon interaction. Unlike the products of excretion, secreted chemical substances are rarely waste products, and are often compounds (such as toxins) that have high metabolic costs. The active nature of secretion is reflected in the highly complex protein structures that have evolved at the cellular level to accomplish this process, focusing on an efficient delivery of substances. In particular, bacterial species have evolved intricate secretion systems to carry out a variety of tasks, ranging from the delivery of toxins to a target cell, to the horizontal exchange of genetic material. This is a vital process in a bacterium's response to its environment and ultimately its survival. In Gram-negative bacteria, secretion systems mainly aid in the export of virulence factors (in pathogenic species) and bacterial conjugation (Pugsley, 1993).

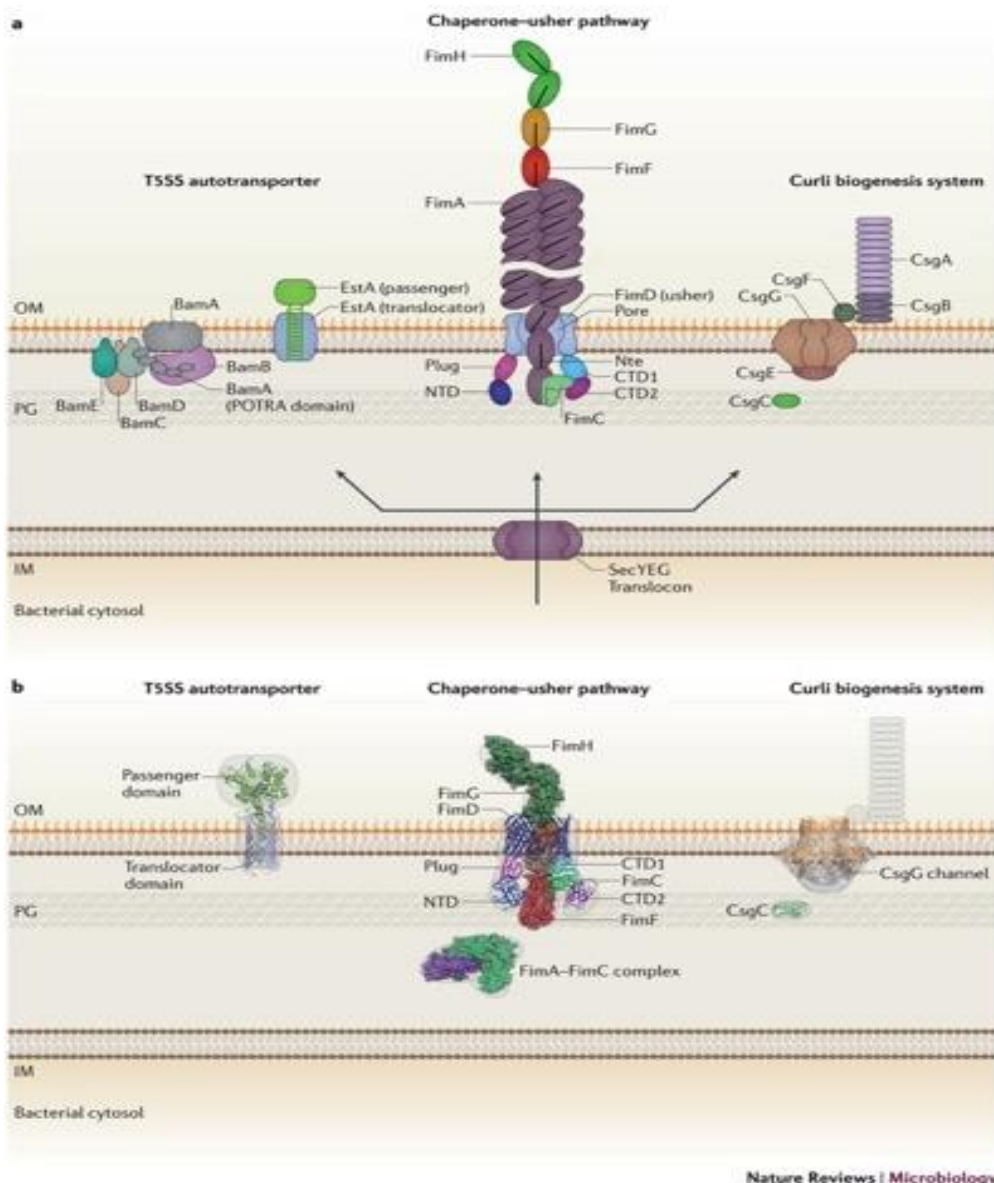
In Gram-negative bacteria, these systems can be dichotomised into two main categories based on whether they span the inner membrane (IM) and outer membrane (OM), or if they span the OM only. Systems that span both the IM and the OM include the Types I, II, III, IV and VI secretion systems (Fig. 1), while the Type V system spans the OM only (Fig. 2) (Costa, 2015 [Gerlach, 2007]).

Figure 1: Structure of double membrane spanning systems in Gram-negative bacteria. (Reproduced from Costa, 2015)



Putative locations of individual protein components are shown schematically in part **a**, and the solved structures are shown in part **b**.

Figure 2: Structure of outer membrane spanning systems in Gram-negative bacteria. (Reproduced from Costa, 2015)



Putative locations of individual protein components are shown schematically in part **a**, and the solved structures are shown in part **b**.

Secretion systems can also be dichotomised according to the number of steps in their secretion mechanism. Types I, III, IV and VI utilise a one-step mechanism where substances are transported directly from the bacterial cytoplasm to the extracellular space or into a target cell. Types II and V utilise a two-step mechanism where substances are first transported into the periplasmic space (via IM-spanning transporters), before transportation to the OM or secretion into the extracellular space (via OM-spanning systems) (Thamh H. Pham, 2012). The main difference lies in the fact that systems utilising the one-step mechanism exclusively secrete unfolded or partially folded substances, whereas systems utilising the two-step mechanism can secrete partially folded and folded substances (Thamh H. Pham, 2012).

Generally, although the individual types of secretion systems, and even the same types of secretion systems found in different species of Gram-negative bacteria, may contain similar or homologous structures, they are unique and highly developed macromolecular complexes that reflect an arms race in bacterial evolution. The degree of diversity in secretion systems is a testimony to the intense competition between bacterial species, and is indicative of the effects evolutionary pressures can have on the development of macromolecular complexes in unicellular organisms. Conversely, despite the highly varied nature of secretion systems, the repeated use of basic proteins and structural designs to achieve similar purposes is clear evidence of sophistication and effectiveness (Costa, 2015).

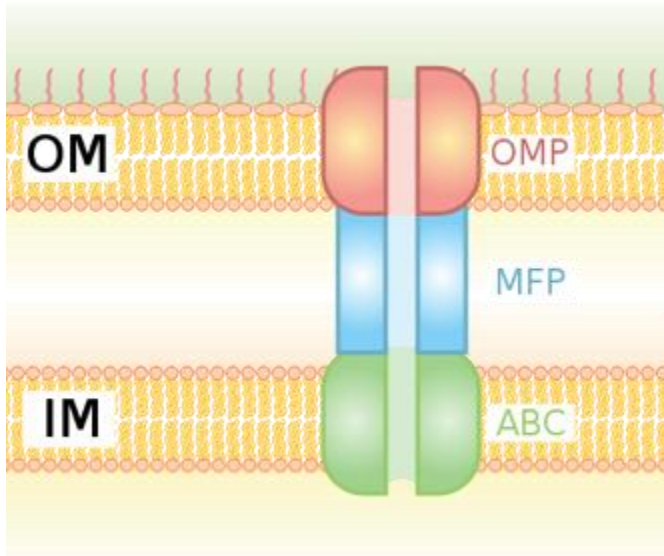
Type I Secretion System

The Type I secretion system (T1SS) in Gram-negative bacteria is involved in the secretion of a wide variety of protein substances from the bacterial

cytoplasm into the extracellular space (Costa, 2015 [Kanonenberg 2013]). In pathogenic Gram-negative bacteria, these proteins are often involved in the acquisition of nutrients and virulence. Structurally, the T1SS is closely related to the resistance-nodulation-division (RND) family of efflux pumps, which secrete small exogenous molecules. These are often antibacterial compounds, which contribute significantly to antibiotic resistance mechanisms. The system forms a tripartite double membrane-spanning channel with an IM component (IMC), a membrane fusion protein (MFP) and an OM channel (OMC). The system forms a thin drumstick-shaped continuous channel that spans the double membrane (Costa, 2015 [Kanonenberg 2013]).

The IMC utilizes energy generated from adenosine triphosphate (ATP) hydrolysis to transport any substances bound to itself from within the bacterial cytoplasm across the IM into the periplasmic cavity of the MFP. This new IMC-MFP complex associates with the OMC, triggering the opening of the channel pore and subsequently allowing the substance to exit the cell into the extracellular space (Costa, 2015 [S. Yum et al, 2009]).

Figure 3.1: Artistic representation of the structure of a Type I secretion system. (Reproduced from Wikipedia's original work).



The three main structures of the T1SS are clearly visible: the IM component (an ATP-binding cassette transporter protein); the fusion pathway between the IM and the OM (MFP); and the OM protein channel, the final exit for a substrate (OMP).

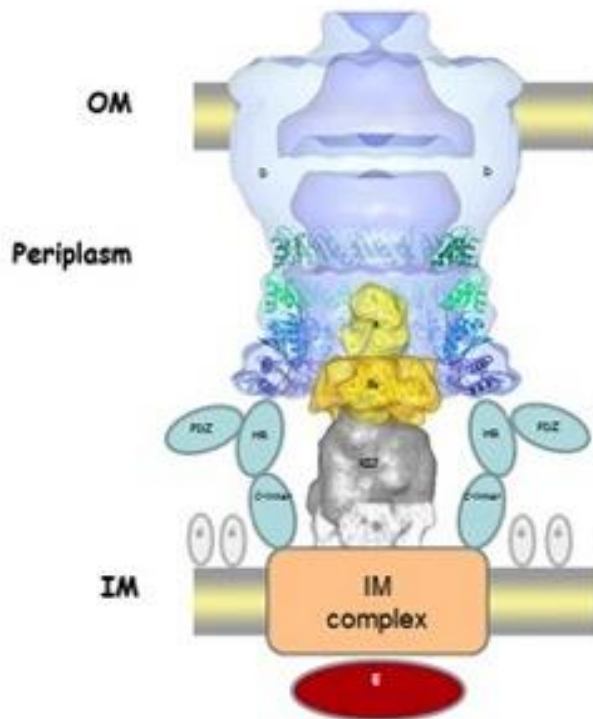
Type II Secretion System

Type II secretion systems (T2SS) are found in both pathogenic and non-pathogenic bacteria, and help to secrete folded proteins from the periplasm into the extracellular space (Costa, 2015 [M. Nivaskumar & O. Francetic, 2014]). These proteins include hydrolysing enzymes (such as pseudolysin of *Pseudomonas aeruginosa*), as well as toxins (such as the cholera toxin of *Vibrio cholerae*). While the structure of the entire T2SS has yet to be determined by research, some of the component protein structures are well documented both structurally and biochemically. These include the four parts of the T2SS: an OM complex, a periplasmic pseudopilus, an IM platform and a cytoplasmic ATPase (Costa, 2015 [M. Nivaskumar & O. Francetic, 2014]).

Unfolded polypeptides within the cytoplasm are transported into the periplasm by a SecYEG translocon, while folded proteins are transported by the Tat transporter. The periplasmic pilus (whose formation is also powered by ATP hydrolysis in a fashion similar to bacterial pili) utilizes energy from ATP hydrolysis to move substances through the OM channel and out into the extracellular space (Costa, 2015 [K. V. Korotkov, M. Sandkvist & W. G. Hol, 2012]). The ATP hydrolysis allows for the extension of the pseudopilus, which physically pushes substances through and out of the double membrane. The pseudopilus does not extend beyond the double membrane. The exact mechanisms between the interaction of substance and opening of the OM channel remain unclear (Costa, 2015 [K. V. Korotkov, M. Sandkvist & W. G. Hol, 2012]).

Figure 3.2: Graphic representation of the structure of a Type II secretion system (Reproduced from Professor Wim Hol, University of Washington).

The Type II Secretion System



Wim Hol

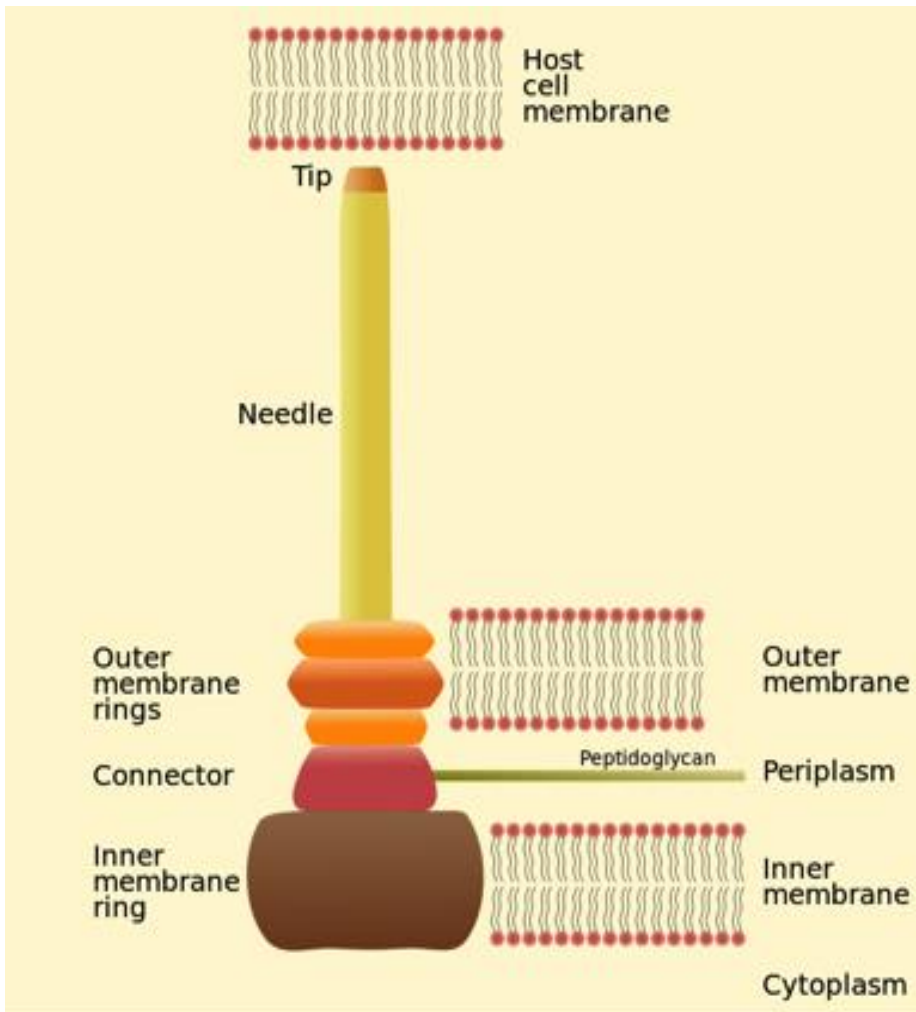
The four main structures of the T2SS are clearly visible: the cytoplasmic ATPase at the edge of the IM; the IM complex in the IM; the periplasmic pseudopilus mid-extension in the periplasmic cavity; and the OM exit channel crossing the OM into extracellular space. The object in yellow represents a cholera toxin molecule, in the process of being secreted by the T2SS.

Type III Secretion System

Type III secretion systems (T3SS), also known as injectosomes, specialise in the targeted delivery of proteins and chemical substances, often directly into or into the surroundings of the target cell (Costa, 2015 [G. R. Cornelis, 2006]). The effective delivery of bacterial effector proteins into target cells promotes invasion and colonisation, and accounts for the complexity of the T3SS. A T3SS includes two main components: a double membrane base containing stacked ring structures (also contains the cytosolic component to which substances bind to enter the T3SS), and a syringe-like filament that extends past the double membrane and into the extracellular space (Izoré, 2011).

Structurally, the T3SS presents with numerous similarities to a bacterial flagellum complex: a base complex embedded in the IM consists of concentric stacked protein ring structures connecting the IM complex (which takes in substances from the bacterial cytoplasm) with the syringe-like filament structure (Galán, 2006). The needle filament is a protein tunnel through which unfolded proteins are transported from within the cytoplasm to the outside of the cell, where it comes directly into contact with the target cell. The filament is capable of extending its extracellular length in a similar manner to the growth of a bacterial flagellum, thus providing substances with a physical bridge to the interior of the target cell (Erhardt, 2010).

Figure 3.3: Artistic representation of the structure of a Type III secretion system (Reproduced from Wikipedia's original work).



The two main structures of the T3SS are clearly visible: the base complex within the IM, connected to the needle complex spanning the periplasmic cavity and extracellular space.

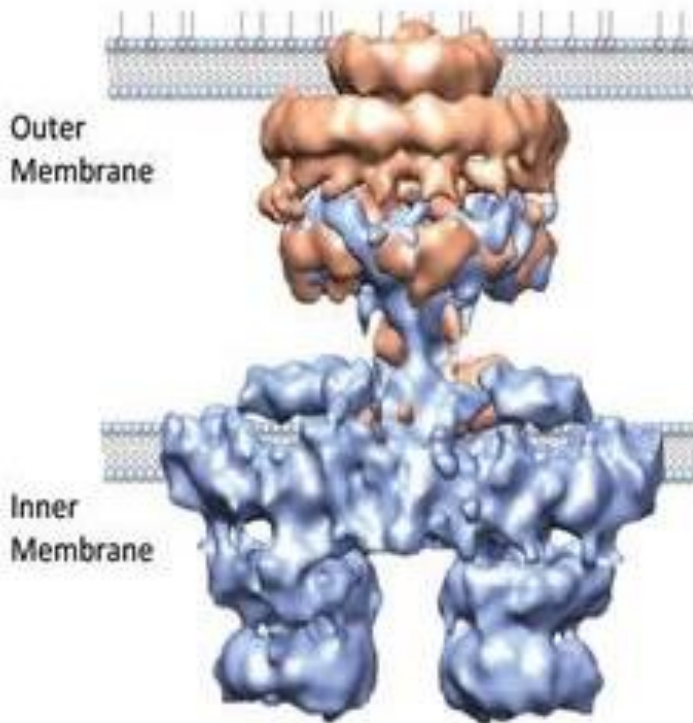
Type IV Secretion System

Type IV secretion systems (T4SS) are considered unique amongst the bacterial secretion systems for their ability to mediate the conjugation of

plasmid DNA into other eukaryotic or prokaryotic cells (Costa, 2015 [C. E. Alvarez-Martinez & P. J. Christie 2009]). T4SS also have the ability to transport bacterial effector protein substances from within the bacterial cytoplasm into extracellular space or other cells; however, the ability to translocate DNA has made T4SS a universal feature for bacteria species. The T4SS is made up of three main parts: an IM complex consisting of two protein barrels that take in substances from the cytoplasm, a stalk that connects the IM complex to the OM complex via the periplasmic cavity, and an OM complex which is responsible for releasing substances into extracellular space (Costa, 2015 [C. E. Alvarez-Martinez & P. J. Christie 2009]).

The core IM complex contains two protein barrel structures on the cytoplasmic side, which interact with substances in the cytoplasm that need to be secreted. The barrels take in unfolded protein substances (or plasmid DNA) and transport them through the stalk (past the periplasmic cavity) to the OM complex (Costa, 2015 [E. Cascades & P. J. Christie, 2004]). ATP-driven signals relay conformational changes to the OM complex's gate, allowing for the controlled secretion of proteins. In the case of plasmid DNA, tubular conjugative pili establish a pathway between cells, and provide a pathway for the transportation of single-stranded DNA between cells (Costa, 2015 [D. E. Bradley, 1980 & M. B. Durrenberger, W. Villager & T. Bachi, 1991]).

Figure 3.4: Graphic representation of the structure of a Type IV secretion system. (Reproduced from Professor Gabriel Waksman, University College London)



The three main structures of the T4SS are clearly visible: the twin barrelled IM complex, the stalk joining the IM and OM complex, and the OM complex gate.

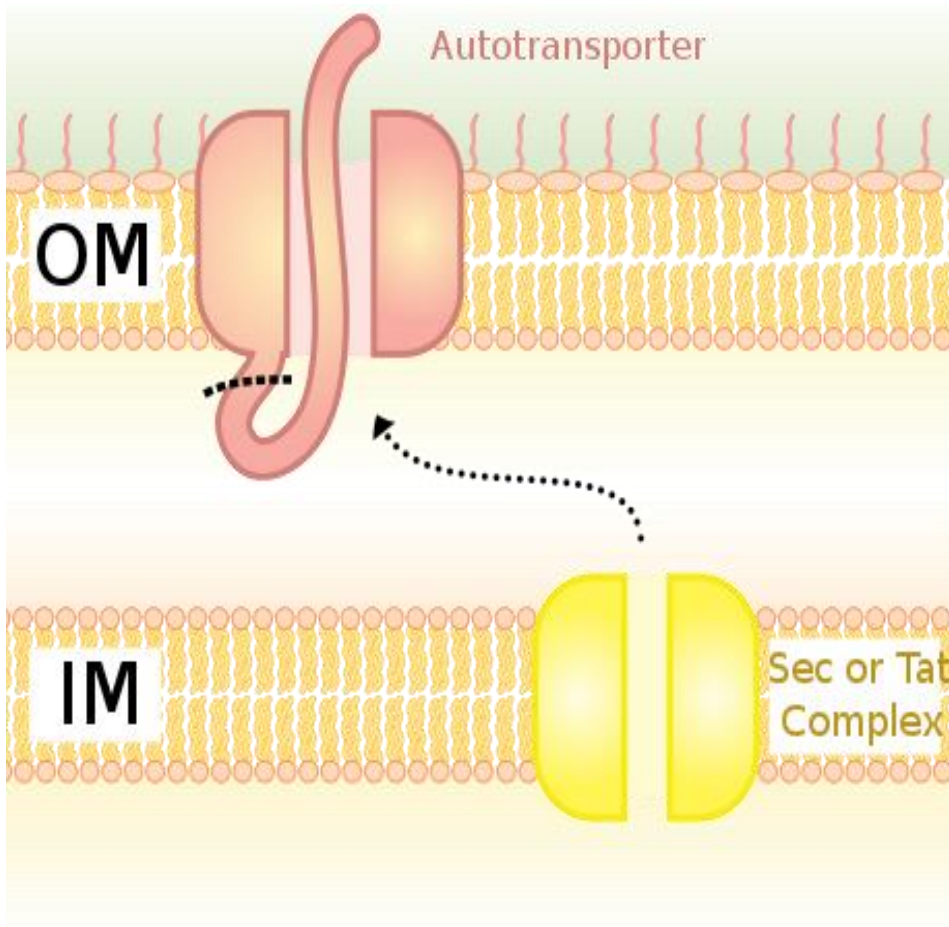
Type V Secretion System

Type V secretion systems (T5SS) are known as ‘autotransporter’ systems, derived from their unique mechanism allowing the fusion of substances to the secretion pore protein, effectively allowing substances to drive their own secretion from the cell (Costa, 2015 [J. C. Leo, I. Grin & D. Linke, 2012]). The T5SS is mainly involved in the secretion of virulence factors, but also participates in intercellular adhesion. Similarly to the T2SS, the T5SS also

requires the SecYEG translocon to transfer an unfolded polypeptide through the IM and into the periplasm, where the polypeptide then continues onto the OM complex via a beta barrel pore. The T5SS comprises two main units: the IM complex (SecYEG in the IM), and the OM complex (autotransporter channel gate). The autotransporter-OM complex, also known as the passenger domain, consists of the secretory substance and the translocator domain, passing through the OM complex as a beta barrel (Costa, 2015 [D. L. Leyton, A. E. Rossiter & I. R. Henderson, 2012]).

Autotransporter polypeptides are maintained in an unfolded state after transport through the IM by the SecYEG translocon. The periplasmic chaperone SurA promotes the assembly and secretion of the beta translocator domain past the OM complex. It has been theorised that transport of this passenger domain involves a hairpin loop, where the domain is threaded through the pore and into the extracellular space (Costa, 2015 [M. Junker, R. N. Besingi & P. L. Clark, 2009]). However, this mechanism is not fully understood, and has not been confirmed by research (Costa, 2015 [M. Junker, R. N. Besingi & P. L. Clark, 2009]).

Figure 3.5: Artistic representation of the structure of a Type V secretion system (Reproduced from Wikipedia's original work).



The two main structures of the T5SS are clearly visible: the SecYEG IM complex in the IM, and the autotransporter OM complex. Note that the secreted substance has bound to the OM complex in a hairpin loop fashion, having been transported across the periplasmic cavity.

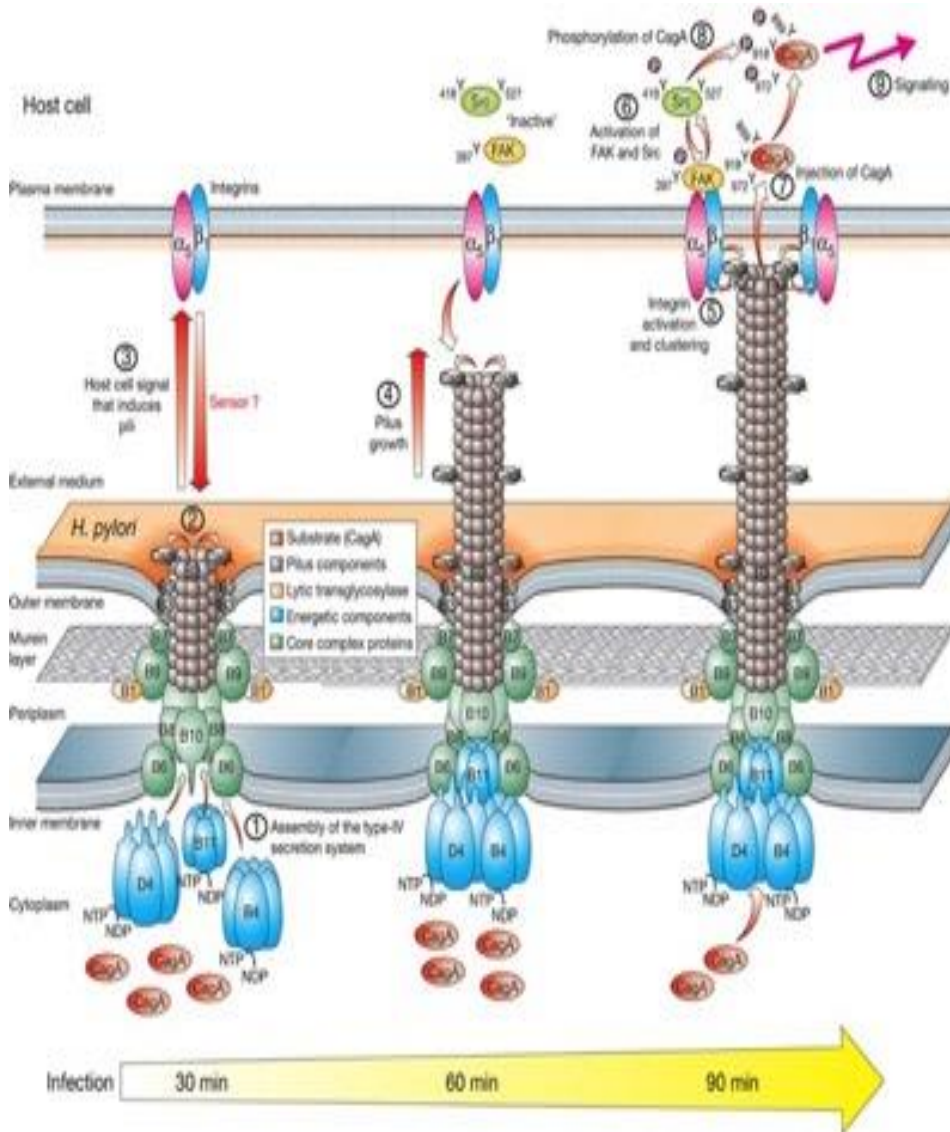
Type VI Secretion System

The Type VI secretion system (T6SS) spans the cell envelope, and is heavily involved in the transportation of toxic effector proteins into target eukaryotic and prokaryotic cells (Costa, 2015 [A. Zoued et al, 2014 & B. T. Ho, T. G.

Dong & J. J. Mekelanos, 2014]). It plays a key role in bacterial pathogenesis, and subsequent competition and survival. The T6SS is a relatively new discovery in the subfield of bacterial secretion systems, and thus many of the solved protein structures remain unknown. However, the T6SS clearly shows similarities to the other secretion systems, consisting of two main protein complexes. The IM complex of the T6SS is homologous to that of the T4SS (Costa, 2015 [L. S. Ma, J. S. Lin & E. M. Lai, 2009]) while the tail complex contains components that are related to contractile bacteriophage tails (tail sheath, inner tube, baseplate) (Costa, 2015 [S. Pukatzi et al, 2006]).

While the exact mechanism of T6SS secretion remains unconfirmed, models suggest that following an extracellular signal, conformational changes in the baseplate of the tail complex trigger a contraction in the tail sheath. This allows for the translocation of multiple effector proteins out of the cytoplasm and past the IM and OM, into target cells (Costa, 2015 [M. Kudryashev et al, 2015]).

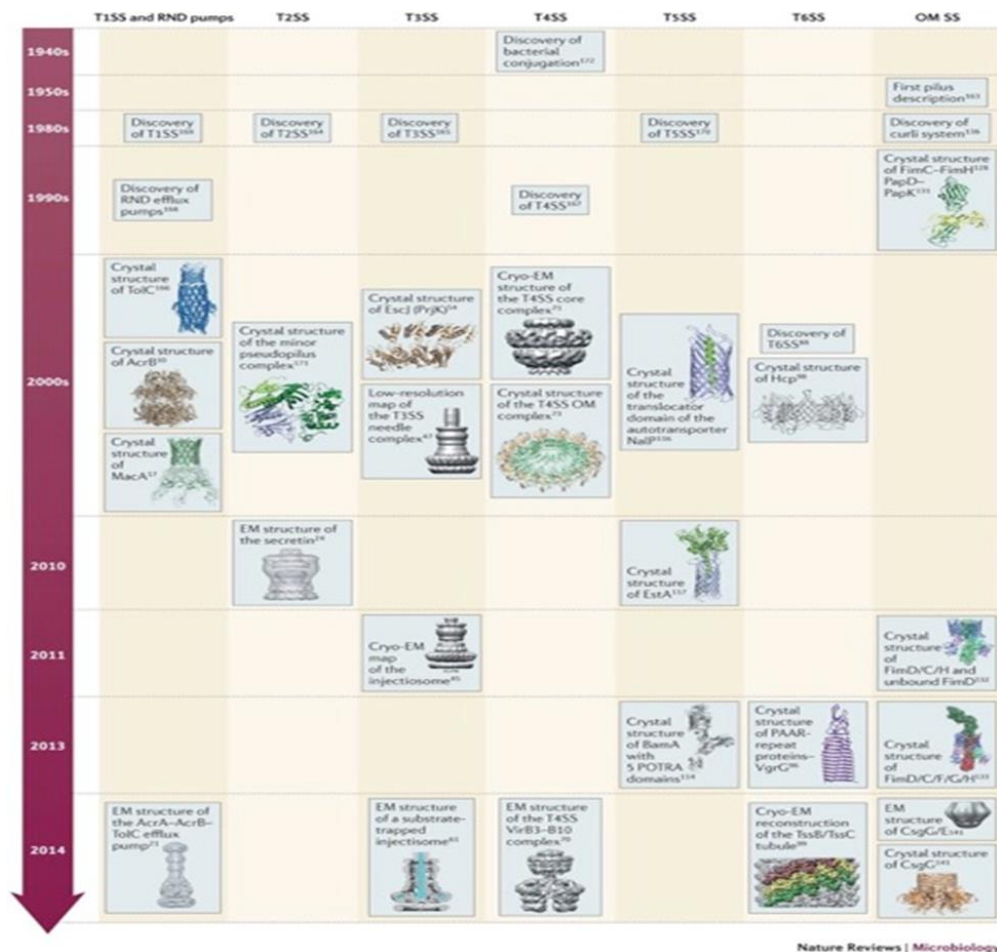
Figure 3.6: Graphical representation of the structure of a Type VI secretion system (Reproduced from Kwok, 2007).



The two main structures of the T6SS are clearly visible: the IM complex to which substrates in the cytoplasm bind to before being transported into the periplasm, and the tail complex, shown to be growing and secreting substances into the extracellular space.

A Brief History

Figure 4: Tabular representation of the discoveries of secretion systems (Reproduced from Costa, 2015).



The above figure provides a brief summary of the history of secretion systems in graphical form, and highlights key discoveries in the field that have proved significant in their contributions to the understanding of secretion.

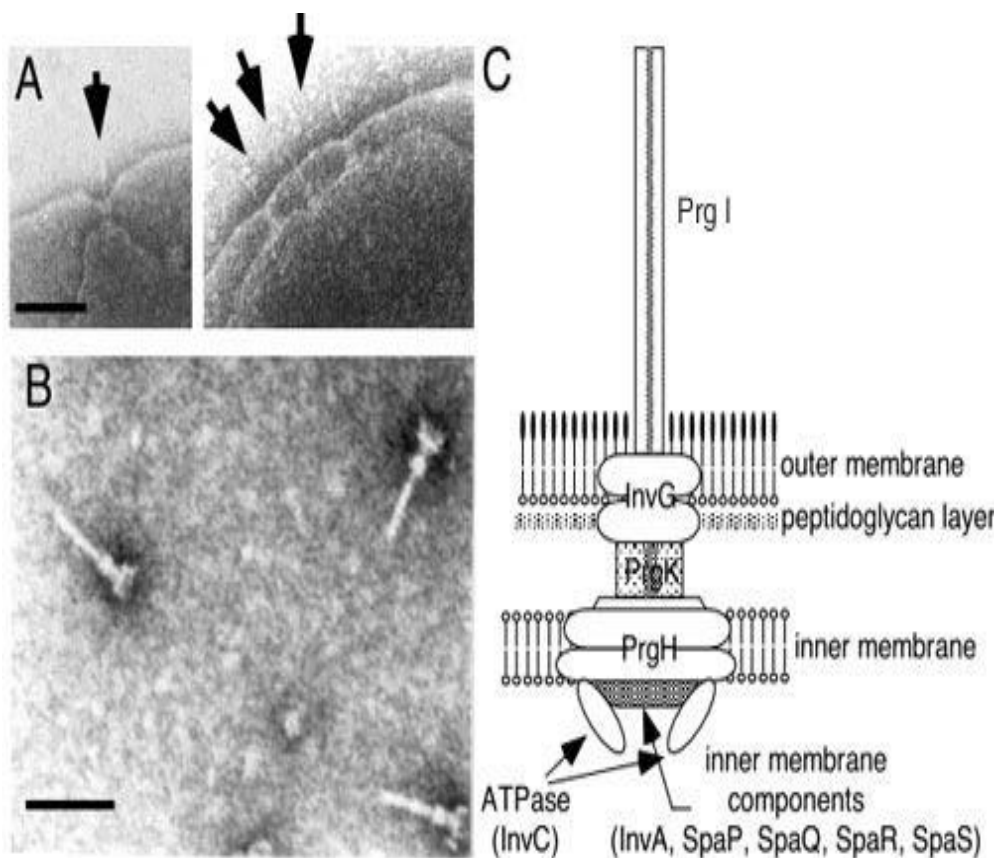
The discovery of bacterial conjugation, while not directly related to secretion and secretion systems, is considered the beginning of an introduction to the

idea of cellular secretion (Costa, 2015). This secretion would not only involve metabolic products such as toxins or other biochemical compounds, but even segments of DNA that conferred evolutionary advantages such as antibiotic resistance. This served to distinguish the processes of secretion and excretion, with the latter simply being a removal of metabolic waste, rather than the intentional release of substances. Further advancements in technology allowed for the discovery and solving of more complex protein structures within the individual secretion systems, providing a greater morphological understanding of cellular biology, as well as the evolutionary pathways of such complex protein machinery. Understanding these concepts has allowed research to focus on the applicability of this knowledge in future situations.

TYPE III SECRETION SYSTEM: STRUCTURE AND ASSEMBLY

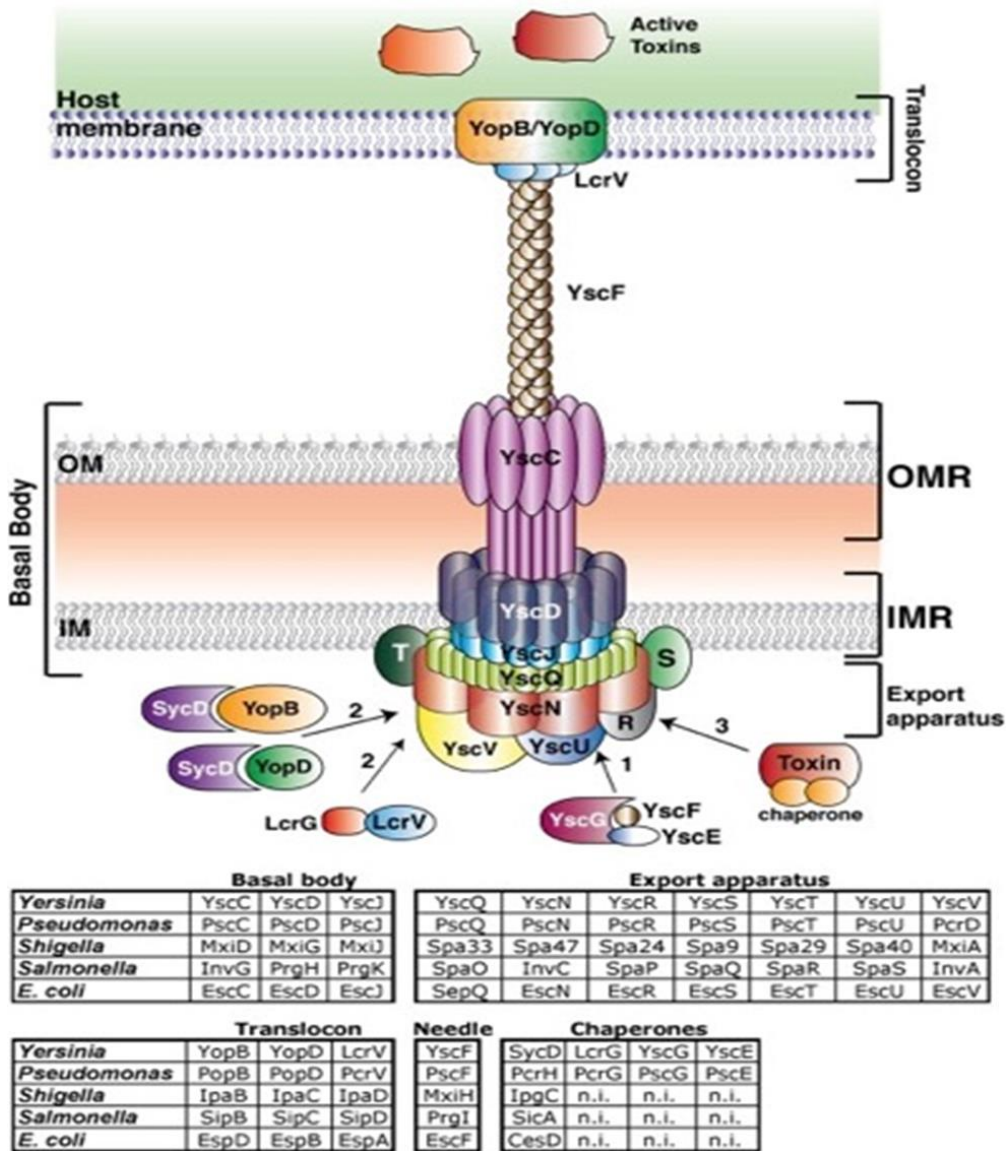
Architecture of the Type III Secretion System

(Next page) Figure 5.1: EM Photograph and representative schematic of a Type III secretion system (Reproduced from Galín).



As mentioned above, the T3SS is a macromolecular machine that allows bacterial species to directly ‘inject’ toxin compounds into target cells, allowing the bacterium to accomplish a wide variety of functions such as pathogenesis and apoptosis (Costa, 2015 [G. R. Cornelis, 2006]). The T3SS is comprised of three (including the translocon pore tip) main parts: the basal body, which is a double membrane-spanning protein structure consisting of concentric stacked rings; the hollow needle complex through which semi-unfolded proteins are transported out of the cell; and the translocon tip, which is a pore inserted into the target cell and forms hydrophilic bridges (Izor é 2011).

Figure 5.2: Detailed structure of a Type III secretion system (Reproduced from Izor é 2011).



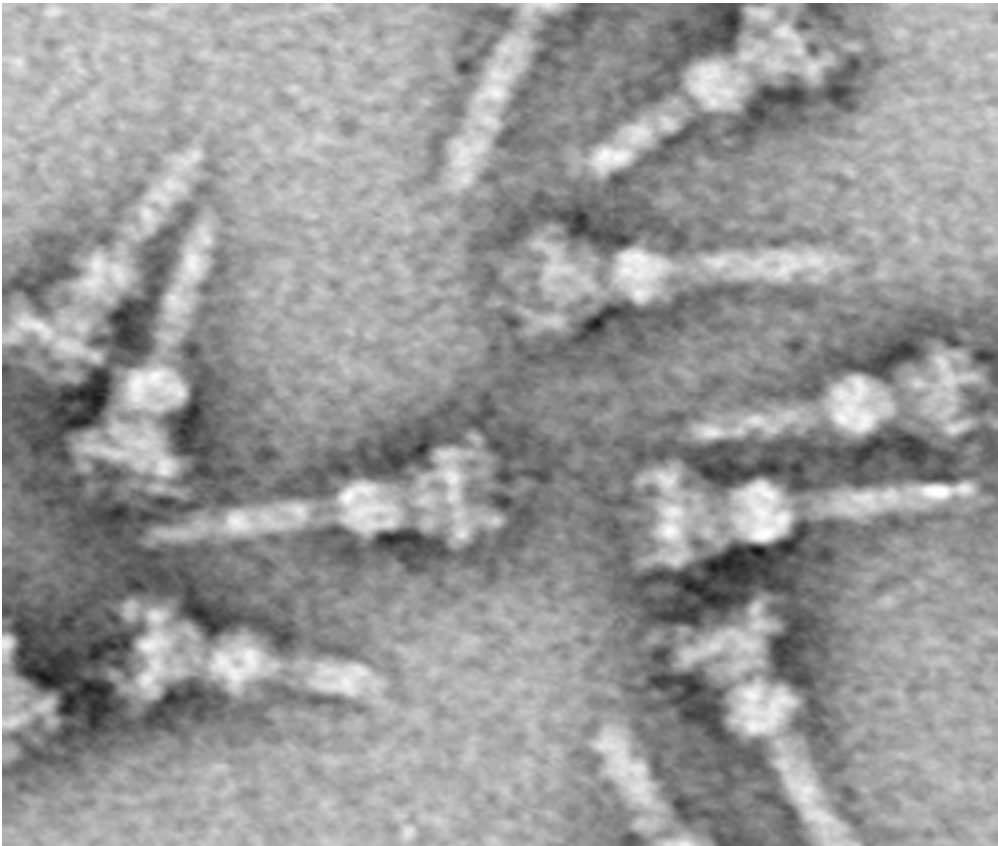
The basal body complex is a protein ring structure in the IM, consisting of concentric stacked rings. The IM ring is organised into two concentric rings

with different diameters, is composed of two proteins (EscD/EscJ), and represents the largest ring visible in the needle complex (Izoré 2011). The basal body is involved in the uptake of substances from the bacterial cytoplasm, as well as the extension of the needle filament via F1F0-ATPase. In this regard, the basal body of the T3SS bears a remarkable similarity to that of the bacterial flagellum: both have an anchor base of proteins embedded in the IM, which secure the ATPase and needle filament inside and outside the IM respectively; and both require ATP hydrolysis to create the vital extracellular component of the system (needle filament and flagella) (Izoré 2011). Additionally, the needle (and its connector to the base) of the T3SS are analogous to the flagellar hook, which connects the flagellar filament to its anchor proteins (Izoré 2011).

The needle filament is perhaps the most recognisable feature of the T3SS, and constitutes the main function of the system (Galán, 2006). The needle filament itself is made of a repeating sequence of a small protein (EscF), and protrudes from the OM as the extracellular pathway for toxins to be delivered into a target cell. This is extremely similar to the structure of the flagella filament, a polymerisation of small protein monomers that extend into extracellular space. The translocon tip of the needle filament is another polymer of an adaptor protein EspA, and is mainly involved in positioning the entrance of substances into the target cell and regulating the secretion of substances (Galán, 2006). However, the specifics of translocon interaction with target cell membranes and later effector protein secretion have yet to be uncovered (Pei-Chung Lee, 2014).

The translocon tip of the needle filament is responsible for regulating the secretion of substances through the T3SS, as well as interacting with the target cell membrane to successfully ‘inject’ toxins directly into the target cell (Tobe, 2006). While the interaction between needle filament monomers and tip proteins is not fully understood, it is clear that the translocon is essential in determining when substances are secreted by the T3SS, and thus essential in the bacterium’s interaction with its competition and environment (Tobe, 2006).

Figure 5.3: EM photograph of Type III secretion system complexes (Reproduced from Schraidt, 2010).



Assembly of the Type III Secretion System

Figure 5.4: Proteins involved in the assembly of the basal body of the Type III secretion system (Reproduced from Izor é 2011).

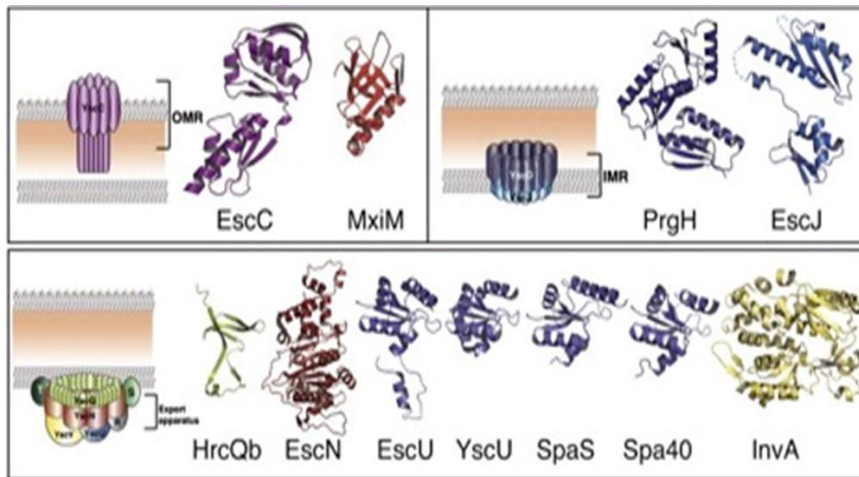
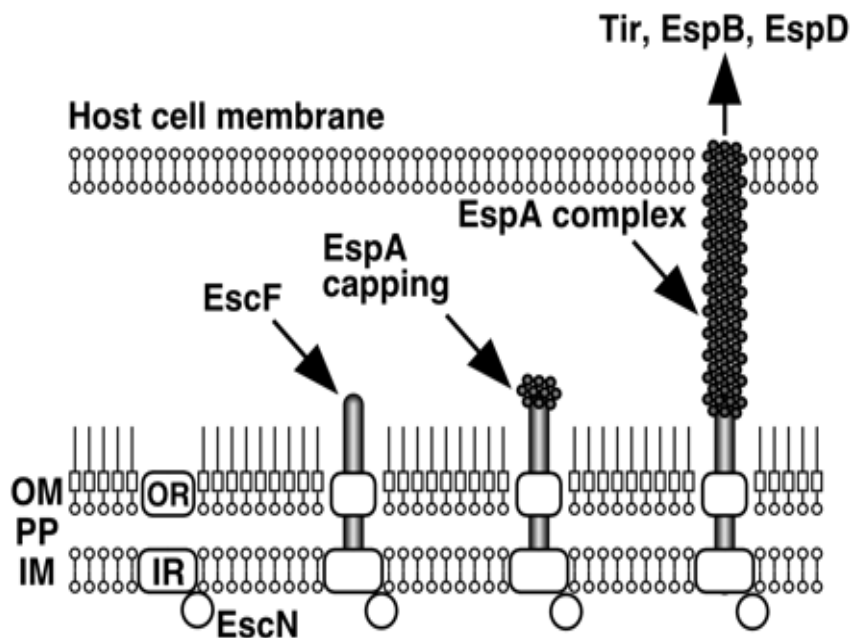


Figure 5.5: Simplified assembly of the injectosome (Reproduced from Sekiya, 2001).



The above figures provide a simple and brief overview of the assembly of the T3SS. The assembly begins with the anchorage base proteins embedded in the IM, eventually forming the basal body structure. Then, the needle filament is formed by polymerising simple monomer proteins in a rotating fashion, driven by F1F0-ATPase, until it extends past the periplasm and OM, into extracellular space. Finally, the polymerisation of adaptor proteins forms the translocon tip in extracellular space. This stage of the assembly results in a functional T3SS (Diepold, 2014). The following stage of assembly involves further extension of the needle filament to reach a target cell (repeated polymerisation), and the interaction between the translocon tip of the needle filament and the target cell membrane, eventually allowing the T3SS to directly transport substances to the interior of the target cell (Diepold, 2014).

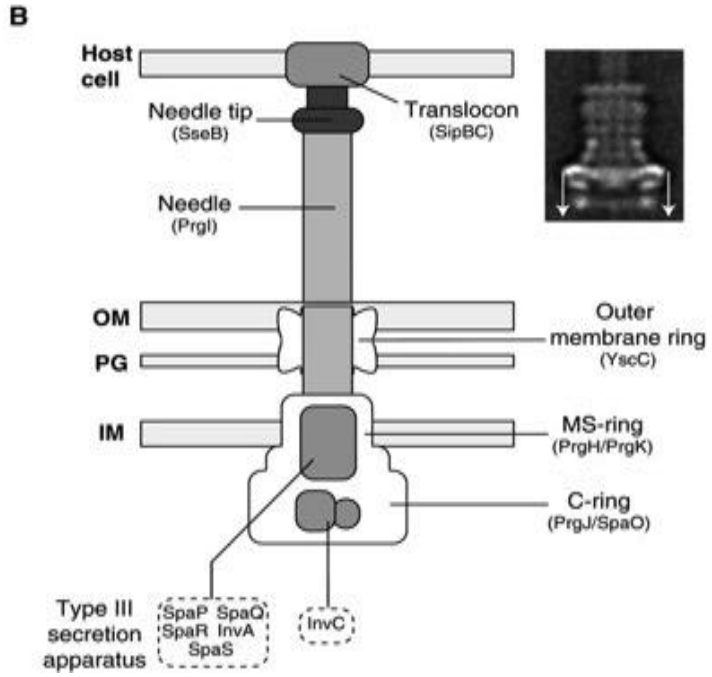
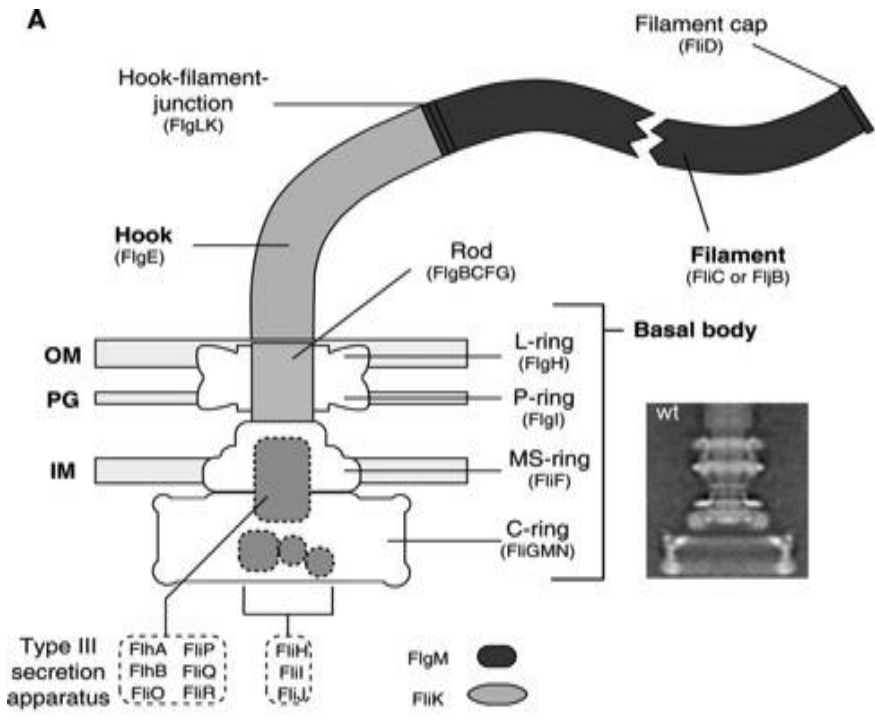
As seen in Figure 5.2, the basal body is formed in distinct sections that anchor the entire complex to the IM in preparation for the assembly of the needle filament. The rings in the basal body are formed from a protein of the secretin family. Each of the two rings is formed separately (the inner ring of SctJ, the outer ring of SctD), and then connected with the SctD ring through its periplasmic domain. Additionally, the export apparatus of the basal body is formed of the proteins SctRSTUV in the area surrounded by the ring structure, where they select substances to enter the translocation channel (Izoré 2011 & Gauthier, 2003). Despite their key roles within the T3SS, little is known about their interaction within the system itself. Though these proteins alone constitute the basal body, it is only complete when the F1F0-ATPase is assembled as well, as the needle filament cannot be formed without it (Gauthier, 2003).

The needle filament is formed by the rotation of the F1F0-ATPase as protons flowing down a concentration gradient across the IM turn the ATPase; each subsequent turn results in a single monomer adaptor protein being polymerised to a growing hollow filament (Sekiya, 2001). This polymerisation continues until the needle filament has extended beyond the OM and into extracellular space, where a different protein monomer is polymerised on the distal end to form the translocon tip (Gauthier, 2003). The EspA protein serves as the translocon tip, as well as the extension complex that allows a pre-existing T3SS to reach for a target cell membrane. However in the initial stages of assembly, the EspA protein simply forms the translocon tip and further polymerisation is withheld (Gauthier, 2003).

The assembly of the T3SS is remarkably similar both in components and sequence to the bacterial flagella, and comparisons between the two macromolecular structures have often been drawn (Aizawa, 2015). Despite the similarities in body anchor proteins and the formation of filaments, the main difference lies in the filament that is synthesised. As seen in Figure 5.4 below, there is a notable difference between the eventual shape and function of the filament (Erhardt, 2010). The hooked shape of the flagella filament reduces its efficiency if an attempt to engage another cell is made, while the straight shape of the needle filament makes for a poor propeller (Erhardt, 2010). As these are complex machines that have resulted from continued evolution, it is easy to see how minute changes in structure resulted in vast differences in function, as though the evolutionary history of each macromolecular complex was played out in its assembly. This comes across

as an unsurprising find when comparing the structure and function of T3SS and bacterial flagella (Erhardt, 2010).

(Next page) Figure 5.4: Schematic comparison of the Type III secretion system and flagella (Reproduced from Erhardt, 2010).



TYPE III SECRETION SYSTEM: ESCHERICHIA COLI

Secretion in Gram-negative species of bacteria is often pathogenic in nature, involving the secretion of toxins into target eukaryotic (human) cells to aid in invasion and colonisation (Shraidt, 2010). These toxins vary greatly in severity, speed of action, affected populations and so on and so forth. However, the commonality between all these toxins is that they require a method of secretion. Secretion systems such as the T3SS or the T6SS specialise in the secretion of bacterial toxin effector proteins, and can greatly enhance the pathogenic qualities of a bacterium. The T4SS also allows for bacteria to gain new toxins (via conjugation of plasmid DNA), whose effects can be deadly when combined with a pre-existing secretion system—as seen in the case of *E.coli* strains that acquired the ability to produce a Shiga-like toxin (Philip I. Tar, 2005).

It is known that not all serotypes of *E.coli* are pathogenic, and amongst pathogenic serotypes of *E.coli*, the severity of resulting pathogenesis varies greatly from simple diarrhea to hemolytic-uremic syndrome.(Philip I. Tar, 2005). Pathogenic strains of *E.coli* are distinguished from non-pathogenic strains by the presence of antigens that elicit an immune response from animals (humans). These are the O, K and H antigens, found in the lipopolysaccharide layer, the capsule and the flagella respectively. The various pathogenic strains are known for causing diarrhea, shigellosis (via production of a toxin almost identical to the Shiga toxin), and hemolytic-uremic syndrome and sudden kidney failure. In the case of the latter two, pathogenesis is dependent on the secretion of specific substances (Shiga-like toxin and hemolysin respectively) into target cells (Tobe, 2006). It is clear

how the design of the T3SS greatly aids the bacterium in accomplishing this task (Tobe, 2006).

In the case of *E.coli* strains that produce a Shiga-like toxin (Stx-1 and Stx-2), the toxin is secreted directly into the extracellular space, where it is taken up into eukaryotic cells (Johnson, 2008). An example is the secretion of Stx-1 and Stx-2 near vascular endothelium cells: the endothelium cells take up the toxin and die when the toxin renders them incapable of synthesizing new proteins. As endothelium cells need to continuously renew themselves, this lack of renewal leads to a breakdown in the vascular lining and subsequent hemorrhage—the classic symptom of bloody diarrhea (Johnson, 2008). The ability of the T3SS to extend the needle filament and reach more cells is key in allowing *E.coli* to secrete the toxin over a larger area (Sekiya, 2001). Similarly, the damage caused by hemolysins is increased when spread over a large area. The secretion of alpha hemolysin results in the death of erythrocytes by oligomerising after binding to their cell membranes and creating pores. In both cases, the T3SS' efficiency in secreting substances into target cells is key in pathogenesis, driven by the successful invasion and colonisation by the bacteria (Sobhanifar, 2015).

FUTURE CONSIDERATIONS

Despite the multitude of advancements in the field, many questions about the fundamentals of secretion systems remain unanswered. Unknowns regarding the interactions between cellular machinery, or the mechanisms behind substance secretion seem almost basic, and yet are infinitely more complex. As with all unknowns, the possibilities of further research are almost

limitless. Amongst popular considerations are the likelihood of new drug targets against bacteria with new antibiotic resistance, or new methods of drug delivery utilising knowledge regarding the secretion systems, especially T3SS and T4SS.

ACKNOWLEDGEMENTS & BIBLIOGRAPHY

This paper would not have been possible without the help and advice of Professor Grant Jensen, as well as the content covered during his classes and the sources cited below.

Works Cited:

- Aizawa, Shin-ichi. "Bacterial Flagella and Type III Secretion Systems." *Oxford University Press*. Oxford University Press: FEMS Microbiology, 1 Aug. 2001. Web. 13 July 2015.
- Costa, Tiago R. D., Catarina Felisberto-Rodrigues, Amit Meir, Marie S. Prevost, Adam Redzej, Martina Trokter, and Cabriel Waksman. "Secretion Systems in Gram-negative Bacteria: Structural and Mechanistic Insights." *Nature Reviews*. Nature Reviews: Microbiology, 15 May 2015. Web. 13 July 2015.
- Diepold, Andreas, and Samuel Wagner. "Assembly of the Bacterial Type III Secretion Machinery." *Oxford University Press*. Oxford University Press: FEMS Microbiology, 1 July 2014. Web. 13 July 2015.
<[http://dx.doi.org/10.1093/femsre.oxfordjournals.org/content/38/4/fiu082.full](http://dx.doi.org/10.1093/femsre/fiu082)>.

- Erhardt, Marc, Keiichi Namba, and Kelly T. Hughes. "Bacterial Nanomachines: The Flagellum and Type III Injectisome." *Cold Spring Harbour Perspectives in Biology*. Cold Spring Harbour Perspectives in Biology, 6 Oct. 2010. Web. 13 July 2015.
<[http%3A%2F%2Fcshperspectives.cshlp.org%2Fcontent%2F2%2F11%2Fa000299.full](http://3A%2F%2Fcshperspectives.cshlp.org%2Fcontent%2F2%2F11%2Fa000299.full)>.
- Gal á n, Jorge E., and Hans Wolf-Watz. "Protein Delivery into Eukaryotic Cells by Type III Secretion Machines." *Nature International Journal of Weekly Science*. Nature Reviews: Microbiology, 30 Nov. 2006. Web. 13 July 2015.
- Gal á n. "The Needle Complex." *The Needle Complex*. Yale School of Medicine, n.d. Web. 13 July 2015.
<https://tools.medicine.yale.edu/galan/www/Pages/galan_figure2.html>.
- Gauthier, Annick, Jose Luis Puente, and B. Brett Finlay. "Secretin of the Enteropathogenic Escherichia Coli Type III Secretion System Requires Components of the Type III Apparatus for Assembly and Localisation." *American Society for Microbiology*. American Society for Microbiology: Infection and Immunity, June 2003. Web. 13 July 2015.
- Geller, Bruce L., N. Rao Movva, and William Wickner. "Electrochemical Potential." *Gold Book IUPAC Compendium of Chemical Terminology* (2009): n. pag. *Proceedings of the National Academy of Sciences of the United States of America*. PN. Web. 13 July 2015.

- Izoré Thierry, Viviana Job, and André Dessen. "Biogenesis, Regulation and Targeting of the Type III Secretion System." *Science Direct*. Cell Press, 10 May 2011. Web. 13 July 2015.
<<http://www.sciencedirect.com/science/article/pii/S0969212611001390>>.
- Johnson, Steve, and Ariel Blocker. "Characterisation of Soluble Complexes of the *Shigella Flexneri* Type III Secretion System ATPase." *Oxford University Press*. Oxford University Press: FEMS Microbiology, 1 Sept. 2008. Web. 13 July 2015.
<<http://femsle.oxfordjournals.org/content/286/2/274.full>>.
- Kwok, Terry, Dana Zabler, Sylwia Urman, Manfred Rohde, Roland Hartig, Silja Wessler, Rolf Misselwitz, Jürgen Berger, Norbert Sewald, Wolfgang König, and Steffen Backert. "Heliobacter Exploits Integrin for Type IV Secretion and Kinase Activation." *Nature International Journal of Weekly Science*. Nature Reviews: Microbiology, 18 Oct. 2007. Web. 13 July 2015.
- Lee, Pei-Chung, Stephanie Elizabeth Zmina, Charles Morgan Stopford, Jonida Toska, and Arne Rietsch. "Control of Type III Secretion Activity and Substrate Specificity by the Cytoplasmic Regulator PcrG." *Proceedings of the National Academy of Sciences of the United States of America*. Ed. Philippe J. Sansonetti. PNAS, 2 Apr. 2014. Web. 13 July 2015.

- Liang, Yi, Nil Celebi, Minyong Chen, and Ross E. Dalbey. "Sec/SRP Requirements and Energetics of Membrane Insertion of Subunits A, B and C of the Escherichia Coli F1F0 ATP Synthase." *The Journal of Biological Chemistry*. The Journal of Biological Chemistry, 19 July 2004. Web. 13 July 2015.
- Pham, Thanh H., Xiaofei Gao, Karen Tsai, Rachel Olson, Fengyi Wan, and Philip R. Harwidge. "Functional Differences and Interactions between the E.coli Type III Secretion Effectors." *American Society for Microbiology* (2012): n. pag. 20 Mar. 2012. Web. 13 July 2015.
- Pugsley, A. P. "The Complete General Secretory Pathway in Gram-negative Bacteria." *National Library of Medicine*. Microbiological Reviews, Mar. 1993. Web. 13 July 2015.
<[http%3A%2F%2Fwww.ncbi.nlm.nih.gov%2Fpmc%2Farticles%2FPMC372901%2F](http://www.ncbi.nlm.nih.gov/pmc/articles/PMC372901/)>.
- Saier, Milton H. "Protein Secretion and Membrane Insertion Systems in Gram-Negative Bacteria." *Journal of Membrane Biology J Membrane Biol* 214.1-2 (2006): 75-90. *Microbe Magazine*. Microbe Magazine, Sept. 2006. Web. 13 July 2015.
- Schraidt, Olivier, Matthew D. Lefebvre, Matthias J. Brunner, Wolfgang H. Schmied, Andreas Schmidt, Julia Radics, Karl Mechtler, Jorge E. Galán, and Thomas C. Marlovits. "Topology and Organisation of the Salmonella Typhimurium Type III Secretion Needle Complex Components." *PLOS*. PLOS Computational Biology, 1 Apr. 2010. Web. 13 July 2015.

<<http://www.plos.org/plospathogens/article?id=10.1371/journal.ppat.1000824>>.

Sekiya, Kachiko, Minako Ohishi, Tomaoki Ogino, Koichi Tamano, Chihiro Sasakawa, and Akio Abe. "Supermolecular Structure of the Enteropathogenic Escherichia Coli Type III Secretion System and Its Direct Interaction with the EspA-sheath-like Structure." *Supermolecular Structure of the Enteropathogenic* (n.d.): n. pag. *Proceedings of the National Academy of Sciences of the United States of America*. PNAS. Web. 13 July 2015.

Sobhanifar, Solmaz. "Clearing the Path for Infection: E Coli Type III Secretion System." *Centre for Blood Research*. Centre for Blood Research, 14 May 2015. Web. 13 July 2015.

Tar, Phillip I., Carrie A. Gordon, and Wayne L. Chandler. "Shiga-toxin-producing Escherichia Coli and Haemolytic Uraemic Syndrome." *Science Direct*. The Lancet, Mar. 2005. Web. 13 July 2015. <<http://www.sciencedirect.com/science/article/pii/S0140673605711442>>.

Tobe, Toru, Scott A. Beatson, Hisaaki Taniguchi, Hiroyuki Abe, Christopher M. Bailey, Amanda Fivian, Rasha Younis, Sophie Matthews, Olivier Marches, Gad Frankel, Tetsuya Hayashi, and Mark J. Pallen. "An Extensive Repertoire of Type III Secretion Effectors in Escherichia Coli O157 and the Role of Lambdoid Phages in Their Dissemination." *Proceedings of the National Academy of Sciences of the*

United States of America. Ed. E. Peter Greenberg. PNAS, June 2006.
Web. 13 July 2015.

Zhou, Mingxu. "*Escherichia Coli* type III secretion system 2: a new kind of T3SS?" *Veterinary Research*. BioMed Central, 19 Mar. 2014. Web. 13 July 2015.



The Motif of Culture: The Costume Culture of the Yi Ethnic Group of Liangshan

Yilin Wen

Author background: Yilin grew up in China and currently attends Chengdu Foreign Languages School in Chengdu. Her Pioneer seminar program was titled “Non-Verbal Communication.”

Overview of the Yi ethnicity

The Yi ethnic group is a distinct and old ethnicity living in southwest China, mainly in Sichuan, Yunnan and Guizhou provinces. It is the sixth largest minority group in China. Within the group, there exists great diversity. People living in different areas have a variety of dialects, styles of dress and religions. Since the place where they live is mountainous and remote, the Yi culture has weathered the influence of the Han, the dominant ethnic group in China with 91.5% of the whole population (National Bureau of Statistics of China, 2011), and other foreign cultures. As a symbol of culture and identity, the traditional clothing of the Yi people is more than just a simple protection from the environment; rather, its importance extends to a cultural and social level. From colors and patterns to texture and shape, Yi clothing sheds light on the depth, diversity and longevity of its culture.

This paper aims to explore the connection between the culture and clothing of Yi people living in Liangshan Yi Autonomous Prefecture (in Sichuan province). It examines various cultural, social and historical factors that influenced the clothing of Yi people and how these characteristics reflect the Yi culture. Also, this paper studies the Yi costume from a

dynamic viewpoint, focusing on the impact Han culture has had on Yi style and sharing thoughts about the future of Yi clothing.

Factors that influenced Yi clothing

Geography and Climate

Because of its high elevation, Liangshan has long hours of sunshine and strong solar radiation. It is located in the subtropical monsoon climatic zone and therefore has distinct dry and wet seasons during the year. The temperature in Liangshan is generally mild, averaging approximately 16 degrees Celsius; however, temperature can vary greatly during the day and slightly over the year. The local agriculture and animal husbandry are influenced by the climate. Because the region is generally dry, Yi people can't grow cotton; thus, they usually make clothes out of woolen fabric and flax (Li, J., ed., 2004). The Yi clothing also represents the Yi farming and herding culture through its decorations. Because the Yi is not a very advanced or modern culture, its production strongly depends on the influence of nature. Today, Yi people still worship and fear nature. They see the sun, the moon and the stars as omnipotent and powerful and thus have incorporated these patterns into their clothing (Wang, 2000). As hunters, Yi people also print symbols of their prey, cockscombs and goat horns for example, on their clothes.



Source: History in Costumes: on Costume Culture of the Yi Ethnic Group in Zhaojue. China Ethnic Photographic Art Press, 2004

1a. Yi men wearing Cha'erwa and the "hero's knot"

Generally, Yi people wear the Cha'erwa, a beautiful felt coat made of wool that protects them from the cold in the morning and in the evening and is also useful during wet seasons (see picture 1 a). The men wear turbans with a knot of hair extending outside, which they call the "hero's knot" (Li, Ji., 2004). The Yi believe that this knot represents the soul and dignity of a man; therefore it is forbidden for anyone else to touch this knot. In addition to the head ornament, Yi men usually wear woolen shirts with baggy flax pants; the width of the bottoms of their pants varies according to region. These baggy pants allow them to move nimbly in the mountains. Women also wear head ornaments, along with woolen shirts and pleated skirts that

vary not only by region, but also with age and social class, which will be discussed later in the paper (see picture 1 b).



Source: History in Costumes: on Costume Culture of the Yi Ethnic Group in Zhaojue. China Ethnic Photographic Art Press, 2004

1b. Yi women

From ravines to plains to mountains, Liangshan has a complex topography that divides the prefecture into different areas. In addition to the broken terrain, the migratory history and language contact with other ethnicities have divided the Yi people into three subgroups with their own distinct dialects--- Yi Nuo (伊诺), Sheng Zha (圣乍) and Suo Di (所地) (Cheng, 2007). Yi Nuo-speaking Yi live in northeastern Liangshan, Sheng Zha in the middle west, and Suo Di in the south. The main difference among

the clothing styles of these three groups is the width of the bottom of their pants. Yi Nuo has the widest pant bottom, Sheng Zha's is medium, and Suo Di's the tightest (Wang, 2000). In addition, women's head ornaments also vary in shape from one region to another, but generally they all have umbrella-like shapes, which are helpful during wet seasons (Li, J., ed., 2004).

As my Yi informants have told me, the older Yi generations can tell with ease a person's origin by his clothing, but their children, the teenagers, are gradually losing this perception. Also, the dress code of different dialectal groups is not as strict as that of different castes; thus people from one place can wear the clothes of another region.

Combative history

In ancient times, Liangshan was a remote and savage land. People of different ethnicities, the Tibetan, the Qiang and the Yi, all shared this mountainous area (Ling, 1997). Instead of being a melting pot, the region was more of a simmering pot, where clashes and wars were constant because people struggled for land and resources (Ling, 1997). In order to expand its territory, conquer foreign ethnicities and protect its people, the Yi developed a strong military. They deemed their soldiers heroes and spared no effort in making beautiful but also practical warrior clothes for the army. The soldier's shirts were made of Pulu (a woolen fabric produced in Tibet for making blankets) on the outside and patched with a layer of lint on the inside (Li, J., 2004) (see picture 2 a). On top of the shirts, armor plates made of animal skins (elephant or cattle), along with skirt-like lower hems composed of several hundred small oblong pieces of leather, protected the upper part of their bodies (Li, J., 2004). Their garments were usually colored with black,

red and yellow which, according to the Yi, are auspicious colors that could keep the soldiers safe and sound.



Source: http://www.china.com.cn/culture/minsu/2009-11/10/content_19376298_3.htm
(accessed on July 2)

3a. Yi soldiers

Social Hierarchy

Although the exact date of its emergence is indeterminate and controversial, the Liangshan Yi people established a slave society around 1600 BC, during the Xia dynasty (Zhu, 2013), and maintained such a system until the 1950s. Their old caste system was complex, dividing people mainly into five categories based on their profession, wealth and kinship: the Zi Mo (tribe leaders), the Nuo He (aristocrats), the Qu Nuo (slaves who had a certain degree of independence), the A Jia (slaves of lower rank than the Qu Nuo) and the Jia Xi (slaves of the lowest rank, subjugated to the A Jia)

(Ling, 1997). Because the Qu Nuo had a certain degree of economic liberty, they could own slaves (A Jia and Jia Xi) and gather wealth through trade. As time went by, more and more Qu Nuo gathered enough wealth to equate to that of the Nuo He, and hence became the nouveau riche. With the rise of the Qu Nuo in the seventeenth century, the original system of five categories gradually evolved into one with only two categories: the Black Yi (aristocrats) and the White Yi (commoners and slaves) (Zhu, 2013).

Clothing, as the most manifest symbol of a culture, also signified the stark differences in social status among the Yi people. Because Black Yi are generally more affluent, their clothes are of the highest quality, usually made of pure wool and cotton (Ma, 2008). They also wear jewels that are made of gold and silver. By contrast, White Yi clothes are usually made of low quality wool or linen. Also, White Yi women's skirts are naturally knee-length, because this allows them to do chores more easily (Ma, 2008). Since the Yi people consider the color black as a symbol of elegance and authority, the Black Yi wear mainly black, or sometimes blue and white robes. On the contrary, the White Yi's clothes are bright and colorful, usually red and yellow and seldom black (see picture 3 a, 3 b).



Source: Provided by informant

3b. A white Yi on the left with three black Yis



Source: Provided by informant

3b. A white Yi bride and a black Yi teenager

With the arrival of Han people in recent years, many ethnic problems have emerged in the Liangshan area. Because the Yi people think that the Han have an unfair and unequal share of wealth and political power in their autonomous prefecture, the divisions between these two ethnicities are growing sharper. The pressure from the Han culture has pushed Yi people closer together and has strengthened the solidarity among them. As interviews with several of my Yi informants have shown, they now identify themselves more with the ethnic group as a whole (e.g. Yi as opposed to Han) than with caste (e.g. Black Yi as opposed to White Yi). Because of the pressure from the Han culture and also the incoming democratic ideas, the stringent dress code for caste is now fading away. People with different castes don't really care about the clothes they wear anymore. Black Yi love to wear the colorful pleated skirts which originally belonged to White Yi, and White Yi wear colors that were exclusive to Black Yi.

However, because the caste system has been in place for almost 400 years, the idea of class and rank is so deeply ingrained in people's minds that today, one class still holds a certain feeling of hostility or superiority over another. Normally, one class wouldn't attend familial rituals (e.g. weddings, funerals) of another, and interethnic marriage is generally forbidden.

Traditions—The Coming-of-age Ceremony and the Torch Festival

From birth to death, Yi traditions and clothing are connected with every stage of people's lives. In this paper, I will focus on the coming-of-age ceremony of Yi girls and the Yi new year celebration, the Torch Festival.

The most prominent mark in Yi culture that denotes the transition from childhood to adulthood is clothing. In Liangshan, when girls step into

adulthood, their families would hold a coming-of-age ceremony called “Sha La Luo,” which means “take off the dress of childhood and take on the dress of adulthood” (Li, D., 2012). The activities of “Sha La Luo” vary from region to region, but across all the Yi regions there is a common feature that involves a change of clothing for teenage girls. Yi children usually wear red and white dresses, hair braided into one plait and simple earrings (picture 4 a). During the ceremony, a girl’s family will usually find an elegant and prestigious woman from their village to help halve the girl’s hair into two plaits, and then put a “hapa” (a square scarf) on her head (Li, J., 2004). On top of that, the girl has to change her red and white children’s skirt into a colorful pleated skirt to complete the ceremony. Wearing adult attire means that a girl can now participate in social activities and date boys. As for Yi boys, the only thing that starkly marks their transition is the way they knot their “hero’s knot.” People of different ages, from children to the elderly, all have distinct ways of knotting.



Source: History in Costumes: on Costume Culture of the Yi Ethnic Group in Zhaojue. China Ethnic Photographic Art

4a. Yi children's traditional dress

Another example which can show the importance of clothing in Yi culture is the Torch Festival. The festival is celebrated as the grandest traditional feast every year on the 24th of the lunar month of June. There are many explanations regarding the origin of this festival, but the widely accepted version is that the Torch Festival serves to commemorate a successful rebellion against a tyrannical landlord (Li, J., 2004). Yi people

hold numerous exciting activities during the festival, including bull-fighting, tug-of-war and dance.

Included in these celebrations, in Liangshan, there is a special activity to celebrate the festival—they hold a beauty pageant to select the most beautiful woman. Women of various ages from all villages gather at the main venue of the festival where people light roaring bonfires (as shown in 4b). They put on their most exquisite clothes and their most precious jewelry. Contestants compete with each other to show the elegance, delicacy and beauty of their outfits, because the Yi deem these qualities to be evidence of one's deftness, diligence and intelligence. The most beautiful and dexterous girl chosen by a panel of prestigious elderly will be given the title, "Suo Ma Hua," which means "azalea" in the Yi language.

Thus, clothing to women in the Yi culture is a reflection of their abilities, education and social status. Beyond the beauty pageant, Yi people also tend to judge women based on their style of dress in daily life.



Source: <http://www.sc.gov.cn/10462/10464/11716/11718/2013/8/2/10271491.shtml>
(accessed on June 28)

4b. The Torch Festival in Xichang

The present and future of Yi Culture

While many other ethnic groups in China, for example the Han people, are putting on modern clothing and taking off their traditional dress, Yi people still preserve their costume culture very well. Various factors have protected the Yi costume from disappearing.

First of all, as mentioned earlier, Liangshan is very remote and far from the centre of the country. The regions where Yi people live are also mountainous, thus minimizing their cultural contact with other ethnic groups. The isolated location is the most crucial factor that has preserved the distinctness of Yi costume, as well as other traditions of this culture. In addition, because Yi tribe leaders had very centralized, strong power and control over their people, the central government's policies hardly reached

the Yi regions (Huang, 2015). A number of emperors in Chinese history, for example emperor Xiaowen of the Northern Wei dynasty, had tried to unify the clothes that people wore; however, because the tribe leaders were so powerful, they resisted and kept on wearing their traditional clothing. Furthermore, the utility of Yi costume has reinforced its popularity even in modern days (Huang, 2015). Because the Cha'erwa, the wool coat that they wear, is very useful in cold or rainy weather, as well as for doing farm work, traditional clothing remains dominant in remote agricultural regions. However, in more developed cities, the Yi refuse to wear their costumes for the same reason: the coats are too heavy, and the ornaments and dresses are too cumbersome.

With industrialization and modernization, the cost of making Yi clothing is decreasing, so that people no longer have to weave wool fabric by hand or embroider motifs stitch by stitch. This technological development makes Yi clothing more accessible and the production more effective.

Many artists also have applied the characteristic patterns of Yi clothing to fashion design, thus making the Yi dresses more modern, beautiful and popular. These exotic designs are very popular among all ethnicities. Some famous Yi artists even wear modified traditional clothes for public appearances, a way through which the Yi can promote their own culture and spread the influence of Yi costume (see picture 5 a).



Sources:

Left: Cover of Ji Jie's music album

<http://ent.163.com/photoview/3ON60003/520242.html#p=9J1O7IDC3ON60003> (accessed on July 3)

Right: Performance on *The Voice China*, Season 1

5a. Left: Ji Jie, a famous Yi singer originally from Liangshan, wearing the “hero’s knot” on a shoot for a fashion magazine. Right: JieKe Junyi, also a singer from Liangshan, wearing a modified Yi-style dress with Yi patterns.

Nevertheless, despite all the factors that ensure the continuance of the Yi costume, this cultural inheritance still faces many challenges. In order to understand the opinions of people of different age groups, I interviewed three Yis: a 17-year-old high school student who currently studies in Chengdu, a 28 year-old woman who married a Han man, and a 43-year-old woman who lives in Liangshan with her family. Based on the data collected through these interviews, I divide Yi people into three main groups based on their various attitudes toward Yi costume.

The first group includes the younger generations aged from 15 to 30. They only wear Yi costumes for traditional festivals and ceremonies, for example during the Torch Festival mentioned earlier or for funerals and weddings. One reason may be the education they received: because education in Liangshan is not as developed as that in big Han cities nearby, many rich parents usually send their children to Chengdu, the capital of Sichuan province, which is just five hours' drive away. The Yi high school boy I interviewed grew up in Liangshan, but his parents paid a huge amount of money to get him into a private boarding school because, as he said, *"They told me that Chengdu is much more international, open and advanced than Xichang. They think that I would have a much better education and better opportunities in Chengdu than in our corner (suggesting Xichang)."* Like him, many other rich Yi families choose to send their children to Chengdu to receive a better education. These Yi students are totally "Hanized," and there is no way to distinguish them from Han people solely based on their appearances and actions. For those teenagers who stayed in Liangshan, they still face a strong Han influence. The state mandates that schools in provinces where ethnic minorities reside must teach Mandarin. Primary and secondary schools in Liangshan also use the same textbooks as other schools in Sichuan province. When asked about her daughter's preferences on clothing, the 43-year-old woman said: *"Her generation is almost all 'Hanized.' Of course she prefers the trendy and cool clothes of young kids. Nevertheless, she is a good kid, she likes to wear traditional dress for festivals. She thinks it is also very pretty."*

The pattern of accommodation of the parents of the first group may be another reason for their assimilation. Because some of their parents are open to the Han culture and deem it superior, as indicated by what the high

school boy said about his parents, they naturally teach their children more about the advanced Han culture than, as the 28-year-old woman described it, the “*backward*” Yi culture. Also, as will be mentioned later, their parents themselves are to a degree influenced by the Han culture. Thus, in their home, they can’t imprint much Yi influence on their children.

People aged 40 to 60 belong to the second group. These people speak the Yi language fluently, communicate with each other in Yi, and eat Yi food but seldom wear traditional dress. When the Chinese Communist Party assumed control of Liangshan in the 1950s, most of these people were still children. As such, they grew up with very traditional parents but also a wave of Han influence. The 28-year-old woman told me that her parents, both belonging to the second group, speak Yi at home and cook her Yi food but wear modern clothes in daily life. Additionally, because most people of the second group work in offices now instead of on farms, this group chooses to wear comfortable modern clothes, and only wears traditional costumes for special occasions like the first group.

The third group consists of the older generation living in more remote and less developed areas. These people rarely have contact with the Han culture, and primarily do farm work for a living. Their location limits their access to the more comfortable modern clothes; their living conditions also force them to continue wearing the traditional clothes, because, as mentioned earlier, Yi clothing is designed for agricultural work. Thus, only the third group has preserved the Yi’s traditional clothing comparatively well.

When I asked my informants why they still put on Yi costume for festivals, they told me that they have the responsibility to respect, protect

and carry on their traditions. To put it in the words of the 43-year-old woman, “*It’s our Yi festivals, of course us Yi need to wear Yi clothes.*” According to V. Turner (1969), gathering for festivals creates “a direct, immediate, and total confrontation of human identities,” evoking a “concrete, if fleeting experience of community.” Although members of the first and second groups are well adapted to the Han culture, they still feel a lack of confidence in their Yi culture when they are surrounded by Han people, as implied by their interviews. Lynch et al. (1995) explained in their paper that, “Within the context of ritual... the conflicts related to cultural change are expressed and debated, and through dress some degree of resolution is attained.” Therefore, the feeling of *communitas* elicited by congregating and putting on traditional dress temporarily relieves many Yi of the stress and struggles of living in a Han society. Switching clothing is therefore given a fluid social function: to blend in and to be seen, as the 28-year-old lady put it, as “*normal*” when Han culture dominates, and to be proud of being Yi when back in their own Yi community.

On top of the demographic differences, there are some more general manifestations that the Yi clothing is facing challenges of assimilation and disappearance.

Firstly, as mentioned several times before, most Yi consider their culture inferior to the Han culture. Before the Chinese Communist Party assumed control of Liangshan in the 1950s, the Yi still had a slave system. It was a stark and hasty transition for the Yi to adapt to the Party’ socialist system without any economic or political preparation. Therefore, many of their less developed and barbarous behaviors and conventions still exist today. Yi people long for the wealth and technical development that the Han

people have, so they try their best to imitate all, including lifestyle, clothing and language.

Also, the Han culture is the dominant culture in China. Han people are the most prevalent group in Xichang, where 74.37% are Han, according to the regional census of the Liangshan prefecture done in 2010. Han people have increasingly moved to Liangshan in recent years, and many Yi work in companies owned by Han people. Due to this increased contact with the Han people, the Yi have had to put away the traditional clothing and put on the Han clothing in order to be “normal.”

Another common reason why the Yi clothing is disappearing is the popularization of interethnic marriage between Yi and Han. Since the 1950s, the Yi people's opinions toward marriage between Yi and Han have changed drastically. Because they think that Han people are richer and more educated, most of the Yi don't object to interethnic marriage anymore, and nowadays some Yi girls even wish to marry Han boys. Because Yi and Han are both patriarchal societies, when a Yi girl is married to a Han boy, she must follow the Han norms of her husband and put behind almost all Yi conventions. The 28 year-old Yi lady follows popular Han norms of her Han husband: they threw a western style wedding and she wore a western-style wedding gown. At home, she also dresses her child in modern Han clothing and speaks Mandarin to her child. Interethnic marriage expedites the disappearance of Yi clothing because children in these families will have much less contact with the Yi culture, hence will identify themselves more with Han culture than with Yi culture. When asked about her opinions on passing on Yi heritage to her daughter, she said: *“Of course I will tell her that her mother is Yi and she is half Yi. I will teach her Yi history and*

culture. But I think that she will more likely to identify as a Han than as a Yi. I don't blame her. I think it's kinda sad but inevitable, as I hope to raise her in Chengdu instead of in Xichang." It is highly probable that children of interethnic parents will gradually lose the obligation to wear traditional clothes even on Yi special occasions.



Source: Provided by informant

5b. An interethnic couple

Final thoughts

Most Yi people in Liangshan are aware of the fact that modern dress is replacing their own costumes. All my Yi informants lament the effacement, but also admit that this trend is unavoidable. They think that in order for Liangshan to develop, adaptation to the Han culture is the most

efficient and the only way. Only by embracing the Han culture, speaking fluent Mandarin and making friends with Han people can they gain access to better education, business opportunities and political power. No matter how much the government tries to advocate for integration and unity, inequality and the feelings of superiority and inferiority will continue to exist among Han and Yi people.

As for the future of the Yi costume, the older generations (people older than 40) are more pessimistic than the younger generations (people younger than 40). The 43-year-old woman deplored that: “[Her daughter] *She belongs to the younger generation, and its way of thinking and living is far different from my generation’s. When she grows up, she’ll probably prefer the ‘cooler’ Han culture to the Yi culture.*” On the contrary, the two young Yis that I interviewed claimed that, although they think that their children will be more Han than Yi, they will still teach them about the Yi culture (e.g history, conventions) and celebrate Yi traditional festivals with them. Though very “Hanized,” the younger generations still feel a sense responsibility to make sure that their descendants know of their Yi origin.

Since the advent of the Han culture in the 1950s, it has taken approximately only 5 generations for some Yi people to be completely “Hanized.” Unfortunately, due to corruption and incompetence, the Liangshan local government hasn’t spent much effort and money to protect the Yi’s intangible heritage. Therefore, if the Liangshan government along with the Yi people don’t take initiatives in protecting their own cultural treasure, the future of the traditional clothing would really be at stake, and it would only be a matter of time before we saw Yi costumes only in museum exhibitions.

Bibliography

Cheng, Huibo. *Research and Development of Yi Costume Culture*. Beijing Institute Of Fashion Technology, 2007

Huang, Jing. On the Factors of Transformation and Inheritance of the Yi People's Costumes. *Journal of Guizhou University for Ethnic Minorities*, Oct. 2015

Ling, Fuxiang. Yi Ethnic Group's History and Present. *China & the World Cultural Exchange Journal* vol.1: 40-41 1997

Li, Ding. The Coming-of-age Ceremony of Yi Teenage Girls. *Vision Magazine* vol. 3 2012

Li, Jin, ed. *History in Costumes: on Costume Culture of the Yi Ethnic Group in Zhaojue*. China Ethnic Photographic Art Press, 2004

Lynch, Anette, Detzner, D., & Eicher, J. Hmong American New Year Rituals : Generational Bonds through Dress. *Clothing and Textiles Research Journal*, 111-120. 1995

Ma, Shan. The Symbol of Culture and History: the Costumes of the Yi Ethnic groups in Liangshan. *Journal of Sichuan University for Nationalities*. vol.17:34-38 2008

National Bureau of Statistics of China. *Communiqué of the National Bureau of Statistics of People's Republic of China on Major Figures of the 2010 Population Census*, 2011

<http://web.archive.org/web/20131108022004/http://>

[www.stats.gov.cn/english/
newsandcomingevents/t20110428_402722244.htm](http://www.stats.gov.cn/english/newsandcomingevents/t20110428_402722244.htm)

Turner, Victor. *The ritual process: Structure and anti-structure* (p. 132).
London: Routledge & Kegan Paul, 1969

Wang, Changfu. The differences of dressing styles in Liangshan's three
dialectical areas. *Journal of Liangshan University*, 86-89 2000

Zhu, Wenxu. The Polarized Rank Problem in Black and White of Yi People
in Liangshan. *Journal of Bijie University* vol.31: 28-32 2013

*All the resources above except for the Communiqué of the National Bureau
of Statistics of People's Republic of China on Major Figures of the 2010
Population Census, and the article by V. Turner and A. Lynch et al. are
translated from the original Chinese



Is Greece's Crisis in the Eurozone a "Tipping Point" in the Development of the European Union?

The Challenges and Opportunities in the Greek Crisis and the Fate of Europe's Integration

Sibei Wu

Author background: Sibei grew up in China and currently attends The High School Affiliated to Renmin University of China in Beijing. Her Pioneer seminar program was titled "European Integration in Comparative Perspective."

I. Introduction

1.1 Abstract: Purpose of the Research

"Tipping point" may be defined as the critical point in a situation, process or system beyond which a significant and often unstoppable effect or change takes place. In this analysis, "tipping point" refers to the dilemma Greece faces. What are some of the questions the Greek crisis in the Eurozone may raise? How did the Greek sovereign debt crisis begin? How will the Greek crisis evolve? How will the crisis influence the future of European Union member states as well as the continent as a whole? How should citizens in European countries understand the crisis?

My research aims to: 1) understand Greece's situation in the Eurozone through an analysis of history and economy; 2) explain the evolution of the Greek crisis from the perspectives of both Greece and the European Union, and 3) analyze challenges and opportunities introduced by the crisis as well as the future of Europe's integration.

1.2 European integration from an economic perspective

European integration defines how countries, starting with the European Coal and Steel Community (ECSC) comprising France, Germany, Italy, Belgium, the Netherlands and Luxembourg, came together after World War II. When analyzing European integration, scholars focus more on economic integration, which after 1954 assumed a leading role in bringing the Original Six closer in the European Economic Community (EEC) and EURATOM, forerunners of the present day European Union (EU). Supranational policies are made in the economic area. Before the years leading up to the Greek crisis, stable and booming economies represented success in the European Communities development. Greece's inability to maintain its economic stability, however, is perceived as a reflection of integration's failure. As a result, when Greece confronts the debt crisis as a member of the Eurozone, the Union must take action to help Greece in order to retain EU prestige and solidarity. In other words, Union member states, as a whole, face the potential threat of the enduring crisis. In the present context, the question of how to address Greece's economic difficulties is the primary challenge of Eurozone integration.

1.3 The Present Debt Crisis – The Past Five Years (Greece in the Eurozone)

Greece has suffered through a vicious circle of recession over the past five years. The Greek economy experienced a series of successive setbacks during this period. Since 2010, the Greek economy contracted more than 20 %. The investment level decreased by 86 % as compared to 2008. (IMF2013b) The continuous drop in GDP, which in 2011 surpassed the historical maximum for the entire postwar period, led to a rapid decline in

domestic demand. Unemployment more than doubled within the first three years of austerity reaching 25.4% in August 2012. More than half the population between 15–24 years of age, or 57%, is unemployed (Eurostat 2012); homelessness increased by 25% from 2009 to 2011 as suicide rates hit record levels, increasing 25% from 2009-2010 and an additional 40% from 2010-2011. (Reynolds 2015) In early 2014, unemployment exceeded 28%; youth unemployment rose to 63%, which led young, educated people to emigrate in large numbers. (Visvizi 2014) Greece's uncertain future remains a troubling predicament on the periphery of the European Union.

II. Background

2.1 The Formation of the European Community/Union: Integration's Twentieth Century Narrative – 'No More War'

As an advisor to Franklin Delano Roosevelt during World War II, Jean Monnet, the former League of Nations official, first conceived of the idea of pooling the resources of France and Germany in the sectors of coal and steel. This functional integration aimed to construct Europe as a community on a solid foundation by making war on the Continent impossible. Technocratic and administrative measures were used by elites to convince societies that European integration was in their best interest. Monnet then led the High Authority as its first President in the newly created European Coal and Steel Community, which was declared to be "a first step in the federation of Europe" (Declaration of 9 May 1950) in the Schuman Plan. Initially hardly visible to the populations in the Original Six member states, integration by sectors was eventually supposed to spill over into high politics, namely, security and defense, establishing economic solidarity and political unity. The functional approach fit the era of "fashion for technocracy," when

citizens did not want to discuss “National Socialism.” Projects were informed by the personal experiences of the Founding Fathers including Adenauer, DeGasperi, Schuman, Monnet, and Spaak. Since each leader advocated federalism, decentralization, and subsidiarity, the need to safeguard and enhance the supranational European legal order, which had emerged since the Treaty of Rome in 1957, increased. The European Economic Community (EEC) was thus created.

In 1958, General DeGaulle, a political leader and a realist, returned to power in France as the first President of the Fifth Republic. A strong opponent of supranational integration, with its emphasis on European institutions, DeGaulle opposed the EEC as a challenge to the sovereignty of France. Instead he envisioned a largely intergovernmental structure designed to coordinate the foreign policies of the EEC member states. Throughout the 1960s, tensions between France and the other members in the Community became serious as DeGaulle aimed to limit supranational power. However, in 1967, the Merger Treaty created a single set of institutions, consisting of the three Communities, which were collectively referred to as the European Community (EC). In 1974, the Community reverted to a traditional intergovernmental style of decision-making with the creation of the European Council. In 1979, a French-German bilateral initiative led to the creation of the European Monetary System (EMS) a forerunner of the Economic and Monetary Union agreed to in Maastricht in December 1991.

2.2 Greece joins the European Community: Consolidating Democracy after the Colonels’ Military Regime

Greece’s membership in the EU has long been defined by the issue of the country’s ‘return to democracy’ with its accession as the tenth member

state of the European Community in 1981. (Visvizi 2013) As the 'colonels' regime' ended, Greece joined the Community because accession was essential to consolidating democracy in the country. (Verney 1987; Valinakis 2012) After signing an Association Agreement with the EEC in 1961, Greece was trapped in a frozen period lasting from 1967-74. However, once the New Democracy government was established in Athens in 1974, the Association Agreement resumed immediately. (Tsinisizelis 2008: 14) As a result, Greece established its relationship with European Union in 1981, becoming the 10th member state.

2.3 The Relationship between Greece and Europe: A Difficult Relationship for a "Footnote State" on the Community's Periphery

During the first year of its accession to the European Community, Greece demonstrated little interest in cooperating with other member states in the European integration process. Rather than actively promote its national interest in the Community, Greece fulfilled its obligations as a member state. At this time, Greece focused almost exclusively on the intergovernmental and bilateral fora of decision-making. (Visvizi 2013)

In the mid-1980s, Greece shifted its attitude toward the European Community. The Greek government showed a willingness to engage constructively with the Community since the domestic political resources available to the PASOK government were limited. However, longstanding cooperation did not anchor the relationship. Greece had to adopt a passive role at the Community level due to its political instability during the period from 1988-91. (Visvizi 2013)

Greece ceased to play the role of the flank in the Western democratic world as communism ended in the former East Germany in 1989. With the disintegration of the former Yugoslavia, and specifically the ‘Macedonian problem,’ Greece waved goodbye to marginalization and isolation on the European scene. (Visvizi 2013)

The Greek government urged active, though not vital, participation with a ‘federalist’ orientation during the negotiations leading to the Treaty of Amsterdam, June 1997, and in the run-up to Agenda 2000. By strengthening its image as a trustworthy and moderate member state, Greece focused on improving its low level of economic development. This aim provided the motivation for the Greek government’s attempt to increase economic aid given to those regions in need from the European budget. Its moderation allowed Greece to succeed in transforming its own passive image while also altering the submissive attitude of its European partners. This approach enabled Greece to be considered as a candidate and eventual member of the Eurozone during 2001. (Visvizi 2013)

The sovereign-debt crisis, however, pulled Greece down from this rising trend. “A fake promise of modernization” and “a warning about the perils of Europeanization without deeper transformation” (Bechev 2012: 6), once again made Greece the subject of criticism. The words ‘lazy,’ ‘profligate,’ ‘irresponsible,’ ‘undisciplined,’ ‘disorganized,’ ‘chaotic,’ ‘corrupted,’ ‘deceitful,’ ‘violent,’ ‘resentful,’ and ‘troublemakers’ were at once associated with Greece. (Bleich 2012; Prokopijević 2010; Rosenthal 2012)

At the beginning of the Eurozone crisis, Greece ignited the EU’s future debates regarding economic governance and the ways to address the crisis in the Eurozone area, thereby placing itself in a marginal position. Although a

longstanding member of the European Union, Greece was trapped and unable to influence the discussion about Europe's future. Greece failed to achieve greater influence on the range of policy issues that shaped Europe's debate, led by Germany, despite its sporadic domestic level articulation regarding the direction and shape of the European integration process. (Visvizi 2013)

Blaming Greece and the Economic Periphery

Since the onset of the crisis in 2009, European member states and institutions showed a lack of satisfaction with Greece's predicament. Countries with accomplished economies were regarded as conscientious member states with taxpayers who work hard, receiving low, yet steady, salaries over the years. However, these laborers also paid for labor in Greece. This is because of the accumulation of southern debt as well as EU-administered bailouts. In headlines such as "Greece dependent on the patience of German taxpayers," the media pointed to Greece's responsibility to survive on its own. (Costello 2012) Greece was portrayed as the root cause of the economic crisis impacting the Eurozone. As such, Greece was persecuted by the outside world. Core-periphery animosity and resentment once again intensified. Scarce resources and persistent antagonisms inside the country were at once recalled, which opened space for the possibility that Greece would not be able to maintain a stable cooperation with other member states. This scenario was referenced by Euro-skeptics in Greece and the Union. In this context, Greece stands clearly on the periphery of the Eurozone.

III. The Greek Debt Crisis

3.1 From Greece's Perspective

This section aims to discuss the root of two aspects of the Greek debt crisis.

1.) How did the development of its financial system pave the way to Greece's economic plight?

Structural weakness

In the case of Greece, the failure to establish a structural balance between the rule of law and democracy before 1974 affected the functioning of the state by decreasing the potential to implement laws effectively, thereby curbing the desire to reform. In the 1980s, the development of a pervasive political system was triggered by the rising principle of the rule of law and democracy. The system, however, proved susceptible to the abuses of nepotism, cronyism, and the improper use of executive authority. The ills of Greece's domestic political system undermined the quality of the relationship between Greece and the European Community. This hampered the impact of the Community on Greek policies and unremittingly raised skepticism about Greece's virtues. Even today, the Community's political system affects the Greek government's ability to embark on reforms and leaves the Greek political system vulnerable. (Visvizi 2013)

During the year 2007, Greece maintained its stable economic growth as one of the only EU countries to avoid the negative impact of the global financial crisis. As a result of its marginal exposure to toxic assets and the stimulus plan enacted by the government, Greece was affected only

indirectly. However, the secondary, domestic, and international implications of the economic downturn struck Greece hard.

Domestically, it was harder for the Greek government to refinance the gross percentage of the debt. The collapse of the US housing bubble led to a global liquidity crunch. It should be noted, though, that in 1981, the gross government debt to GDP ratio was only 25%; this percentage increased dramatically in the 1980s stabilizing at about 100% of GDP in the 1990s. Following Greece's entry into the Eurozone in 2001 and the resulting access to cheap loans, Greece's debt as a percentage of GDP increased gradually to reach 110.7% in 2008 and 127.1% in 2009. (EC 2011b)

Internationally, lower consumption and investment levels led to inflation, which lessened export opportunities. As a result, its competitors, such as Turkey and Spain, replaced Greece's domination of the tourism industry. Revenues from tourism fell to €10.4 billion in 2009, marking a decrease of 11% as compared to 2008. With the tourism industry's contribution to the GDP estimated at 15.2% in 2009 and with the tourism multiplier valued at 2.184, this drop had a substantial impact on the entire Greek economy. (SETE 2010: 18–25)

Facing pressures from both inside and outside the country, the Greek government failed to modify the proper economic structure or adopt immediate actions to minimize the possibility for loan default. This scenario brought Greece to the brink of insolvency.

2.) Did the elite lead the country and its leaders into crisis?

Leadership failure

The Greek government's position regarding the global financial depression was to emphasize the protection of deposits in all Greek banks, specifically with a €28 billion bank liquidity scheme. The scheme included capital injections by way of sale of preferred shares to the State, guarantees on debt issuance by banks, and liquidity support via special government bonds. (Brown, 2009) Following the European Economic Recovery Programme (EERP), approved by the European Council on December 11-12, 2008, the Greek government worked hard to reform the economic structure in order to shelter Greece from the financial storm brought on from other parts of the world. The governmental leadership aimed to maintain a stable, safe, and long-term economic boom for the country while strengthening its legitimacy in the eyes and hearts of the citizenry.

Yet, ongoing conflict between the parties resulted in the plan's stagnation. The Conservative Party *Nea Demokratia* (ND) regarded austerity measures and economic structural reforms as the only way to pull Greece out of difficulty. The Opposition Movement (PASOK) insisted on more expenditure, such as spending money on subsidies. After PASOK was in power, too much emphasis was placed on revenue-enhancing measures, such as levies for profitable enterprises, luxury goods (yachts, swimming pools), private income above €60,000 as well as presumptive taxation (Hellenic Stability and Growth Programme SP2010). Too little emphasis was placed on the expenditure-reducing measures, for example, downsizing the public sector and reducing other expenditures. (Visvizi 2012)

“Rather than tackling the roots of the problem, Greek Prime Minister Papandreou initiated a strategy aimed at convincing the political establishment in the EU and the US that Greece had fallen prey to massive speculative attacks. Papandreou argued that financial regulation should be strengthened to prevent market manipulation.” (Visvizi 2012) Shifting the responsibility for the dilemma to the previous government, Papandreou even regarded himself as a responsible politician. It is, therefore, understandable how a once prosperous EU member country evolved into a failing state under the leadership of one man.

3.) How did domestic expenditure drive Greece to this quandary?

High military expenditure

Greece had been spending far more on its military than the average in other European states for several decades. According to NATO's calculations, the average expenditure of the European NATO states in 2009 was 1.7% of GDP whereas Greece spent 3.1% of its GDP on armaments. If one compares the overall military expenditure of the original member states from 2000 to 2008 on the basis of the Stockholm International Peace Research Institute (SIPRI)'s figures, the average is 1.6% of GDP whereas Greece spent 3.6% of its GDP on the military during the same period. (Jan Grebe and Jerry Sommer BICC Focus)

3.2 From the EU's Perspective

The crisis should also be viewed from a panoramic perspective in the EU context.

1.) Weakness of structural economic support

Dating back to the founding period in the 1970s, flaws in design were buried in the later development of the European Monetary System (EMS). In fact, the present crisis was predictable. The member states in the Eurozone had very different fiscal and economic policies. This context explains why they failed to constitute what economists define as an Optimal Currency Area (OCA). Booms and busts are two results of the free market. During the boom period, a country needs flexibility to push down rising prices and wages in order to return to the balancing point with its trading partner. Countries with sovereign monetary policies, nevertheless, simply print more money and devalue their currencies. (Woods n.d, 12-13)

Additionally, fiscal integration, labor mobility, and wage decreases are essential parts of a stable economy. Yet, politicians are loath to curtail the wage within their constituencies owing to political considerations. As Paul Krugman states, “Europe isn’t fiscally integrated: German taxpayers don’t automatically pick up part of the tab for Greek pensions or Irish bailouts. And while Europeans have the legal right to move freely in search of jobs, in practice imperfect cultural integration – above all, the lack of a common language – makes workers less geographically mobile than [for example] their American counterparts.” (2011)

In other words, loose convergence criteria and the Stability and Growth Pact (SGP) cannot stop state fiscal policies and economic cycles from diverging. In 2002 and 2003, the Stability and Growth Pact lost its efficiency and power. While the French and German central bank governors preferred lower interest rates at that time, they “were easily outvoted by a coalition of small states and the ECB Executive Board German and French governments had to run public deficits in this period, and hence break the

rules of the Stability and Growth Pact.” (Hix & Hoyland 2011, 263) Breaking the rules allowed Greece to enter the Eurozone when “in reality, it had not met the . . . convergence criteria of 3 per cent of GDP ceiling on the government deficit.” (Featherstone 2011, 199) The Greek public debt level has also been consistently high: “it has fluctuated around the equivalent of 100 per cent of GDP since 1993” even though Greece did not join the Eurozone until 2001. (Featherstone 2011, 198) Greece was then led to a worse financial disaster with high debt as well as periods of high inflation. (Woods n.d, 8-9)

Moreover, the 2008 banking crisis exacerbated the structural problem, which reflected the vulnerability of Greece's economy. Although the Eurozone was not an Optimal Currency Area, the project went on because “the idea of the euro had gripped the imagination of European elites. Except in Britain . . . political leaders throughout Europe were caught up in the romance of the project, to such an extent that anyone who expressed skepticism was considered outside the mainstream.” (Krugman, 2011) As Jeffrey Miron of Harvard University writes, “Greece will never change its misguided policies if the EU and IMF infuse it with new cash, just as no teenager who has overspent an allowance will reform if the parents merely expand that allowance.” (Woods n.d, 9-10)

2.) Political leadership and crisis management.

Europe's economic policy has been shaped by political will from the beginning of the EMU. “While [Frenchman Jacques] Delors' aim was to design a project that would be irreversible, the governor of the German Bundesbank, Karl Otto Pöhl, wanted to be certain that the single currency would be as stable as the *Deutsche Mark*, arguing for constraints on national

deficits and a fully independent European Central Bank.” (Hix & Hoyland 2011, 250)

The Stability and Growth Pact was also based on a political compromise, aiming to assure a stable relationship between member states and European preferences. Had the founding states of the Economic and Monetary Union not been weakened in 2002-03, the Pact would have been strict enough to support economic convergence. Even if Greece intentionally misrepresented its public debt and deficits, the European Union did not ensure the report’s authenticity, thus benefiting Greece by allowing its entry into the EMU. (Woods n.d, 14-15)

Some political leaders, particularly German Chancellor Angela Merkel, have been criticized for their indecisiveness in granting a bailout to Greece thereby igniting a severe euro crisis. (Pidd, 2011) As an election on May 9, 2010, was fast approaching, Merkel’s government was under domestic pressure not to “give away” German money to Greece. However, “that is only half the truth. Merkel is also afraid of German voters. Rarely has Europe been as unpopular with German citizens as it is now, and getting people to accept the need for unpopular policies has never been Merkel’s strong point.” (Müller et. al., 2011)

Obviously, the policy of austerity imposed on Greece by its creators has done more harm than good. Greece should be allowed to either temporize its debt or exit the Union. The negotiations, however, simply go around in circles because the members have no focal point. Member states point the finger of blame at one another, divide Europe into North and South and do not feel like a community helping each other in need, hence “driving away the mistakes they made, away from the symbiotic economic interconnections

between the Eurozone countries and away from the structural problems that cannot be specifically blamed on Greece." (Guérot 2015, URL)

IV. Challenges and Opportunities in the Greek Crisis

4.1 Internal (Domestic – Greece)

The Greek government faces a number of main challenges.

The challenge of restoring growth

There are two points people should keep in mind. First, Greece's economy is in recession and its political system is in crisis. Second, the crisis is the result of the fiscal adjustment program defined in the Memorandum of Understanding (MoU), which did not tackle the major problems of the Greek economy. Instead of limiting the size of the public sectors and introducing reforms, such as privatization and liberalization, the fiscal adjustment only channeled the private sector by increasing taxation and introducing tax measures. Thus, the methods discussed above were detached from Greek's income and the ability to pay for the debt. (Visvizi 2012)

In order to maintain political stability, economic predictability, and foreign investment, the Greek government requires such methods as decreasing taxation, privatization, downsizing the public sector, deregulation, and liberalization. However, what challenges the Greek government is "the climate of social dissatisfaction, disenchantment with [a] so far fruitless reform process, and severe opposition from specific stakeholders that seek to protect their vested interests." (Visvizi 2012)

The challenge of domestic turbulence

Sources undermining domestic stability exist in Greece. SYRIZA declared its position as an active opposition. On the one hand, “the 71 parliamentarians of SYRIZA will seek to block any law proposal by the new government. On the other hand, SYRIZA is likely to organize massive protests, demonstrations and strikes, which could once more paralyze the country.” (Visvizi 2012) The Independent Greek Party is also likely to impact the reform drive “due to a dynamic characteristic of new spontaneously formed groupings and due to the anti-bail-out rhetoric that they employ.” (Visvizi 2012) The Golden Dawn Party, too, is able to stir the waters of political debate in Greece. (Visvizi 2012) Therefore, how stable, how consistent, and how effective is the coalition government?

Venizelos, the leader of PASOK, faces the challenges of consolidating the government and rejuvenating his party. PASOK used to have 44% of popular support, standing in stark contrast to today’s daunting 12.28%. As a result, PASOK risks being swallowed by the greater government while being exposed to the critique of the opposition. (Visvizi 2012)

The challenge of reestablishing Greece’s credibility and prestige on the European and international scene

Greece’s negative publicity was affected by two issues. On the one hand, the recessing fiscal tide of Greece and its inability to fulfill the obligation to pay its debts shattered Greece’s credibility. On the other, several politicians made misleading statements describing Greece as a country of tax dodgers, thus ruining Greece’s image and credibility. Tax evasion was presented as the main cause of Greece’s economic crisis. In a

word, it is a challenge for the Greek government to restore its credibility and reverse its negative image. (Visvizi 2012)

4.2 External (Regional – European Union)

The results of the parliamentary elections in Greece create a number of opportunities for the Eurozone and the Union.

The tide for growth

After the electoral success of ND, its leader, Samaras, sought communication and consensus with other countries in the European South, including Spain, Portugal and Italy. In the spring of 2012, during a series of visits to Portugal and Spain, Samaras discussed the challenges, similarities, and differences behind the economic problems and the opportunities these countries could develop together. In the following visit to Italy, both leaders emphasized the necessity to help each other with economic growth. Monti, the leader of Italy, expressed his willingness to hold a multilateral summit in Rome. In other words, Monti began to cut into the core of Europe's problem, seeking to build a new coalition for Europe beyond the weakened Franco-German tandem. Such strategy would build a consensus for the future of the Eurozone, deepen European integration, and open up the space for solutions. All in all, the failures in handling the debt crisis can provide an incentive for member states' leaders to reinvigorate the discussions on growth in Europe as well as search to create more cooperation with periphery countries. (Visvizi 2012)

Maintaining stability as a possibility

It can be argued that following the establishment of the coalition government under Antonis Samaras, the Eurozone will maintain considerable stability. Once a climate of stability returns to the Eurozone, the opportunities created for EU leaders to discuss the future of economic cooperation and strengthen the Union's economic governance may proliferate. With such opportunities, more communication among member states can also lessen misunderstandings, thereby strengthening close binding ties within the Union.

4.3 Impact on the Future of European integration

Since the founding of the European Coal and Steel Community (ECSC), economic integration has been the glue of the European construction process. From the Greek crisis, we see that the illnesses of the political system in Greece are in essence a reflection of the illness of the Greek economic system. The conspicuous economic disparity was also reflected in the crisis. With the bailout, benefactors complain about Greece's inferior ability to sustain a stable economy. The recipients complain about the domination of economically strong member states which cut off all opportunities. The blame within member states has led people to worry about whether economic integration will still work in the Eurozone. Without doubt, more economic cooperation and communication are strongly urged. To seek and share opportunities among the member states as well as European institutions becomes the main mission in the future of integration.

V. Conclusion

Why has Greece been drawn into to economic crisis? What are the impacts of the Greek crisis?

This analysis concludes that five reasons explain the Greek crisis: 1) economic and political domestic structural weaknesses 2) a failure of leadership to carry out efficient methods to deal with the crisis 3) an unreasonably high military expenditure 4) flaws in the EU economic structure and 5) a lack of conformity between EU leaders and crisis management. Three challenges result: 1) restoring growth 2) handling domestic turbulence 3) finding ways to reestablish Greece's credibility and prestige. Two opportunities emerge at this stage: economic growth and stability within the Eurozone.

Although the Greek crisis had a great many negative impacts on Greece and the other EU member states, the lessons observers should learn from the crisis are even more important. Economic integration is always the glue to maintain European integration. Thus, the search for more cooperation between member states should be placed as a high priority for the future of European integration.

References

Bechev, D. (2012). *The Periphery of the Periphery: the Western Balkans and the euro crisis*, Policy Brief 60, European Council on Foreign Relations (ecfr).

Bleich, E. (2012). *Say goodbye to PIGS and GIPSIs*. [online] Aljazeera.com.

Available at:

<http://www.aljazeera.com/indepth/opinion/2012/12/201212392653337846.html>. [Accessed 29 Aug. 2015].

Brown, M. (2009). *Global Financial Markets Initiative - Experience -*

Mayer Brown. [online] Mayerbrown.com. Available at:

<http://www.mayerbrown.com/FinancialMarkets/article.asp?id=7849&nid=9774>, retrieved 19.05.2011. [Accessed 8 Sep. 2015].

Costello, P. (2012). *Greece dependent on the patience of German taxpayers*.

[online] The Sydney Morning Herald. Available at:

<http://www.smh.com.au/opinion/politics/greece-dependent-on-the-patience-of-german-taxpayers-20120619-20m0b.html#ixzz2RHYbqSiC>. [Accessed 6 Sep. 2015].

EC (2011b). European Economic Forecast Spring 2011.

Featherstone, K. (2011). The JCMS Annual Lecture: The Greek Sovereign

Debt Crisis and EMU: A Failing State in a Skewed Regime*. *JCMS: Journal of Common Market Studies*, 49(2), pp.193-217.

Grebe, J. and Sommer, J. (n.d.). Greece: High military expenditures despite the financial crisis. *BICC Focus*, p.1.

Gu érot, U. (2015). *Europe as a republic - Ulrike Gu érot The story of Europe in the twenty-first century*. [online] Eurozine.com. Available at:

<http://www.eurozine.com/articles/2015-07-10-guerot-en.html> [Accessed 2 Sep. 2015].

Hix, S. and Hayland, B. (2011). *The political system of the European Union*. Basingstoke: Palgrave Macmillan.

IMF. (2013b). IMF Executive Board Concludes 2013 Article IV Consultation, Public Information Notice (PIN).

Krugman, P. (2011). *The Road to Economic Crisis Is Paved With Euros*. [online] Nytimes.com. Available at: <http://www.nytimes.com/2011/01/16/magazine/16Europe-t.html?pagewanted=all> [Accessed 5 Sep. 2015].

Müller, P. and Schult, C. (2011). *Rudderless EU: Chancellor Merkel's Dangerous Lack of Passion For Europe - SPIEGEL ONLINE*. [online] SPIEGEL ONLINE. Available at: <http://www.spiegel.de/international/europe/0,1518,775085,00.html> [Accessed 3 Sep. 2015].

Ntampoudi, J. (2014). The Eurozone crisis and the politics of blaming: cases of Germany and Greece. pp.1-20.

Pidd, H. (2011). . [online] the Guardian. Available at: <http://www.guardian.co.uk/business/2011/dec/02/angela-merkel-eurozone-fiscal-union> [Accessed 3 Sep. 2015].

Prokopijevic, M. (2010). Euro Crisis. *PANOECONOMICUS*, pp.369-384.

Reynolds, L. (2015). *The Greek Economic Crisis, The Social Impacts of Austerity. Debunking the Myths*. [online] Global Research. Available at: <http://www.globalresearch.ca/the-greek-economic-crisis-the-social-impacts-of-austerity-debunking-the-myths/5431010> [Accessed 31 Aug. 2015].

Rosenthal, J. (2012). [online] Available at:

<http://www.worldaffairsjournal.org/article/germany-and-euro-crisis-powerhouse-really-so-pure> [Accessed 1 Sep. 2015].

SETE, (2010). Greek Tourism 2010: Proposal for the New Development Model.

SP, (2010). Report to the implementation of the Hellenic Stability and Growth Programme and Additional Measures in response to Council Decision.

Tsinisizelis, M. (2008). Greece in the European Union: A political/institutional balance sheet. pp.145-158.

Valinakis, Y. (2012). Greece's European Policy Making. *Greece Paper*, No.63.

Vernardakis, X. (2012). From Mass Parties to Cartel Parties: The Evolution of the Structure of Political Parties in Greece through Changes in their Statutes and Systems of Financing. *Working Paper Series on the Legal Regulation of Political Parties*, No. 27.

Verney, S. (1987). Greece and the European Community. *K.Featherstone and D.K.Katsoudas (eds.)*.

Visvizi, A. (2012). The crisis in Greece and the EU-IMF rescue package: Determinants and pitfalls. *Acta Oeconomica*, 62(1), pp.15-39.

Visvizi, A. (2012). The June 17 Elections in Greece: Domestic and European Implications. *Policy Paper*, No.31, pp.4-9.

Visvizi, A. (2013). Democracy or demagogy? The Greek political actors on the sovereign debt crisis. pp.4-6.

Visvizi, A. (2014). From Grexit to Grecovery : The Paradox of the Troika's Engagement with Greece. *Perspectives on European Politics and Society*, 15(3), pp.335-345.

Wind, S. (n.d.). Eurozone Sovereign Debt Crisis. *SSRN Electronic Journal*.

Woods, S. (n.d.). The Greek Sovereign Debt Crisis: Politics and Economics in the Eurozone.



**Women Uniting on the Web:
Social Media Contributions Against Gender-Based Violence
in Turkey**

Ashley Zhang

Author background: Ashley grew up in the United States and currently attends Crean Lutheran High School in Irvine, California. Her Pioneer seminar program was titled “Women’s Rights in the Middle East and South Asia.”

Author’s Note:

I would like to express my very great appreciation to Professor Sylvia Maier for being an engaging and motivating mentor who guided with enthusiastic encouragement and helpful critiques of this research work. I would also like to thank Pioneer Academics for this research opportunity.

This paper is dedicated to the girl I witnessed experiencing physical abuse from her father and male relatives on the streets of Linfen, China.

**Women Uniting on the Web:
Social Media Contributions Against Gender-Based Violence in Turkey**

Ozgecan Aslan lost her life while taking the bus home when the bus driver “allegedly bludgeoned her with a crowbar, stabbed her to death, and cut off her hands to hide the evidence” (Khazan 2015, 1) because she

resisted rape. According to a UNICEF report, “globally, women aged between fifteen and forty-four are more likely to be injured or die as a result of male violence than through cancer, traffic accidents, malaria and war combined” (Vlachova 2005, 5). Gender-based violence (GBV) is a global problem. GBV includes any act of verbal or physical force, and life-threatening actions that may result in “physical, sexual or mental harm or suffering to women” (United Nations, Declaration on the Elimination of Violence against Women). GBV includes legal exploitation, forced abortion, forced marriages, child marriages, hymen repair, genital mutilation, abduction for marriage, rape, honor suicides, and honor killings at the most extreme.

My paper will illustrate the issues that are prevalent in Turkey in regard to gender-based violence towards women, share stories of victims and survivors, and analyze the way women’s rights activists are using social media to raise awareness and eliminate unjust customs. I approach this project from an emancipatory research perspective from which I will share and promote opportunities for others who read my paper to engage with and support the women in Turkey in their fight for their rights, especially regarding domestic violence and honor killings.

The paradox of Turkey—a country with one of the most comprehensive sets of laws to protect women’s rights, yet also a country with one of the highest recorded numbers of GBV—prompted me to investigate the cases of violence there. A particular incident that occurred with a college student, Ozgecan Aslan, also drew my attention to how serious GBV is in Turkey. Aslan is in my age range, and the fact that such savagery was committed on a young student stirred me. Furthermore, the

social media movement started by women activists that arose after the incident impelled me to contribute to the effort to raise awareness of GBV in the hope of eventually ending it.

I will argue that anti-GBV activism creates the greatest impact when done through social media, such as Twitter and Facebook, which both allow concise and precise status posts and updates that efficiently convey a message. Additionally, I will argue that Twitter—with the use of hashtags—has been a watershed to quickly mobilize a group of activists as well as empower women worldwide and pressure the public to talk about GBV in Turkey. In this paper, I will examine how women’s organizations in Turkey fight for women’s empowerment using social media. Specifically, I will focus on their struggle against gender-based violence, domestic violence, and honor killings.

My project will seek to answer the following questions: How are women activists using social media as an outlet to revolutionize the perceptions of domestic violence in Turkey? What methods and techniques have been used on social media to bring a global perspective to this issue? How effective has the use of social media been in raising awareness, promoting mobilization, and reducing violence? In what ways can others contribute to terminating domestic violence and honor killings?

Social media facilitates and provides a sense of community, emotional support, collaborative endeavors, and a forum for the exchange of knowledge and ideas. As such, it is a great power and can connect people from all around the world who share the same passions—for example women’s empowerment. Women activists from all parts of the globe engage with each other on social media to increase understanding of the brutalities

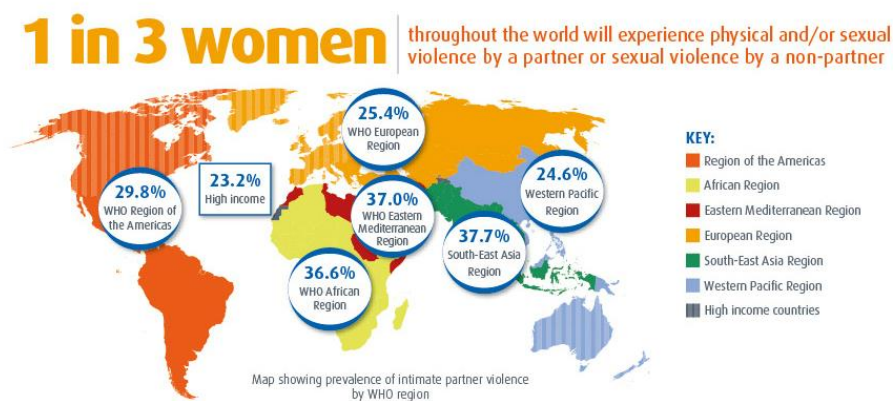
that occur on a daily basis in Turkey and elsewhere. Social media is also strategically used in Turkey to initiate boycotts, protests, and demonstrations.

In section one, I will discuss the prevalence of GBV worldwide and provide key statistical data. I will also focus on the situation of women, women's rights, and GBV in Turkey. Section two will discuss how social media activism is being used to reduce GBV in Turkey with emphases on the Ozgecan Aslan case and social media strategies. Section three will include a progress report on the effectiveness of social media and how it has played a role in improving GBV situations in Turkey plus a call to action with my own thoughts, solutions, and opportunities to help reduce GBV. My research includes browsing and analyzing anti-GBV social media posts and profiles from Twitter, Facebook, Tumblr, and blogs. I supply examples of those activism accounts, as well as case studies drawn from conducting desk research. My desk research includes examinations of news articles from both Turkey and the United States: The New York Times, The Los Angeles Times, The Huffington Post, Today's Zaman, and Haaretz, to name a few. I also analyzed web-based media interviews and documentary films, such as the short film entitled *Dying to Divorce*. In addition, I have interviewed specialists on GBV and women's rights in Turkey such as Turkish anti-GBV activist Goncagul Ay. The purpose of my project is to explore the social media methods women's rights activists, especially those in Turkey, are utilizing to foster auspicious change, so I can inspire others with my findings to join the movement and reduce occurrences of domestic violence in their communities and worldwide.

Gender-Based Violence: A Global Scourge

Gender-based violence is a global problem that has exceptionally harmful effects on women, both physically and psychologically. The consequences for a victim of GBV may include fatal outcomes such as homicide and suicide; physical health damage such as injuries, permanent disability, HIV, and unwanted pregnancy; mental health problems such as depression, anxiety, and eating disorders; and negative health behaviors such as smoking and alcohol/drug abuse. A woman affected by GBV may be permanently scarred and feel robbed of her innocence, purity, and virtue. The roots of GBV can be linked to countries established on patriarchy, and the self-interested misinterpretations of religious texts—the Bible and Quran—by men often keep women in a subservient position to them. Misinterpretations of religious texts may be caused by illiteracy, self-interest, or deceitfulness. Men can combine fallacies and scripture to assert that women are in a secondary position to them in the eyes of their deity. This keeps men in an allegedly divinely sanctioned superior position and gives them authority to enforce that position and power through violence. The factors behind GBV are composed of beliefs in family honor and a woman's sexual purity, ideologies of male superiority, men witnessing family violence when growing up, marital dissatisfaction, and difficulties in communication between partners. Through surveys conducted by research teams at Turkish universities, it is seen that “women with higher incomes than their spouses are at double the risk of beating, and two of every three women who contribute more to the household income than their spouses are subjected to physical violence” (Altınay, *Violence Against Women in Turkey—A Nationwide Survey 2009*, 65). This demonstrates that men who have lost their power to provide economically may resort to violence to reassert their masculine authority.

Gender based violence is most prevalent in the Eastern Mediterranean Region, South East Asia, and Africa. Attached is a figure with data provided by the World Health Organization in 2010 with the percentages of intimate partner violence by region.



Gender-Based Violence in Turkey

“Turkey is one of the world’s worst countries to be a woman. Between 2002 and 2009 the murder rate of women skyrocketed by 1,400 percent” (Johnson, Al Jazeera English 2014). Turkey is located in the Eastern Mediterranean Region, which is an area with one of the highest percentages of GBV: “One out of every three women in Turkey has been physically abused by her husband and this figure has not changed over the past decade” (Altinay, Violence Against Women in Turkey—A Nationwide Survey 2009, 2). Yet, Turkey has established a number of laws and reforms that harshly punish violence against women by any means. Three main problems still remain: lack of enforcement, lack of government support for gender equality and women’s rights, and pervasive patriarchal customs and attitudes that suborn GBV and women’s oppression. For example, the Law

for the Protection of Family was passed in 2012, which allows a judge to press a restraining order on an abusive spouse to protect the abused spouse and children, but as shown in the Violence Against Women in Turkey—A Nationwide Survey, 42.8 percent of Turkish women have never even heard of the Family Protection Law. And for those who are aware of the law, only 7.5 percent of those surveyed knew someone who has benefited from the law. The laws to help maintain women's autonomy may be presented perfectly on paper, but are rarely executed with effort. The issue of domestic violence is not treated with urgency, which causes women to remain silent about their encounters with violence. Most women do not report their incidents either out of shame and embarrassment or out of hope that their experience was a one-time incident.

The idea of violence is implanted in many Turkish men's heads as a solution to conflicts they may face with a female. As presented in the 2013 *Hurriyet Daily News* survey, "34 percent of Turkish men think violence against women is necessary, while 28 percent say violence can be used to discipline women" (Khazan 2015, 2). Even though Turkey is a secular country, the traditional religious lifestyle is still predominant. As a 40-year-old Turkish man stated, "We grew up in Eastern Anatolia hearing that when a woman makes a mistake, she is always wrong. That it's even okay to kill her... and some people in our community still believe this" (Khazan 2015, 2). In addition to the "general public" conforming to customary conduct, government officials also express male-dominant power. The current President of Turkey, Recep Tayyip Erdogan, stated—in a speech at a women's rights conference in November of 2014—that "he did not believe men and women are equal and that feminists don't accept the concept of motherhood...[He] believes in equal opportunities, [but] men and women

are different because it goes against the laws of nature” (O’Grady, Foreign Policy 2014). The list of Turkish officials who have made asinine statements goes on with Prime Minister Arinc banning women from laughing in public because it is not chaste; and saying that “a rape victim must have wanted it,” and “it is perfectly fine if parents give permission to their 15-year-old daughter to marry a 45-year-old-man” (Sassounian, The Huffington Post 2014). Ruling party parliamentarian Ahyan Sefer Ustun also declared, “A rapist is more innocent than a rape victim who chooses to have an abortion”(Johnson, Los Angeles Times 2015). Some other stories of GBV that reflect pure horror include a guilty man who was found innocent of stabbing his wife because he claimed that she was wearing provocative leggings; a father kidnapping his daughter so he can sell her off to another relative through marriage; and a girl who was buried alive as a form of ‘honor’ punishment by her relatives because she was talking to boys. It is also a sad truth that the cases of GBV and domestic violence that make it into court—one out of 30 women interviewed by Ozlem Albayrak and Alp Ardic—are often decided against the female. The unfair treatment can be seen when a judge ignored a woman’s complaints and gave her advice to go home and kiss her husband to solve the issue. If women are not taken seriously in legal situations, they are less likely to feel comfortable and confident in speaking up. Even if a male is charged for his crime of violence, he can get reductions from his sentence if he ‘claims’ that he was ‘provoked’ by a woman’s clothing or behaviors. Proof is rarely required from a male who claims that he was ‘provoked’ into violent actions, yet it is an acceptable statement in court. The prejudicial treatment women receive in court causes them to feel stigmatized because of their situation. In reality, a

woman is not responsible for violence committed against her and should never feel humiliated to speak up and seek help.

The Idea of Honor

Honor killings are one of the most common GBV cases in Turkey. The following story is of a woman who lost her life in an honor murder. Her account is unfortunate, but her case has marked a turning point for other women and the authorities in Turkey. Ceyan Sosyal was a nineteen-year old who eloped with her husband. They ended up having a baby girl together, but her husband abandoned her. In order to care for her daughter, she returned back home, “only to be killed by her brother while the rest of the family waited in the living room. Her father walked into the room several times to check why the murder was taking so long”(Shafak, The Guardian 2011). This case has marked a milestone in how Turkey authorities deal with events of GBV and honor killings. Instead of arresting the one who conducted the murder, as the procedure had always been, all of Sosyal’s family members were taken into custody. This decision can protect more women from honor murders because families will be less willing to conduct killings if they would have to sacrifice themselves to jail time. Typically, families organize the honor murder by using the youngest male in the family as a scapegoat (Shafak, The Guardian 2011). With the Sosyal verdict, there is a higher chance that families will think twice about the situation and avoid perpetrating an honor homicide.

Ozgecan Aslan Incident: A Revolution in anti-GBV and Turkish Social Media Activism

Social media activism for GBV has been an increasingly popular

method for combating all forms of violence towards women. A spark in the movement was ignited after the murder of Ozgecan Aslan, a 20-year-old college student. Aslan was taking the bus home and the bus driver attempted to rape her. As she fought back, the bus driver callously attacked her and dumped her body into a forest. After Aslan's body was found, Twitter erupted with the hashtags #OzgecanAslan and #sendenlat—you must also tell. Using the hashtag, #sendenlat, women in Turkey shared their personal stories and experiences with GBV via Twitter. Tweets ranged from their personal experiences with violence and harassment to the sharing of their personal tactics for self-defense and self-protection. #Sendenlat included over 440,000 tweets and became a trending topic worldwide in a matter of a few days. Another hashtag, #ozcecanaslan, was even more popular, trending as the highest topic worldwide with more than three million tweets.

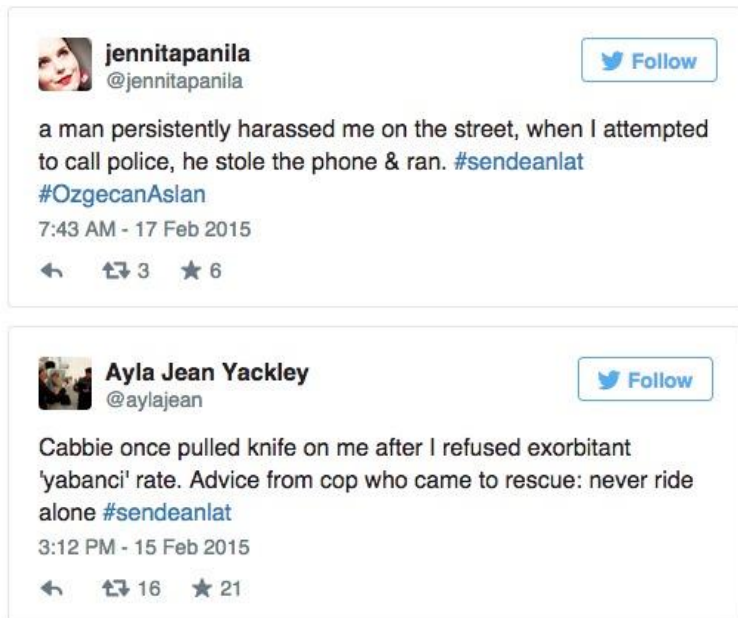


Figure 1- Tweets by female Twitter users who share personal stories of GBV and advice using the hashtags #sendenlat and #OzgecanAslan

Even groups of Turkish men joined the effort by wearing miniskirts to symbolize their rejection of conventional gender standards, protesting on the streets and online with the trending hashtag, [#ozgecanicinminietekgiy](#) (wear a miniskirt for Ozgecan). The street and online activism—especially with male participants—regarding Aslan’s case are significant because it shows that a woman wearing a miniskirt should not be used as an excuse for provoking acts of GBV like rape.



Figure 2—Men post pictures of them wearing miniskirts on Twitter to raise awareness for GBV as well as to show their opposition to traditional gender norms



Figure 3 -Twitter users who participated in physical protests also posted pictures online to broadcast their activism to a bigger audience

The Ozgecan Aslan incident has kindled a revolution in the perceptions of GBV in Turkey because usually both Turkish society and government try to find excuses to avoid confronting cases of rape and killings, but there was no alibi in Aslan's case. Aslan was completely innocent; she was merely a bus passenger who was on route to return home. This got many other women engaged in fighting against GBV because they realized it could have been any other women, including themselves, on the bus that day. In order to create pressure on the government and encourage them to increase the attention brought to GBV, activists and supporters used sites like Twitter and Facebook to post about protests and rallies in honor of Ozgecan Aslan. Once the news spread, protests were held in major cities such as Istanbul and Ankara with thousands of women and men marching in the streets while shouting slogans to remember Aslan.

The use of hashtags and its immediate ability to spread awareness of GBV and anti-GBV activism has proved successful in changing the

perceptions and perspectives of many Turkish citizens and officials, including President Erdogan himself, who once depreciated women's image through a public speech "declaring [women] were not equal to men," and said "the guilty deserved the most severe punishment" (Williams 2015). He continued to say in a televised speech "violence against women is an open wound on our society... I hope the awareness that has been raised with the death of Ozgecan will be the beginning of a new era" (Williams 2015). Joining President Erdogan, Turkey's Prime Minister, Ahmet Davutoglu added "the trauma we've experienced is so massive that no penalty is severe enough to match such a crime"... [he] expressed hope the crime would mark a turning point for Turkey" (Williams 2015). These statements from government officials serve as a positive example and stimulus for Turkish people to come together and combat violence against women.

Social Media Effects on GBV in Turkey

Social media has provided a forum for activists and supporters to share their opinions and beliefs, to organize and mobilize, and to educate those women who are unaware of what their rights should be. Women activists in Turkey share that social media "has had a massive impact, and if it wasn't for social media we wouldn't have the right information on anything. It's been our saviour" (O'Mahony, Telegraph 2013). Knowledge is power, and women who are educated about their rights will have a higher tendency to help stop GBV. A survey conducted twenty years ago with "a representative sample of 572 women in Istanbul, nearly half of the women responded "Yes" to the question, "Do you think there are situations in which a woman deserves to be beaten by her husband?" (Altinay, Violence Against Women in Turkey—A Nationwide Survey 2009, 34). However, women's

views on the legitimacy of domestic violence have radically changed over the course of twenty years. Two decades later, “an overwhelming majority of women agree that ‘there is never a valid justification for beating’” (Altınay, *Violence Against Women in Turkey—A Nationwide Survey 2009*, 36). Since the mass of women in Turkey believe GBV is unacceptable under any circumstance, there will be more women standing up for their rights, and thus, contributing to the declaration from activists towards government officials that women need to receive equal treatment and that laws protecting them should be fully implemented.

Before social media, women who experienced domestic violence or sexual violence only had a few online chat rooms in which they could share what they had to undergo. Those who were brave enough and attempted to talk to someone about it were simply turned away by the police and usually further punished by their male counterparts for seeking help. In particular, a woman who could no longer stand the violence said:

“I was battered so much that I felt like a lamb agonizing in the jaws of a wolf. I was ashamed to tell neighbors or relatives. So, I would go to the police station to tell them of my plight. The police would do nothing. I went to the police station four or five times, and each time I got a more violent beating at home for going there.” (Hurtas, *Al-Monitor 2015*)

In the past few years, women have utilized the web and united through social media to share stories of GBV incidents through blogs and videos as well as posting powerful visuals that illustrate the detrimental impacts of GBV on women.



Figure 4—An example of a GBV graphic posted on Tumblr

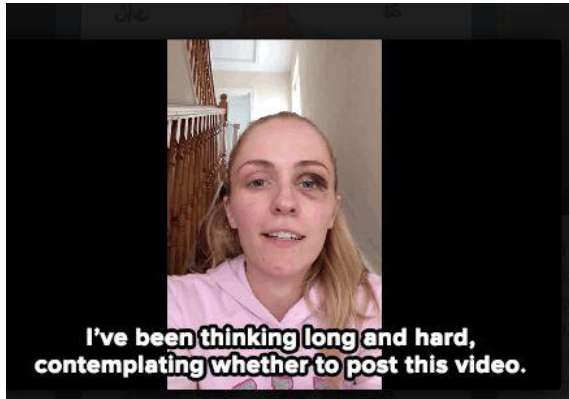


Figure 5— A woman sharing her story with GBV through a video posted on Tumblr

There are many Tumblr blogs—withal Twitter accounts and Facebook pages—dedicated as activism accounts of Turkish women and violence. Below is an example of a Tumblr blog dedicated to women in Turkey who have been abused and/or experienced violence. The blog is set up so survivors can submit their personal stories and have them posted without context. This way the GBV survivor's identity remains undisclosed, and the numbers of audience members who will view the survivors' stories are broadened, thus, reaching out to more people who might not have been aware of how harsh and serious GBV is in Turkey. This is similar to the Civil Rights Movement in America and the impact of the widely broadcast Selma March. Almost everyone knew about segregation at the time and that it was a morally unjust practice. However, many of those who were not near it did not know the full story and held a naïve perspective until they saw the violence on television.

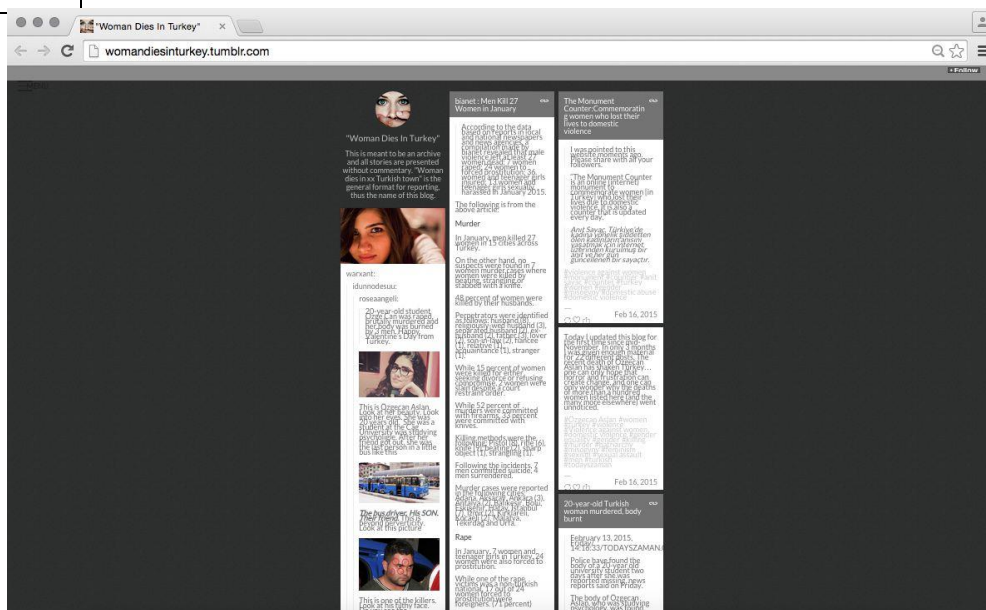


Figure 6- A Tumblr blog, "Woman Dies in Turkey" is an archive of stories from women who have experienced GBV

Social media has also served to reveal the corruption behind ruling-party figures in Turkey by broadcasting the brutalities of GBV to other countries. The accessibility of social media sites worldwide can make the Turkish government fearful and ashamed of what goes on in their country, which then causes them to enforce more laws and regulations in hopes of preventing and alleviating the conditions and cases of GBV. Social media users promoting women's empowerment play a pivotal role in contributing to the pressure that the Turkish government faces regarding GBV and protecting the country's women.

Strategies on Social Media

Since social media has the power to rapidly convey information and can provide on-the-spot timeliness and reporting, the authorities such as the

police can also quickly navigate where protests are happening. Some strategies activists have developed include using private group messaging, private Facebook groups, and blogging on Tumblr. A social media blogging site, Tumblr offers the ability for individuals to post anonymously online. A survey found that “one out of every three women who are subjected to violence do not share their experience with anyone else but instead have to deal and struggle with domestic violence on their own” (Altınay, *Violence Against Women in Turkey—A Nationwide Survey 2009*, 64). A third of the women surveyed never spoken of their violence because they were either fearful of their abusers, worried they would be further punished if they spoke up, wanted to protect themselves or their husbands, were ashamed to talk about it, or thought they would not be supported or taken seriously. Sites that provide an anonymous option to post have allowed women to share their stories with zero apprehension that they will be further punished or killed.

Another useful tactic has been hashtags on Twitter. Hashtags make it easier for users on social networking sites to find messages, statuses, or tweets with a specific theme or content, such as the #sendenlat and #OzgecanAslan hashtags involved with the Aslan case. Hashtags on Twitter also expedite the speed content can travel on the web worldwide. For example, when a Twitter hashtag is tweeted a certain number of times, it will become a trending topic and be seen on every Twitter user’s homepage. Thus, more people are exposed to a certain activism topic, which increases the awareness towards an issue requiring activism, and/or expands the audience and participants for a situation such as politicking against GBV.

Twitter has also been used to reach out to more audiences and gain supporters for online petitions that call for women’s rights, especially

regarding violence. Sharing the online petition links on social media increases the number of signers within Turkey and globally. This way the process of completing the required number of signatures for a petition is accelerated. Especially after Aslan's case, online petitions and the call for support have been amplified, such as "an online petition for a bill designed to curb violence against women managed to attract over a million signatures"(Sanderson, DW 2015). A statistic of over one million signatures symbolizes the growing amount of backers of anti-GBV and guarding Turkish women from violence.

The screenshot shows the Care2 Petitions website interface. At the top, there is a navigation bar with the Care2 logo, 'PETITIONS', and links for 'START A PETITION' and 'BROWSE'. A user is logged in, indicated by 'login / join'. The main content area features a petition titled "Demand Protection and Equal Rights for Turkish Women!". The petition is by author A. Ward and targets Turkish President Recep Tayyip Erdoğan. It has 81,390 signatures, with a progress bar showing this is 81.39% of the 100,000 target. A "Sign Now" button is prominent. To the right of the petition details is a sign-up form with fields for First Name, Last Name, Email, United States (dropdown), Street Address, City, State (dropdown), and Zip. There is also a checkbox for "Share with my Facebook friends" and a "Sign Now" button. A privacy notice at the bottom of the form states: "Your personal information remains private. By signing, you accept Care2's terms of service." The footer contains the Care2 logo, navigation links, and a "site feedback" link.

Figure 7-An example of an online petition that can be signed by people worldwide and shared on social networks

Progress Thus Far

Social media postings have pressured the Turkish government and society to pay more attention to helping women and their experiences with violence. Progress can be seen in the development of many women's

organizations comprised of activists who embark on the quest for equal rights between the sexes. For instance, Women for Women's Human Rights-New Ways is an independent women's non-governmental organization (NGO) that aims to promote women's human rights as well as equality and non-discrimination in Turkey and on the international level. Likewise, there are ventures from HASNA, a non-profit organization, such as Project SHINE (Self Help through New Exchanges) to empower and encourage NGOs set up to aid Turkey's women. What is unique about Project Shine is that it includes both female activists and their male allies. This is extremely important because "if [society] want[s] to build peace between the genders, it cannot only engage a single gender in the dialogue!" (Saad, HasNa). What is even more interesting is that both of these organizations—as well as many others—have their own social media such as Twitter and Facebook where they constantly post tweets and status updates to spread their message and accomplish the mission to keep every woman in Turkey protected. As a companion to women's organizations that are present to provide aid—based on the 2014 report from the Bureau of Democracy—there are also 90 women's shelters in Turkey. Simultaneously, there is a hotline to provide aid and advice; the 183 hotline is set up to help women who are suffering from violence by giving them an opportunity to call for legal advice and counseling from experts. The hotline has been a hit; it "received more than 100,000 calls in Turkey last year (2014)" (Domestic Violence, DailySabah 2015). Another influence that has provided a positive future for GBV situations is the combined involvement of media and Turkish celebrities and public figures with their support to raise awareness on violence against women and to help prevent it in the future. In April of 2015, well-known and respected Turkish actresses Türkan Şoray, Zerrin Tekindor and Serra Yılmaz

performed a show titled “Wounded to Death.” The show summarized stories of abused women to draw attention to the prevalence of violence against women in Turkey. These celebrities utilized their talents to expose a topic that is stigmatized for stage performance, and even casual conversation with peers. Their pursuits to shine a light on the darkness and evil behind GBV in Turkey has helped educate more of the population in regard to what goes on behind closed doors. Their performance shows how, as a community, they can come together to bring attention to GBV. There are also several platforms—a network of women across Turkey who are fighting for change—in existence. One of them is the Stop Women Murders Platform. They make documentary films such as *Dying to Divorce* to attract more supporters to help the women in Turkey fight for justice. They include English subtitles on their films, so more people around the world can be knowledgeable about what is happening in regard to GBV in Turkey.

Cell phone social media apps have made it even easier to access help for domestic violence. A revolution in solutions to helping GBV sufferers is the Turkish Easy Rescue app, which allows women to seek help by connecting them to emergency services and government hotlines without notifying their attacker that they have called for help. The app's function is subtle. All a woman has to do to activate the app is shake her phone. The app has proven to be a success; so far “at least 100,000 women have used Easy Rescue and more than 250,000 have downloaded it”(Jackson, Fortune 2015). However, in order to spread the news about the app without informing men about it, strategies had to be applied. The app company used marketing tricks such as promoting the app in the midst of a makeup tutorial video or other women’s interest shows. Many steps were taken into consideration for the presentation of the app. Campaign designers “made sure to place the

messages about the app far enough into the video that men were unlikely to overhear or stumble across them” (Jackson, Fortune 2015). Women activists and their male allies have come up with many approaches to successfully overcome any barriers they might encounter when promoting activism and empowerment on the media. As seen, advancements to help GBV victims and survivors are developing and stabilizing, but there are many more opportunities to participate in expediting the efforts to end GBV in Turkey and worldwide.

Help Stop GBV

A prominent platform—We Will Stop Femicide Platform (The Platform)—has been set up in Turkey and has received much attention after Aslan’s case; it is created to help stop femicide. The Platform is used to set up meetings or mass rallies in front of court hearings and these meetings or rallies are set up using social media accounts. This creates a sense of unity among women to have their voices heard on some of the female murder cases and pressures the justice system into punishing the perpetrators. The Platform supports women in their quest for their rights as well as their quest for protection and healing from violence. Solutions that The Platform has proposed for stopping femicide include receiving full support from the President, Prime Minister, and other leaders of the political parties to condemn violence against women; strictly enforcing the Family Protection Law; and establishing a Ministry of Women.

Other procedures that must be performed to ensure the security of women’s rights and protection from violence are installing more female judges in court that interview or interrogate victims of GBV. This way, women will feel more comfortable when relaying their stories. Furthermore,

the police must be trained to convey a more caring and responsive manner and express urgent concern by taking actions to keep the violent men away from the victims. The same goes for prosecutors and judges who deal with women and GBV cases. This will ensure that women will feel safe and be protected against violent scenarios, and it is also taking another step towards ending GBV.

Any act of activism can help save and better many lives from GBV. No effort is too small. Join the movement and help those who are experiencing GBV and domestic violence by getting online and using social media to enforce positive change: retweeting and tweeting activism hashtags, sharing Facebook groups and posts, following and commenting on blogs to express encouragement. In addition, one can share news article and blog post links regarding GBV in Turkey directly to his/her Twitter, Facebook, or Tumblr. Most news article sites and online blogs will have the option on the page itself to post an article to a social media site or even send the article or post as an email. Further endeavors one can engage in include participating in protests, activities, meetings, and organizations; joining a branch of The Platform or starting a branch in a region not yet represented; and supporting The Platform, non-profit organizations, women's organizations, and NGO's with donations.

Final Thoughts

It is probable that each individual on this planet knows a woman who is facing or has survived domestic violence. Gender based violence is unacceptable and needs to be terminated worldwide. It is a growing problem in many areas of the world, especially Turkey. It is an enigma that the rates

of GBV in Turkey are high because it is one of the few countries with codified laws to protect women. The underlying issues behind the Turkey puzzle include lack of enforcement and the lack of genuine governmental support to help women. Fortunately, social media has played a positive and pivotal role in augmenting the power of activism and creating greater response towards the issue of GBV with online activism, protests, rallies, and petitions. The success of social media (Facebook posts and Twitter hashtags) can be seen in the Ozgecan Aslan case. Ozgecan Aslan suffered a dreadful fate, but her story has revolutionized Turkish mindsets on GBV and forever changed the way Turkey's governments view GBV situations. The social media movement that was kindled after Aslan's death is one of the most impressive online undertakings with worldwide trending hashtags on Twitter, and it even changed the President's expressed views on women's rights. Social media has served as a spark to light a bigger flame in addressing GBV. More people around the world have been exposed to information about GBV in Turkey due to Twitter hashtags, Facebook statuses and activism groups, and Tumblr blogs dedicated to sharing the stories of GBV victims and survivors. Due to global awareness of violence against women in Turkey, the Turkish government and society have been pressured to act on the situation by passing laws and approving women's organizations that protect GBV victims. Anti-GBV activists have gained the support of men and women worldwide, including celebrities and public figures, who are tweeting and posting about women's organizations; sharing GBV news articles, video, blogs, photo links; and donating to platforms set up to provide aid to women experiencing violence. No action is too small; each contribution is crucial. Get online and promote positive change to help stop gender-based violence in Turkey!

Bibliography

"Woman Dies In Turkey" "Woman Dies In Turkey" Accessed July 20, 2015.

<http://womandiesinturkey.tumblr.com/>.

Altinay. "Violence Against Women in Turkey—A Nationwide Survey."

Kadinayoneliksiddet. Accessed July 20, 2015.

http://www.kadinayoneliksiddet.org/KYS_ENG.pdf.

Asquith, Christina. "We Will Stop Women Murders." The Daily Beast.

Accessed June 14, 2015.

<http://www.thedailybeast.com/articles/2015/03/02/we-will-stop-women-murders.html>.

Asquith, Christina. "Turkish Men Get Away With Murder." The New York

Times. February 24, 2015. Accessed June 14, 2015.

http://www.nytimes.com/2015/02/24/opinion/ozgecan-aslan-and-violence-against-women-in-turkey.html?_r=0.

Ay, Goncagul. Email interview by author.

Bar'el. "Lifting the Veil on Violence against Women in Turkey." Haaretz.

Accessed July 20, 2015. <http://www.haaretz.com/misc/iphone-article/.premium-1.655944>.

Bilefsky, Dan. "How to Avoid Honor Killing in Turkey? Honor Suicide."

The New York Times. July 15, 2006. Accessed June 14, 2015.

http://www.nytimes.com/2006/07/16/world/europe/16turkey.html?pagewanted=all&_r=0

Burton, Elise K. Email interview by author.

Davidson. "Rape and Murder of Young Woman Sparks Mass Twitter Protest in Turkey." *The Guardian*. Accessed July 20, 2015.

<http://www.theguardian.com/world/2015/feb/17/turkish-woman-ozgecan-aslans-sparks-anti-violence-campaign-sendeanlat>.

"Declaration on the Elimination of Violence against Women." UN News Center. Accessed July 20, 2015.

<http://www.un.org/documents/ga/res/48/a48r104.htm>.

"Domestic Violence in Turkey at 40% Says New Gov't Report | Clarion Project." *ClarionProject.org*. Accessed July 20, 2015.

<http://www.clarionproject.org/news/domestic-violence-turkey-40-says-new-gov-report>.

"Domestic Violence Hotline Planned for Turks Abroad." *DailySabah*. Accessed July 20, 2015.

<http://www.dailysabah.com/nation/2015/05/01/domestic-violence-hotline-planned-for-turks-abroad>.

"Domestic Violence in Turkey Reaches Boiling Point | Europe | DW.COM | 11.07.2015." *DW.COM*. Accessed July 20, 2015.

<http://www.dw.com/en/domestic-violence-in-turkey-reaches-boiling-point/a-18576377>.

Enloe, Cynthia. Email interview by author.

"FOR ENGLISH." *FOR ENGLISH*. Accessed July 20, 2015.

<http://www.kadincinayetlerinidurduracagiz.net/for-english>

Dying to Divorce. Turkey 2015. Film. This short film follows the work of Aysen, an advocate for 'Stop Women Murders Platform'; a network of women across Turkey who are fighting for change.

Hurtas. "No Justice for Turkish Women Who Kill in Self-defense - Al-Monitor: The Pulse of the Middle East." Al-Monitor. February 25, 2015. Accessed July 20, 2015. <http://www.al-monitor.com/pulse/originals/2015/02/turkey-violence-against-women-husbands-speak-out.html>.

Jackson. "How Women in Turkey Are Using an App to Fight Domestic Violence." Fortune. July 14, 2015. Accessed July 20, 2015. <http://fortune.com/2015/07/14/domestic-violence-app-turkey/>.

Johnson. "Turkey: The Life of a Battered Woman." Al Jazeera English. Accessed July 20, 2015. <http://www.aljazeera.com/indepth/features/2014/02/turkey-life-battered-woman-2014257594926477.html>.

Johnson. "Brutal Slaying Leads to Protests over Violence against Women in Turkey." Los Angeles Times. Accessed July 20, 2015. <http://www.latimes.com/world/europe/la-fg-slaying-of-turkish-woman-leads-to-street-protests-over-womens-rights-20150214-story.html>.

Jones, Sophia. "Despite Legal Gains, Turkish Women Are Still Dying Over The Right To Divorce." The Huffington Post. Accessed June 14, 2015. http://www.huffingtonpost.com/2014/11/19/turkish-women-divorce_n_6133470.html.

Karatas. "Turkey's Leading Women Hit the Stage, Draw Attention to Violence against Women." *Today's Zaman*. Accessed July 20, 2015. http://www.todayszaman.com/anasayfa_turkeys-leading-women-hit-the-stage-draw-attention-to-violence-against-women_378784.html.

Khazan, Olga. "In Turkey, Not Even Posters of Women Are Safe From Violence." *The Atlantic*. March 25, 2015. Accessed July 20, 2015. <http://www.theatlantic.com/international/archive/2015/03/domestic-violence-against-women-video-turkey/388544/>.

Korkman, Zeynep K. Email Interview by author.

Merry, Sally Engle. *Human Rights and Gender Violence: Translating International Law into Local Justice*. Chicago: University of Chicago Press, 2006.

"MOR MUHABBETLER." *Erkekler Dijital Dünyada Kadına Şiddet Uyguluyor*. Accessed June 14, 2015. <http://www.morcati.org.tr/tr/anasayfa/312-erkekler-dijital-dunyada-kadina-siddet-uyguluyor>.

O'Grady, Siobhan. "Erdogan Tells Feminist Summit That Women Aren't Equal to Men." *Foreign Policy*. Accessed July 20, 2015. <http://foreignpolicy.com/2014/11/24/erdogan-tells-feminist-summit-that-women-arent-equal-to-men/>

O'Mahony. "Turkey Protests: How Activists Stay One Step Ahead with Social Media." *The Telegraph*. Accessed July 20, 2015. <http://www.telegraph.co.uk/technology/internet-security/10098353/Turkey-protests-how-activists-stay-one-step-ahead-with-social-media.html>.

"Petition: Demand Protection and Equal Rights for Turkish Women."

Petition: Demand Protection and Equal Rights for Turkish Women!

Accessed July 20, 2015.

<http://www.thepetitionsite.com/219/399/461/demand-protection-and-equal-rights-for-turkish-women/>.

"Platforms | Women for Women's Human Rights." Women for Womens Human Rights. Accessed July 20, 2015.

<http://www.wwhr.org/platforms/>.

Saad. "Responding to the Ozgecan Aslan Tragedy: HasNa's Story on the Fight against Gender-based Violence." HasNa. February 23, 2015. Accessed July 20, 2015. <http://hasna.org/responding-to-ozgecanaslan-tragedy-hasnas-story-on-the-fight-against-gbv-sendeanelat/>.

Sabloff, Nicholas. "Turkish Girl Buried Alive For Talking To Boys." The Huffington Post. Accessed July 20, 2015.

http://www.huffingtonpost.com/2010/02/05/turkish-girl-buried-alive_n_450019.html.

Sarkar. "Men Wear Miniskirts to Support Women's Rights - CNN.com." CNN. Accessed July 20, 2015.

<http://www.cnn.com/2015/02/22/europe/turkey-men-miniskirts/>.

Sassounian, Harut. "Turkish Deputy Prime Minister Tells Women Not to Laugh in Public." The Huffington Post. Accessed July 20, 2015.

http://www.huffingtonpost.com/harut-sassounian/turkish-deputy-prime-mini_b_5656807.html.

Scaturro, Michael. "Talk Therapy for Wife Beaters in Germany's Turkish Community." *The Atlantic*. September 2, 2013. Accessed July 20, 2015. <http://www.theatlantic.com/international/archive/2013/09/talk-therapy-for-wife-beaters-in-germanys-turkish-community/279006/>.

"Selected Links on Gender Based Violence." HHRI. Accessed July 21, 2015. http://www.hhri.org/thematic/gender_based_violence.html.

Shafak, Elif. "Turkey Opens Its Eyes to Domestic Violence." *The Guardian*. Accessed July 20, 2015. <http://www.theguardian.com/commentisfree/2011/aug/23/turkey-domestic-violence-honour-killings>.

"Social Media Helps Minorities in Turkey Connect to Their Culture, Survey Finds." *Today's Zaman*. Accessed June 14, 2015. http://www.todayszaman.com/national_social-media-helps-minorities-in-turkey-connect-to-their-culture-survey-finds_384706.html.

"Turkey." *Stop Vaw*. Accessed July 20, 2015. http://www.stopvaw.org/uploads/ending_dv_in_southeast_europe_and_turkey.pdf.

"Turkey Social Media Ban Raises Censorship Fears - BBC News." *BBC News*. Accessed July 20, 2015. <http://www.bbc.com/news/world-europe-32204177>.

Vlachova, Marie. "Women in an Insecure World: Violence Against Women Facts, Figures, Analysis." UNICEF. Accessed July 20, 2015. http://www.unicef.org/emerg/files/women_insecure_world.pdf.

"Why I Believe the Mistreatment of Women Is the Number One Human Rights Abuse ." Jimmy Carter:. Accessed July 20, 2015.
https://www.ted.com/talks/jimmy_carter_why_i_believe_the_mistreatment_of_women_is_the_number_one_human_rights_abuse.

Williams, Stuart. "Turkey Vows Action after Outcry over Female Student's Murder." Yahoo! News. Accessed July 20, 2015.
<http://news.yahoo.com/turkey-arrests-3-over-students-murder-outcry-grows-085623804.html>.

"Özgecan Aslan: Men Kill Women – Research Turkey." Research Turkey. Accessed July 20, 2015. <http://researchturkey.org/ozgecan-aslan-men-kill-women/>.

