A Philosophical Defence of Thought Experiments in Political Philosophy

by Patrick Johnson Mendie BA (Hons), MA

Submitted for the Degree of Doctor in Philosophy

June 2022

Department of Politics, Philosophy and Religion

University of Lancaster

Abstract

***A Philosophical Defence of Thought Experiments in Political Philosophy,* submitted by Patrick Johnson Mendie for the degree of Doctor of Philosophy, June 2022.**

This thesis provides a defence of thought experiments in political philosophy, also known as '*political thought experiments'* (PTEs). Different problems cast doubt on whether PTEs can be considered as being trustworthy. Critics think that the use of unrealistic hypotheticals in PTEs is problematic, especially when scenarios are completely detached from 'real' and 'recurring' situations as they occur in the actual world. I resolve this problem using the reflective equilibrium method, demonstrating how we can establish the logical equivalence between the unrealistic scenarios and the real-world counterparts with which they can be compared to understand their real-world implications. I also argue that since PTEs are arguments, the use of unrealistic premises does not affect the validity of their arguments. Critics also think that PTEs yield varying responses from different readers, meaning their intuitions are neither stable nor objective. I respond to this concern using John Norton's ' *argument view'* which proposes the idea that scientific thought experiments are disguised arguments, consisting of tacit premises and conclusions. I extend this view to PTEs, arguing that some PTEs contain valid arguments ruled by a system of logical inference. I examine some examples of PTEs by reconstructing them as valid arguments, arguing that PTEs with valid arguments are capable of evoking stable and objective intuitions. Finally, I discuss the concern of whether PTEs can motivate readers, in particular, how PTEs can be used in responding to real-world dilemmas. I establish the relationship between PTEs and narratives arguing that PTEs with 'narrative transport' would reasonably motivate readers. Using the problem of corruption in Nigeria as a case study, I show how PTEs can provide action guidance in responding to the issues of corruption in Nigeria. In conclusion, the original contribution of this research lies in the defence of PTEs, and the claim that PTEs can be useful in solving real-world dilemmas.

Acknowledgements

Throughout the writing of this thesis, I have received enormous support from many people who greatly helped me in completing the thesis. First of all, I would like to thank my supervisors Professor Rachel Cooper and Dr Nick Unwin. I am also grateful to my examiners Professor Simon Beck (External examiner) and Dr Patrick Bishop (Internal examiner) for their comments during the viva voce. I appreciate those criticisms and comments in the build-up of this thesis. Their expertise was invaluable in motivating me in formulating both the research topic and methodological approach and the final submission of this thesis.

I greatly benefited from different stimulating discussions with many colleagues at the University of Calabar and the University of Uyo, Nigeria, to whom I am very grateful. These are Prof. Dorothy Oluwagbemi-Jacob, Prof. Kyrian Ojong, Dr Chris O. Akpan, Dr Darty Emmanuel, and Dr Emmanuel Iniobong. I would also like to thank my colleagues at Lancaster University, Dr Victor Chimela, Dr Itoro Udoh and Dr Okezie Ikonne for encouraging me during the writing of this thesis.

I am grateful for the financial support from Tertiary Education Trust Fund (TETFund) Nigeria, for providing part of the sponsorship that enabled the completion of this thesis. I am also grateful to Lancaster University's Department of Politics, Philosophy and Religion for providing a supportive environment throughout my studies.

In addition, this thesis would not be possible without the support and motivation of my mother, Mrs Comfort Mendie and my siblings Blessing Mendie, Imaobong Essien, Rose Uwoh, Isidore Mendie, Mr Emmanuel Essien, Mrs Etido Bassey, and Patience Mendie. These people have supported me in various ways throughout this long, difficult, and rigorous process.

Finally, I would like to thank my dear wife Ifure Patrick Johnson and my daughter Janelle Patrick Johnson for their unparallel love, support, patience, and motivation that made the completion of this thesis possible. Those I have forgotten to mention here are not less important. Indeed, they are deeply appreciated for all their contributions toward the success of this work.

Declaration

I hereby declare that this thesis is my work and has not been submitted for any other degree of qualification. The work in this thesis has been produced by me, except where due acknowledgement is made in the text. I also confirm that this thesis does not exceed the prescribed limit of 80,000 words, including the main text but excluding the bibliography.

Patrick Johnson Mendie

CONTENTS

CHAPTER ONE

1. POLITICAL THOUGHT EXPERIMENTS: AN INSIGHT INTO THE POWER OF IMAGINARY CASES IN POLITICS

1.1 Introduction

This thesis examines the use of imaginary/hypothetical situations in moral and political philosophy also known as *'political thought experiments'* (PTEs). It explains how thought experiments can be used legitimately in responding to political or moral issues. Thought experiments have recorded huge successes across disciplines, especially in science, and there has been quite a lot of philosophical work defending their usage in science (see Norton, 1996, 2000, 2004, 2018; Brown 1999, 2008, 2010, 2012; Authur, 2018; Peacock, 2018; Corcilius, 2018 and a lot more). But, while their successes in other disciplines have continued to be acknowledged, particularly in physics, their counterparts in philosophy, especially moral and political philosophy have remained controversial. To date, thought experiments in political philosophy have continued to receive serious criticisms and are sometimes rejected mostly by empirically minded scholars. My contribution in this thesis is to defend the use of thought experiments in political philosophy, in particular, the legitimacy of intuitions that are elicited via political thought experiments. The thesis also offers responses to some of the key questions/concerns surrounding their usefulness. Furthermore, the thesis investigates how political thought experiments could provide moral and political guidance to their audience in combating real-world issues. This will involve discussions on whether PTEs can reasonably motivate their readers and also if they can be institutionally possible, that is if PTEs can be used to make changes in our political institutions starting from our present political situations. Findings on the motivational possibility are used to provide insight into how thought experiments in political philosophy could be redesigned to better provide political motivations, in particular, in shaping the moral and political behaviour of its audience. My claims on the institutional possibility in political thought experiments will be used in this thesis to look at how PTEs might help in a real-world setting, taking the issue of political corruption in Nigeria as a case study. Below, I will elucidate what 'thought experiments' are, including their essential features. Subsequently, I will also engage in the subject 'political thought experiments' and their uses from where the main problems of this thesis are discussed.

(a) Thought experiments and their uses

Briefly, before I proceed to the main issues of the thesis, it is important to explain some of the essential ingredients and general use of thought experiments. The idea is to demonstrate what sort of hypothetical or imaginary scenarios could count as a thought experiment. This will further explain what thought experiments are, how they work, the nature of their intuitions, and finally why we need them in normative politics. Herein, I consider thought experiments in this thesis to mean: those processes that involve imaginary reasoning (sometimes presented as stories) carried out by anyone to explain the nature of things. They are often performed within the context of well-articulated hypothetical situations used mostly by philosophers and scientists to answer complex scientific puzzles and to respond to philosophical issues concerning nature and our general way of life. Thought experiments according to Brendel (2018) "aim to improve or undermine existing theories by revealing contradictions, paradoxical implications, intuitive or unintended results of theories and also pave way for better ones" (Brendel, 2018, 281). During thought experiments, imaginary arguments are often employed basically to evoke intuitions used in conducting a mental assessment of how we ought to live, and what actions we ought to take given a particular imaginary situation. Intuition in this sense is considered to be the knowledge we get during the thought process via thought experiments. Intuitions come to the mind of its readers as mental images or as ideas in the reader's brain and do not need to involve new data from experience. Sometimes, intuitions come as a spontaneous judgement to the reader during imaginary exercise, without the need for conscious reasoning. For instance, the moral/political judgement we get during any imaginary exercise is a result of the elicited intuitions. In this thesis, I will argue that the intuitions elicited when employing political thought experiments could influence our real-world actions, capable of making people (audience) come to certain intuitive conclusions or behave in a way they usually do not. In addition, these intuitions can influence how we think and respond to complex issues, in particular, how we respond to normative claims. To date, the power of intuitions via thought experimentation has continued to play a significant role in how we see the entire world, political debates/theories have been shaped by it, and have continued to influence our moral and political way of life.

The question is, do all forms of imaginary exercise, stories or hypotheticals in the mind of the reader count as thought experiments? Certainly, across disciplines not all forms of hypotheticals do count as one, perhaps those to be considered as thought experiments must be

very persuasive to evoke the needed intuitions, and must meet some fundamental criteria. I, therefore, endorse Brun's (2018) analysis of thought experiments that explains the three fundamental steps to look out for when justifying what sort of imaginary exercise counts as a thought experiment across disciplines. According to Brun (2018), the three key steps to look out for during thought experimentation include:

i. A scenario and a question are introduced.
ii. The experimenter goes through (by imagination or thinking about it, etc.) the scenario and arrives at some result.
iii. A conclusion is drawn with respect to some targets. For example, an ethically relevant claim or distinction etc. (see Brun, 2018, 196).

From Brun's illustration, (i) and (ii), are considered the core features of a good thought experiment because these steps contain the imagined premises that give rise to the conclusion (iii) in a thought experiment. Perhaps, intuition represents the result, and it is what justifies what sort of conclusion the thought experimenter wants to evoke in the mind of its readers. Generally, most successful thought experiments provide self-evident truth, in this thesis, I will argue that intuitions via PTEs could move the audience without the need for any physical experiments. In other words, through intuitions, successful thought experiments contain some persuasive force that is capable of influencing the moral and political choices of their readers without reference to real-world action. The claims about the nature of intuitions in PTEs will be developed further as we proceed in the thesis. However, apart from the intuitive components of a thought experiment, most thought experiments are significant because of their special ingredients that involve argumentation, abstraction, narratives, and the use of unrealistic hypotheticals. These shall all be explained as we proceed in the thesis.

(b) Thought Experiments in Moral and Political Philosophy

Thought experiments have long been accepted in different fields, such as science. In physics scholars like Galileo, Newton, Einstein, Schrodinger, and many others have in numerous ways employed the use of hypotheticals in making significant contributions that ushered in great changes in science. For instance, there are famous scientific thought experiments that have impacted the way we think about the universe which includes: *Galileo's thought experiment on falling bodies*-this thought experiment was successfully used by Galileo (1971) to explain the fact that all bodies no matter their sizes fall at the same time. This idea

also displaced the Aristotelian idea that a heavy body falls faster than a light body (see Galileo, 1971, 66f). Another successful scientific thought experiment was *Newton's bucket and absolute space* thought experiment. Before this time, Leibniz had claimed that space only exists in relation to bodies, and this idea questioned the existence of absolute space. Using this thought experiment, Newton was able to debunk Leibniz's claims and thereafter, developed new assumptions that justified the existence of absolute space (See Principia, 7). Einstein's famous thought experiments on *riding the elevator in freefall* and *chasing the light beam scenarios* are also of great significance in science. These two thought experiments helped to shape the way scientists/ modern physicists think about velocity and the nature of light and were quite useful in Einstein's later theory of relativity in modern physics (See Einstein, 1938; 1949). *Schrodinger's cat experiment* is another famous scientific thought experiment that is remarkable in the development of quantum physics. It was used by Erwin Schrodinger to demonstrate how critics were misinterpreting quantum theory, especially the superposition claim (see Schrodinger, 1935).

Similarly, thought experiments have also been widely used in areas such as epistemology, ethics and philosophy of mind. For instance, in ethics, Peter Singer uses *the pond case and the drowning child* thought experiment to show that there is no moral difference between a child drowning in front of you and one starving in some far-off land. He concluded that since we can save a drowning child with little or no cost then we also have the responsibility of donating to a charity organisation far away which might cost us very little (Singer, 2009). The *Chinese Room* by John Searle is also famous in the discipline. This thought experiment concludes that computers do not have minds, they are strictly limited to the exchange of symbolic strings and are not capable of understanding language (Searle, 1999). Equally, *Mary, the colourblind neuropsychologist* is also famous and was used by Frank Jackson to justify his claims that kick against the view that the universe is made up of only physical entities, hence his idea that non-physical properties do exist (see Jackson, 1982). Thought experiments have also been used in political philosophy; they are commonly expressed using hypothetical scenarios in form of events and stories concerning an imagined political arrangement. Some famous ones are the *Original Position* by Rawls; *the Republic* by Plato; and the *Social Contract* by Rousseau are good examples of PTEs that are used to explain different notions of justice and political arrangements. I will explain in detail some of these thought experiments as we proceed.

However, the use of thought experiments in political philosophy (political thought experiments) has been more controversial and subjected to serious scrutiny. Critics are worried about the role of intuitions that are often elicited via political thought experiments (PTEs), in particular, they are also refusing to acknowledge the roles of thought experiments in normative theories, even though these thought experiments were instrumental in different political debates and political arrangements far back in the ancient Greek era. Miščević (2018) is one of the notable proponents of the use of thought experiments in the field, and the phrase *'political thought experiment'* was first coined by him (see Miščević, 2018, 154). He uses this term to engage readers on the different philosophical issues about the legitimacy of imaginary cases in political philosophy. Miščević's conversation and claims seek to explain how intuitions from the discipline were legitimately used by philosophers to serve a variety of functions. He consistently thinks that the intuitions from political thought experiments can be used "to establish a *prima facie* source of evidence for normative truth"; and that "they can serve as heuristic and testing device for political theories" (see Miščević, 2018, 167). Furthermore, Miščević also claims that the use of thought experiments in political philosophy can be considered as "one source of political theorizing and can be used to examine just arrangements and other just related issues in politics" (see Miščević, 2018, 159). Miščević further endorsed different works of political thinkers and their ideas as political thought experiments. For instance, he thinks that *the Republic* by Plato, Jean Jacque Rousseau's *Social Contract,* and *the Original Position* by John Rawls are typical examples of political thought experiments of inestimable value that offer "possible arrangements to intuitions and judgements of justice"; and that "they aim at offering the reader an understanding of proposed normative social ideals" (see Miščević, 2013, 510).

1.2 Key Issues and Challenges

Despite Miščević's account, the use of imaginary scenarios in political philosophy has continued to face serious criticisms and rejection. Most empirically-minded scholars believe that intuitions that are elicited via PTEs are often problematic (see Thaler, 2017; Klampfer, 2017; Tobia, 2018). These critics think that intuitions generally are not trustworthy because thought experiments in political philosophy have suffered from a whole lot of issues like the use of unrealistic hypotheticals, the use of abstraction in PTEs, the unstable nature of their intuitions, the question of bias and prejudices, and finally the issue of generalization and extrapolation in PTEs. I will explicate some of the key issues below.

Briefly, Thaler (2017) thinks that the use of unrealistic hypotheticals in PTEs is problematic, especially those that are physically implausible, whose scenarios are completely detached from 'real' and 'recurring' situations as they occur in the actual world. Thaler presents his argument in two ways. Firstly, he argues against the moral permissibility of abstraction in political thought experiments. He contends that abstracting away real-life details during thought experiments in political philosophy can be morally dubious. Primarily, Thaler believes that it could be difficult if not impossible to establish an empirical connection of PTEs with the real world of politics if abstraction becomes permissible. The second way Thaler presents his argument is by assessing how thought experiments are supposed to provide us with action-guiding moral norms. Thaler shows how abstraction from some thought experiments could possibly take us away from important moral details. Using a few illustrations, Thaler shows how thought experiments are supposed to direct our moral and political actions in the real world. He argues that in the current literature on political philosophy, there are no thought experiments that can fit this demand, which implies that unrealistic thought experiments are generally not useful in political philosophy.

Most empirically minded scholars also think that intuitions via thought experiments often go wrong because they are unstable, and prone to be affected by our cultural values, language, age, discipline and the religious beliefs of the reader (see Klampfer, 2017; Tobia, 2018; Roojen, 2014; Harman, 1973; Machery, 2011). For instance, Klampfer (2017) has questioned the philosophical usage of intuitions, in particular, the ones that are evoked by means of thought experimentation in moral and political philosophy. He believes that any experiments, whether conducted in the laboratory or reader's mind, are supposed to yield intuitive results that are stable and should not be doubted. His concern is that intuitions generated through moral and political thought experiments produce indeterminate results, whose conclusions are often controversial and divisive; by different people. For a similar reason, Dave (2001) avers that "it would be a huge mistake to conclude anything from philosophical intuitions" (see Dave, 2001). These critics think that their usage for political theorizing, or as a *prima facie* source of evidence for normative truth or political arrangement is mistaken. From the above exposition, critics have therefore concluded that intuitive outcomes from the use of thought experiments in moral and political philosophy are inadequate, and that thought experiments in political philosophy must be rejected.

There are also problems relating to the motivational and institutional possibilities of PTEs. Proponents of political thought experiments have suggested that for imagined scenarios to be successful in political philosophy, they must be motivationally and institutionally possible (see Miščević, 2018). Meaning the individual and political institutions are supposed to be motivated by such projects, and that the proposed political arrangements via thought experiments should also be useful for institutional changes/development in the real world. This requirement also calls for some specific questions concerning the general efficacy of political thought experiments. For instance, can the *Original Position* thought experiment by John Rawls be motivationally and institutionally possible? How do we interrogate the motivational power of political thought experiments in responding to political issues in the real world as we know it? Similarly, how can moral and political thought experiments shape the feelings and political behaviour of their audience? How can we redesign PTEs to suit these purposes? However, critics are continually worried about whether the intuitions from unrealistic hypotheticals in PTEs could motivate its audience and whether such motivations could also lead to action guidance in the real world. For instance, how could political thought experiments motivate individuals and political institutions in combating the problem of political corruption and its negative impacts in the world, in particular, in a multicultural setting such as Nigeria? And what sort of institutional possibility via thought experiments could help eradicate corruption in Nigeria? With these questions in mind and the doubt occasioned by the use of unrealistic hypotheticals in the field, the motivational/institutional possibilities of political thought experiments remain unclear and have become another controversial aspect of the use of thought experiments in political philosophy that is yet to be settled. This thesis shall engage in different findings/conversations based on these questions to prove the general efficacy/defence of some thought experiments in political philosophy.

To date, there is no response or possible explanation in defence of the present use of thought experiments in political philosophy vis-a-vis the above concerns. Therefore, the crux of this thesis is drawn to fill these gaps. Firstly, to explain why we do not need to reject the use of unrealistic hypotheticals in PTEs. Secondly, to explain how to evoke reliable intuitions from PTEs through the study of the existing relationship between thought experiments and argument. Also, I will establish whether PTEs can be hoped to reasonably motivate their audience by means of their motivational power. I will also use Rawls's thought experiment on the *Original Position* to argue for the importance of factors such as collective action, rule of law, freedom of speech and democracy. These factors will be used in discussing the institutional possibility

of PTEs. To achieve these aims, the thesis believes that political theorists have in many ways used thought experiments more efficiently for political reasoning and that if we study them properly using the right approach, we could understand how they could be assessed to explain current political realities. The thesis endorses Miščević's account of hypotheticals to establish a framework for understanding, assessing, defending, and providing a clear description of what sort of imaginary situations should count as a reliable way of thought experimenting, and what kind of intuitions can be considered adequate in political philosophy. I further contend that the claims against the reliability of intuitions in PTEs are mistaken because there is a potent interconnection between political thought experiments and arguments that have not been acknowledged by these critics. This is to say that 'argument' is at the very heart of most successful thought experiments in political philosophy. I argue that, if one understands these connections, then one can understand the trustworthiness of intuitions in political thought experiments. In presenting my views, I endorse '*the argument view'* by John Norton to examine the conditions under which thought experiments can validly yield logical/instructive results. The *'argument view'* is used in this thesis to respond to the questions relating to the unstable nature of intuitions via PTEs and to defend why current unrealistic political thought experiments cannot be rejected. This study shows that the use of unrealistic hypotheticals (as premises) does not in any way affect the intuitive outcome of thought experiments in political philosophy. This is because arguments do not need to have realistic premises or conclusions for them to be valid.

Furthermore, the thesis will be looking at corruption in Nigeria as a case study to explore how PTEs might be useful in bringing about actual political change. As such, the Nigerian people to date are practically denied justice and the benefit of their common good by the few political elites, in what I also denote as political corruption. With political corruption in Nigeria, I will expose how leaders in the country have used their various position of power for personal benefits, an action that is affecting individuals and various institutions in the country. The analysis of corruption in Nigeria is used in this thesis as a case study to examine whether PTEs can be used in responding to real-world issues. Meanwhile, findings from the issues of political corruption in Nigeria have given rise to the following fundamental questions: how do we eradicate political corruption in Nigeria? How can political corruption be eradicated especially when it has become a moral norm in modern political settings? This thesis suggests that we can achieve this through behavioural changes by means of political thought experiments. Using McVey's (2020) findings on *narratives and philosophical arguments*, the

thesis believes that there is something so special about thought experiments that can be hoped to motivate readers, such that it can be used to shape the feelings and political behaviour of its audience given any complex moral or political situation. The motivational roles of political thought experiments (PTEs) are employed in this thesis to establish how individuals and institutions via PTEs can bring both moral and political changes in responding to the behavioural lapses of a people in a heterogeneous society as is the case in Nigeria.

Meanwhile, while I believe that political thought experiments are arguments, the thesis also claims that some thought experiments in political philosophy can be considered as narrative, through which moral and political thought experiments can be hoped to motivate their readers during thought experimentation. I endorse the relationship between narratives and political thought experiments to explain the motivational possibility of PTEs. I argue that most thought experiments in political philosophy are presented as narratives as well as arguments. As narratives, readers are prone to getting more psychologically involved by the thought experimental claims, such that it is capable of facilitating *narrative transport*. Narrative transport is the key mechanism that shapes our beliefs which might result in changing behaviour during narrative exercise (see McVey, 2020). I argue that by using the narrative approach, political thought experiments can hope to motivate readers, because of their narrative and argumentative powers. Through this medium, it suggests that PTEs with the right features can facilitate narrative transport, meaning we can as well gain real-world political changes with the aid of imaginary situations via thought experimentation in political philosophy. Meanwhile, I interrogate these claims with Rawls's thought experiment on *the Original Position*, which offers us a legitimate means to discuss how some PTEs can facilitate narrative transport and how to redesign future PTEs to meet the required ingredients to facilitate narrative transport in order to aid motivations.

The thesis also thinks that some important factors that make up the conclusion of most successful PTEs can as well be extrapolated to solving real-world issues, especially, in providing policy guidance. For instance, in Rawl's thought experiment, factors such as collective action, freedom of speech, rule of law, and democracy can all be intuited for such purposes. I argue that extrapolating these factors and using them as policy guidance during thought experiments, can make us live a better life. This strategy has been adopted in the thesis to offer a model explanation of the institutional possibility of thought experiments and is used to explain how PTEs can be useful in solving complex institutional/political problems in the real world as we know it. To justify this claim, I used these factors to interrogate the

institutional possibility of Rawls's thought experiment with the problem of corruption. Using Nigeria as a case study, I also defend the institutional possibility of PTEs and argue that *the Original position* can produce useful factors that can be deduced in understanding how political corruption can be eradicated in Nigeria. The idea is that the *Original position* thought experiment can provide important factors such as collective action, rule of law, and freedom of speech to the mind of its audience for the betterment of our political life and in responding to our political challenges. I argue that these factors could provide a good avenue for moral education, they can also serve as a guide for the implementation of the collective action framework which is required to establish political reforms in the fight against corruption in Africa and Nigeria in particular. In addition, the information or conclusions from PTEs can as well be extrapolated and used by lawmakers as heuristics for law-making purposes and policy guidance to address the real issues of political corruption for a long time.

1.3 Positioning of the Thesis

This thesis is a multidisciplinary project covering different fields (such as politics, logic, epistemology, and behavioural studies) with the aim of defending the use of thought experiments in political philosophy. Primarily, as a work in political philosophy, the thesis examines how the device of imagined situations can be used more effectively in different political arrangements. Thought experiments play an essential role in political philosophy, and this can be seen in the works of Plato, John Rawls, Jean-Jacques Rousseau, and a lot more. Miščević (2018) has claimed that the concept of intuitions is central to the effective use of thought experiments in political philosophy, but his claims have raised serious gaps in the discipline as demonstrated in section 1.2. The thesis attempts to fill these gaps by engaging the loopholes and providing philosophical explanations on the role of thought experiments in political philosophy using the different tools of logic, behavioural studies, epistemology, and politics.

As claimed above, this thesis involves the use of logical tools in defending PTEs. The idea is that there is a close relationship between political thought experiments and logical arguments. The argument of this thesis overlaps with some claims that have been discussed by some authors, but the position of this thesis differs significantly from these authors in many ways. For instance, I employ *the 'argument view'* by John Norton as a foundation for some of my claims. Norton also thinks that scientific thought experiments are nothing more than "dressed-up arguments" (Norton, 1991, 1996, 2002, 2004). Although primarily Norton's work

focuses on scientific thought experiments, the claims of this thesis focus on moral and political thought experiments. Many critics of this view think that thought experiments are not arguments because imagined scenarios are not like premises as is the case in standard logic. My contribution to knowledge is in part to defend *the argument view* and to reconstruct some moral and political thought experiments as arguments following Norton's *reconstruction thesis*. Findings from this view are used to develop a framework for evaluating the use of hypotheticals in political philosophy and at the same time filling the existing gaps in the use of thought experiments in political philosophy.

This thesis also explores some key concepts in behavioural studies as a tool for explaining the internal mental processes and states of the thought experimenter during thought experiments. Part of my claim is that there are reasons to hope that political thought experiments could motivate their audience, and perhaps, based on what goes on in the head of every reader while reading thought experiments, PTEs could as well shape their moral and political behaviour. Thus, to determine whether a political thought experiment is capable of motivating or shaping the moral and political behaviour of readers, the thesis also evaluates the internal mental states of readers of hypotheticals; to unravel how they would react during thought experimentation to justify their behavioural outcomes. To achieve this aim, the thesis explores the tools of the narrative approach that explains the existing relationship between thought experiments and narratives. The relationship between thought experiments, narratives, and their assumptions offers a plausible way of considering thought experiments as stories. As stories, they can facilitate narrative transport. The scope of this finding is based on McVey's (2020) studies on the relationship/impacts of narratives and philosophical arguments in behavioural changes. The original contribution of this thesis is that political thought experiments can be studied as narratives and it focuses on the claim that via the narrative approach, PTEs with the right narrative features could facilitate narrative transport, capable of shaping the moral and political behaviours of the audience. In this thesis, I present the view that Rawls's thought experiment can be examined as a narrative. Based on this finding, the thesis believes that Rawls thought experiment is promising because its argument is valid, and it has some of the required ingredients to facilitate narrative transport. Meaning the thought experiment can be hoped to change the behaviour of the audience. In addition, some factors from Rawls's thought experiment are deduced and considered useful tools in discussing the institutional possibility of PTEs. This discussion will help to explain how PTEs can be useful

in providing real-world strategies in responding to political problems especially to address the issue of political corruption in Nigeria.

1.4 Research Objectives

This thesis offers a detailed philosophical defence of the use of thought experiments in political philosophy, in particular, how PTEs can be useful in responding to the corruption problems in Nigeria. The principal aim of this thesis is to address the challenges described in Section 1.2. I am of the view that thought experiments can be legitimately used in political philosophy and that they can motivate political change. My objectives are as follows:

- To argue that both realistic and unrealistic PTEs can be legitimately employed in political philosophy.
- To defend the use of 'reflective equilibrium' as a method that can be employed to connect moral intuitions derived from unrealistic thought experiments with real-life cases. I will also use this method to show how conclusions from good moral and political thought experiments can be relevant to practical matters.
- To defend John Norton's *argument view* and argue that good moral and political thought experiments can be reconstructed as valid arguments.
- To reconstruct some moral and political thought experiments as arguments to justify their validity. This shall be used to explain how intuitions that are elicited via PTEs could be considered objective and stable. Similarly, it shall also be used to offer a response to the various criticisms of intuition in political thought experiments.
- To examine the connection between narratives and political thought experiments drawing from Mcvey's (2020) studies on *'the moral importance of narrative: a philosophical analysis of narrative transport'.* Findings from this study will be used to unravel whether political thought experiments can reasonably be hoped to motivate readers to change their actions. It shall also examine fundamental reasons why PTEs could be redesigned to facilitate narrative transport in order to shape the political behaviour of the audience in the real world.
- To assess the problems of political corruption in Nigeria. Findings from this study shall be used as a framework for further investigation of how PTEs can be useful in solving such problems in the real world as we know it.
- To assess how *the original position* thought experiment could be used to suggest how justice might be brought about in the real world. I use this thought experiment to argue

for the importance of factors such as collective action, rule of law, freedom of speech and democracy. Findings from this assessment will be used to suggest novel ways of establishing anti-corruption reforms in responding to the problems of political corruption in the real world using Nigeria as a case study.

1.5 Synopsis

The thesis comprises ten substantive chapters. While the current section (chapter one) introduces us to the general discussion of the entire research, including the problems, positioning (scope), and objectives of the research, the rest of the chapters are hereby summarized as follows:

Chapter 2 considers an analysis of the nature of political thought experiments that demonstrates what conditions should imaginary scenarios have to count as thought experiments in political philosophy. The aim is to participate in conceptualizing the method of thought experiments and its implications in political philosophy. Focusing on Nenad Miščević's (2018, 2013a, 2013b, 2013c, 2012) claims the chapter analyses the nature of this project to explain the scope, structure, and definition of thought experiments. The overall aim of this chapter is to examine Miščević's notion of hypotheticals, in particular, his desiderata of political thought experiments. I endorse his proposals to offer a general explanation and purposes of thought experiments in political philosophy. I then explore how the arguments highlighted in Miščević's account could help explain the following key questions in the field: (i) What kind of hypothetical or counterfactual scenarios should we count as political thought experiments? (ii) What are the purposes/functions and the epistemic significance of thought experiments in political philosophy? However, because Miščević's account provides justifiable support for the use of thought experiments in political philosophy, I also explore some of the arguments made by critics of his notion. I then suggest how these concerns may affect our study of thought experiments in political philosophy. Yet, some of the key issues and challenges of this thesis emanate from the lapses created by Nenad Miščević's account of political thought experiments.

Chapter 3 examines the various debates and criticisms in moral and political thought experiments. With Miščević's account of thought experiments in the fore, much remains unclear about the use of thought experiments in political philosophy. The nature of their intuitions and their usefulness has been seriously questioned, in particular, their usage for normative evidence has remained a subject of scrutiny by most empirically minded scholars. I present Thaler's (2016) account of hypotheticals in the field in detail focusing on the step-by-

step arguments for why Thaler thinks thought experiments in political theory are untrustworthy. Similarly, in the same chapter, I discuss the concerns of Klampfer's (2017) views and outline his claims about why intuitions from moral and political thought experiments are unstable and should be considered problematic.

In chapter 4, I endorse *the argument view* by John Norton (1991, 1996, 2004) to provide a theoretical explanation of the existing relationship between thought experiments and arguments. I use this view in the entire chapter to argue that some political thought experiments are ruled by systems of logical inference that contain premises and conclusions and are capable of providing valid arguments. The idea is that we can evoke stable intuitions from PTEs if we understand the potent interconnection between imaginary situations and arguments that have not been acknowledged by critics of thought experiments in political philosophy. Perhaps, if one understands these connections, then one can better understand the intuitive stability of thought experiments in political philosophy. Using data from Rawls's *Theory of Justice*, I demonstrate how John Rawls used the *Original Position* thought experiment to argue for the notion of justice. Perhaps, *Original Position* acts as an operator that provides premises for a valid argument on the notion of justice. To defend the reliability/stability of intuitions via moral and political thought experiments, I adopt the reconstruction strategy of John Norton's *argument view* by reconstructing *the Original Position* as a valid argument.

In chapter 5, I respond to Thaler's criticisms of the use of unrealistic hypotheticals in political thought experiments. Before tackling the issues raised by Thaler, I provide a detailed explanation of the various roles of unrealistic hypotheticals in political development. Drawing from the *argument view*, I argue that the use of unrealistic hypotheticals as premises during political thought experiments does not affect the validity of the moral and political claims that are embedded in its conclusion. I also defend the use of abstraction, which enables the reader to abstract away unrealistic situations to gain the right intuitions. I argue further that because thought experiments are arguments, the use of abstraction via thought experimentation has a legitimate role to play in the development of political theory regardless of the bizarre nature of some of their scenarios, as such, political philosophers can legitimately use unrealistic hypotheticals to direct us on how to live a better life. I defend the relationship between unrealistic thought experiments with the real-world, by endorsing the *reflective equilibrium* method, which connects moral intuitions with real-life cases using the logic of comparison of their intuitive equivalence. I use this method to argue that unrealistic cases can be connected with their real-life equivalence to aid action guidance in politics. The chapter also addresses

the question: What would successful or failed unrealistic thought experiments in political theory look like? I suggest that readers of unrealistic thought experiments must look out for their intuitions to establish a reflective equilibrium between the imagined scenario of hypotheticals and the real world for which they can both be compared. I further contend that, as arguments, the use of unrealistic hypotheticals does not in any way change the epistemic value of thought experiments in political philosophy. Because the premises of arguments do not need to be realistic to provide a valid argument, as such, moral and political thought experiments are not controversial as critics might claim.

In chapter 6, I offer direct responses to the various criticisms surrounding the reliability of intuitions in PTEs. I start by responding to unreliability claims by Klampher and other criticisms from notable empirical-minded scholars. For instance, Klampfer (2017) has argued that political thought experiments evoke intuitions that yield varying responses from different readers, meaning that the use of thought experiments in political philosophy cannot be justified because of the unstable nature of their intuitive claims. His concern is that unstable intuitions cannot be used as *prima facie* evidence for normative judgments. In responding to this claim, this chapter adopts *the argument view* of PTEs to offer a possible solution to the debate. The idea is that if the argument in a PTEs happens to be valid, therefore, thought experiments in political philosophy are capable of eliciting a stable intuition. I argue further that *the Original Position, the Violinist Case and the Pond and Envelope Case* are examples of such thought experiments that are capable of evoking objective intuitions in their conclusions. The idea is that some moral and political thought experiments can yield a logical conclusion which means they can also evoke legitimate intuitions that can be considered as being stable, objective, and reliable, and would as well be used in responding to complex moral/political issues in society. I also use these claims to respond to the issues of generalization and extrapolation, and the issue of bias and prejudice that were raised by Klampfer (2017) against the use of thought experiments in political philosophy.

In the subsequent part of this thesis, I develop this argument further. As such, Chapters 7 and 8 provide insight into how political thought experiments can be useful in the real world, in particular the issues of political corruption. In chapter 7, I examine the motivational possibility of political thought experiments to address the impact of PTEs on the reader's behaviour, in particular, whether PTEs can reasonably motivate human behaviour. First, I examine some traditional philosophical arguments to illustrate the point that we often intellectually accept valid and sound arguments, and yet we do not act as they suggest we

should. The question is: is there something so special about political thought experiments that can be hoped to motivate readers to act or change their actual behaviour during thought experimentation more than other types of traditional philosophical arguments? With this concern, I argue that there is something special about thought experiments that can trigger motivation especially, those PTEs that can be expressed as narratives and contain the right features that facilitate narrative transport. This chapter argues as follows, that most political thought experiments are both arguments and at the same time narratives. As narratives, readers are prone to getting more psychologically/emotionally involved in the thought experimental claims, such that it is capable of facilitating narrative transport. Narrative transport is considered in this thesis as the key mechanism that shapes our beliefs which might result in changing behaviour during thought experimentation. To explain this claim, the chapter evaluates Mcvey's empirical findings on the moral importance of narrative and philosophical arguments. Mcvey's (2020) work shows that readers are prone to changing their behaviour when exposed to narratives that facilitate narrative transport. This study also makes us believe that stories with higher narrative transport triggers changes in behaviour, meaning political thought experiments with such distinctive features can be hoped to motivate action. In conclusion, the chapter demonstrates the importance of redesigning future thought experiments in political philosophy to meet the various elements of narrative transport in order to understand the possibilities of getting motivated by means of PTEs.

In chapter 8, I expose the nature of corruption- to show how political leaders engage in different levels of corrupt practices that are affecting our society. I begin this chapter by providing a clear definition of corruption, especially in governance, I use this view to investigate how corruption is being carried out by political leaders globally. Presenting Nigeria as a case study, I focus on the issue of political corruption in the country. Findings from this chapter show how political corruption affects Nigerian society negatively, especially on issues relating to (i) institutional and (ii) individual behaviours. Both effects have led to an increase in mortality rate, economic instability, and poor infrastructural development in Nigeria, which has become a source of great concern. I suggest in this chapter that for us to eradicate corruption, the individual members of the society, first of all, must adopt good moral and political behaviour, while politicians also need to make good policies to address the institutional engagement on corruption. I finally argue that it is also pertinent to direct the study of hypotheticals in political theory to complex political issues especially the problem of political corruption in Nigeria. The various findings of this chapter are used respectively in

chapters eight and nine to examine (i) whether political thought experiments could hopefully be used in shaping corrupt moral and political behaviour. (ii) To explain the impacts of PTEs in responding to the corruption issues in Nigeria. I adopt the institutional possibility of PTEs to demonstrate how thought experiments can provide an alternative approach to the problem of political corruption in Nigeria. In the same chapter, I show how factors (such as collective action, freedom of speech, and rule of law) as deduced from Rawls's *original position* can translate to policy guidance using the collective action approach in responding to corruption issues in Nigeria. This chapter also draws ideas from the proposal of Persson et al (2010) concerning the problem of anticorruption reforms in Africa. Persson's proposal is used to provide empirical facts on why the corruption problem in the continent should be conceived as a collective action problem as is also the case in Nigeria. However, the present fight against corruption in Nigeria has failed due to the predominant principal-agent framework which has been used to date as a strategy in anticorruption policies. The reason for this failure is that given the principal-agent framework, "the agent betrays the principal's interest in the pursuit of his/her personal interest", this has affected the fight against corruption in Africa including Nigeria. Based on Persson's findings which highlight the problem of corruption in Africa as that of collective action, I have argued that Rawls's thought experiment provides a better framework for explaining how the collective action problem could be resolved. Using the collective action approach and demonstrating it within the logic of Rawls's thought experiment, the chapter shows how to resolve the issue of corruption in Nigeria as discussed in chapter 7. While Rawls's case provides the idea that it would be good if people cooperate in responding to corruption in Nigeria, the study of Hamada et al (2017) is also used to offer empirical insight providing reasons why the rule of law and freedom of speech as suggested by Rawls can be used to enforce cooperation from all individuals in a state.

In chapter 9, I conclude the findings of the thesis. I also summarize the main point providing answers to my main research questions. I reflect on other questions concerning the reliability of political thought experiments in our contemporary world that would be interesting to address in future research.

CHAPTER TWO

2. AN APPRAISAL OF NENAD MIŠČEVIĆ'S ACCOUNT OF POLITICAL THOUGHT EXPERIMENTS

2.1 Introduction

This chapter aims at discussing Nenad Miščević's account of political thought experiments (PTEs) in a bid to analyse the nature of his project, and also explain the scope, structure, and definition of thought experiments in political philosophy. I intend to use Miščević's account to demonstrate what conditions imaginary scenarios should meet to count as successful PTEs.

Briefly, Nenad Miščević (2018, 2013a, 2013b, 2013c, 2012) in a series of articles has extended his general account of thought experiments to the field of political philosophy. In these works, Miščević seeks to demonstrate the legitimacy of thought experiments in political philosophy and also outlines a few desiderata through which thought experiments could be considered successful in political philosophy. In explaining his proposed notion of *'political thought experiments'*, Miščević uses the works of Plato, John Rawls, and Jean Jacques Rousseau to demonstrate the impact of thought experiments in the field of political philosophy. Miščević considers two traditions, (i) the platonic and (ii) the social contract traditions as key to his explanation. Miščević's main contention is that the extension of thought experiments in the field of political philosophy should not be considered problematic because they evoke an understanding of how political affairs could be properly arranged.

To capture Miščević's claims in detail, the rest of the chapter will proceed as follows: 2.2 will examine Miščević's definition and scope of "political thought experiments". His definition of political thought experiments will enable us to discuss what kind of hypothetical scenario should be considered as PTEs. In section 2.3, I will explicate the two traditions employed by Miščević in discussing how the works of John Rawls and Plato fit into his account of PTEs. Miščević uses this view to argue that political theorists made use of counterfactual or hypothetical scenarios to develop their political theories which would be considered as an explicit example of PTEs. Section 2.4 will examine Miščević's desiderata for thought experiments. These desiderata are used by Miščević to present those conditions under which PTEs could be considered successful. Section 2.5 offers some responses to the various challenges of Miščević's desiderata of political thought experiments. And finally, in 2.6, I will

look at the general summary of Miščević's, claims and will thereafter examine his fundamental assumptions and presuppositions of PTEs.

2.2 Politics in the Armchair: An Overview of Nenad Miščević's Account of Political Thought Experiments (PTEs)

What are political thought experiments? Miščević believes that his idea of PTEs can be classified as a subspecies of ethical thought experiments that draw their normative values from both "moral and political principles" (Miščević, 2013a, 509). He thinks that PTEs "are meant to facilitate our ascent to normative (and/or conceptual) facts or truth, as well as the internal structure of normative realities" (see Miščević, 2004). Miščević conceives the idea of PTEs as having to do with those hypothetical scenarios in political theories that have to do with "justice and other justice-related qualities of political arrangement and the principles that govern them" (see Miščević, 2013b, 200). He outlines that PTEs "belong to political epistemology, the branch of social epistemology concerned with how people come to know (truth) about politics and political life, in particular, general truth" (see Miščević, 2013b, 191). The idea of truth from Miščević's claims focuses on moral truth, in particular, those that have to do with political principles that one may acquire through thought experiments (see Miščević, 2013a, 509).

Miščević thinks that PTEs can be characterised as follows:

i. "Thought-experimental reasoning involves reasoning about a particular set of social and political circumstances, which may be specified in more or fewer details.
ii. The thinker's mode of access to the scenario is via imagination rather than via observation.
iii. Contemplation of the scenario takes place with a specific purpose: coming to a judgment about some politically relevant theoretical proposal" (see Miščević, 2018, 155).

Miščević then considers how PTEs achieve their primary goals. Miščević believes in the role of intuition in political philosophy, and he thinks that intuitions are vital to PTEs. He claims that the purpose of thought experiments in political philosophy is to provide or to act as "a guide to intuitions" about a particular or general political arrangement; to serve as a means of "political theorising, and to improve our reasoning in matters of political theory" (see Miščević, 2018, 157). In this case, intuitions are used as prima facie sources of evidence for normative claims. In Miščević's view, for any hypothetical scenarios to count as PTEs, means

they must be able to provide intuitions that are vital for political arrangements or other normative judgments in politics (see Miščević, 2018, 154).

2.3 The Traditions of Thought Experiments in Miščević's Account of Political Thought Experiments (PTEs)

Using the works of Rousseau, Plato, and John Rawls, Miščević identifies two major traditions (The *Platonic* and *Social Contract*) to explain the efficacy of thought experiments in political philosophy. This section shall elucidate these traditions of PTEs and their epistemic significance in the study of thought experiments in political philosophy.

2.3.1 The Platonic Tradition

Briefly, Miščević situates the '*Platonic tradition'*, within the framework of Plato's *Republic*. He thinks that the entire *Republic* is a macro (political) thought experiment that contains "a series of smaller thought experiments, to be integrated into a large whole of the scenario of the ideal society" (Miščević, 2013b, 196). This means Plato's *Republic* (as a utopia) is in itself a thought experiment that consists of different micro-thought experiments that are important for political reasoning. He considers *the Republic* as "a blueprint for our world" that would help us to understand the nature of justice (see Miščević, 2012, 161). Miščević thinks that Plato uses different scenarios to develop an elaborate explanation of an ideal state and at the same time guides our understanding of how our society ought to be ethically arranged.

Miščević considers *the Ring of Gyges case* as a typical example of a political thought experiment within the Platonic tradition. The *Ring of Gyges case* is one of Plato's famous (micro) thought experiments, one that demonstrates the nature of morality in *the Republic*. In this case, Plato uses the characters of Glaucon and Socrates to examine the nature of justice and morality with the aid of an imaginary situation. In the dialogue between Socrates and Glaucon, Socrates claimed that justice is the highest good, and is intrinsically and extrinsically valuable. Glaucon disagrees with Socrates' claims by using a thought experiment called *the Ring of Gyges*. In this scenario, Glaucon claims that all persons are naturally egoistic and selfish. He asks us to imagine giving two rings that make the wearer invisible to both just and unjust individuals. Using the rings, the wearer could do things without notice or paying for their crime. Glaucon argues that they would act in the same way in protecting their interests. In this scenario, Gyges is seen as a person who has acquired one of the rings that makes him invisible, allowing him to commit crimes/injustice like rape, stealing and murder without being noticed. Glaucon claims that if the second ring is given to a just person, the same outcome will

be achieved, such that we could secretly conceive the person as being foolish if s/he does not act that way. Intuitively, he thinks that people will not resist the temptation of acting 'unjustly' if given such power. They act 'justly' because society has made some strict laws that force people to choose actions that are just. Glaucon uses this case to argue against Socrates that people cannot act justly willingly. They can act unjustly very easily if they have such powers (the power to commit injustice without being questioned), meaning absolute power corrupts us easily. Glaucon states that society values justice, not as 'a *good'* or what they like but because they are too weak to commit injustice with impunity. With these claims, Glaucon thinks that justice is not intrinsically and extrinsically valuable as Socrates had considered. Ethically, these claims suggest that *the Ring of Gyges* is vital in political philosophy as it poses important questions for political reflection. For example, 'can justice be valued in and itself?'; 'Is it better to act justly than to act unjustly?' These are some of the thought-provoking questions that Plato attempts to answer in *the Republic* about the nature of justice and morality in society.

It can be argued that *the Ring of Gyges case* creates an intuition on how human beings ought to behave in their different political settings. Miščević thinks that *the Ring of Gyges case* is not problematic; and further claims that such imaginative exercises are done in the armchair, without physical experimenting or observation, are aimed at discovering the truth about some given issues, and are central cases of thought experiments (Miščević, 2012, 156). Miščevic believes that *the Ring of Gyges* is aimed at discovering psychological facts to elicit moral intuition needed for social and political arrangement (Miščević, 2012, 157). However, Miščević also thinks that *the Republic* also contains other scenarios involving " the community of goods, of women and children, education system and the philosopher rule" (see Miščević, 2012, 153). Given this understanding, he firmly believes that *the Republic* is a macro (political) thought experiment, that contains some micro scenarios that provide us with the general intuition of a macro arrangement of an ideal state called the *Republic*.

2.3.2 The Social Contract Tradition.

Miščević uses the *Social Contract* tradition to demonstrate the prevalent use of hypothetical imaginations by different philosophers to explain how individuals in the state of nature take over their political rights through mutual agreement. What is a social contract? In a political sense, a "social contract is an agreement between the government and its citizens, defining the rights and duties of each" (see Encyclopaedia Britannica, 2019). Miščević believes that these hypothetical agreements of individuals involving "an idea of justice and right"

produced through the use of imagined scenarios are examples of political thought experiments (see Miščević, 2013a, 510). He thinks that the entire social contract theory of different political philosophers that at least contains hypothetical construction would also be described as a possible example of a political thought experiment.

Miščević extends the idea of 'social contract tradition' to the philosophical work of Jean-Jacques Rousseau. In his seminal paper titled, "*In search of the reason and the right-Rousseau's social contract as a thought experiment*" (2013a); Miščević, considers "the social contract as a piece of normative abduction"', expressing partly historical and imaginary reasoning (see Miščević, 2013a, 515). Miščević thinks that within the understanding of today's notion of thought experiments, "Rousseau is one of the first, if not the first thinker who seriously addressed the idea of the social contract as hypothetical" (Miščević, 2013a, 515). Miščević argues that Rousseau's social contract is one of the pillars of the contractarian tradition of political thought experiments which has "become central in modern political philosophy" (Miščević, 2013a, 510). In buttressing his point, Miščević believes that Rousseau's case; "fulfils the necessary requirements put by (hopefully) intuitively acceptable definition of a thought experiment and fits in the contractarian tradition that has been experimenting with hypothetical arrangements since its start" (Miščević, 2013a, 509). He claims that it is a thought experiment "primarily in the area of political norms", whose duty "offers us a criterion telling us under what condition an arrangement is legitimate" (Miščević, 2013a, 510). Miščević claims that this can be used to assess actual arrangements and would help in the development of modern political reasoning.

In Rousseau's social contract theory, we are asked to "imagine men in the state of nature reaching a point when the impediments that endangered their survival prevailed by their resistance over the force that each individual could use to survive that state. At this point, these primitive conditions could no longer subsist, and the human race would perish unless it changed its mode of existence" (cited in Miščević, 2013a, 516). Rousseau contends that it is a result of the situation when "men could not create any new forces, but only combine and control those that do exist that they have to form by aggregation a sum of forces which may prevail over the resistance, to put them in action by single motive power, and to make them work in concert" (cited in Miščević, 2013a, 516). Rousseau's hypothetical scenario explains how men were able to transcend from the state of nature to the attainment of their political rights, a method that Miščević generally considers as "a forward normative abduction" using hypothetical illustrations (see Miščević, 2013a, 516).

Furthermore, Miščević identifies Rawls's thought experiment (the *Original Position)* as a good example of a political thought experiment within this tradition (see Miščević, 2013b, 191). John Rawls uses some hypothetical scenarios to propose a social contract theory on how to attain a just society. Rawls asks us to imagine some group of men and women (in the original position) who had to decide on how to arrive at a fair distribution of goods. In the *Original Position,* Rawls claims that agreement on an egalitarian society could be reached by those negotiating from behind "*the veil of ignorance"* (see Rawls, 1971, 118). Behind the veil of ignorance, parties rationally choose the society they would prefer but they are completely ignorant of their personal characteristics and their conception of goods. Rawls argues that if those men and women are rational, they will employ a *maximin rule* that enables them to rationally choose a fair distribution of goods including what inequalities would be allowed. The maximin rule is a principle for making choices when parties at the original position are not sure of the outcome that will result from their choice. It states that in "choosing their arrangement parties are to evaluate each option in terms of the worst possible outcome that could result from choosing that option and to pick the option that offers the best worst outcome" (Rawls. 1971, 134-5). According to Rawls, employing maximin is the only way they can avoid risking losing out when they come out from the veil of ignorance. As such, they will choose to have political liberty that is compatible with the liberty of all, and will only accept those inequalities in power, wealth, income and resources that benefit the worst off. They will work to secure the best possible worst-off scenario when they finally step out of the veil. Miščević considers Rawls's hypothetical scenario to be a good example of a political thought experiment that can guide our thinking about justice.

Many critics of hypotheticals might hold that *the social contract* is more of a historical event, but Miščević considers it a thought experiment. He believes that the social contracts should be regarded as "hypothetical events" which means their construction is dominated by the use of imagined scenarios, one that cannot be considered as the actual history of political revolution (see Miščević, 2013a, 515). Furthermore, Miščević claims that his two traditions of political thought experiments have played vital roles in political philosophy. In buttressing his point, he notes that: "Plato's work is surely central for the ancient political thought, inspiring Stoic and neo-Platonic political treatises". Similarly, "the social contract tradition in its turn is of central importance for modernity" (Miščević, 2013b, 198). The idea is that these thought experiments provide intuitive insights into how political affairs could be properly arranged.

The following key assumptions can be deduced from the social contract tradition which is important for the study of political thought experiments: (i) the social contract encompasses "imagining in thought" (ii) it involves "a particular set of social and political circumstances to be arranged for" (iii) it contains intuitive details of what can be achieved, perhaps asking us whether we could sign for such a contract (as it is the case is Rousseau's social contract). As we proceed in this thesis, I will be arguing that the social contract is not what we can confirm by means of experimentation in the real world, but to be thought out imaginatively. I shall contend that such imaginative exercises are capable of providing the reader with an intuitive understanding of a social arrangement. Such intuitions guide our normative principles, particularly when we consider which political arrangements will become useful for our society.

In the section below, I will proceed to explain Miščević's desiderata of political thought experiments. I will use these desiderata to further demonstrate what should constitute a successful thought experiment in political philosophy.

2.4 Miščević's Desiderata of Political Thought Experiments

In several articles, Miščević has also proposed some considerations called "the desiderata of political thought experiments" and used them to explain his notion of political thought experiments. He believes that these desiderata outline the fundamental conditions that hypothetical scenarios must satisfy in order to be considered successful thought experiments in political philosophy (see Miščević, 2013a, 519).

2.4.1 The Desideratum of Relevance

According to this desideratum, Miščević thinks that the imagined situations used in the development of PTEs are relevant if they contain important information or factors to do with "political arrangement" (see Miščević. 2013a, 519). In discussing the basic requirement of this desideratum, Miščević considers the following questions: what kind of (empirical) information should a political thinker include in his or her proposed political arrangement during thought experimentation? Or is all considered information relevant to the thought experiments? In responding to these questions, Miščević contends that for thought experiments to be relevant in the field, political thinkers "have to consider the importance of matters" within the imagined scenarios such that they contain information that enables the actualization of specific political goals or arrangements in the real world (see Miščević, 2018, 161). He thinks that this desideratum remains a significant part of every political thought experiment such that, "it helps to figure out the importance of the actual thought experiment which is crucial in the moral-

political debate" (Miščević, 2013c, 51). In summary, given this desideratum, imagined situations need to capture useful information (factors) that offers an explanation of why a particular political thought experiment is proposed; notably, it must address matters to do with political arrangements for such thought experiments to be considered relevant in political philosophy.

2.4.2 The Desideratum of Realisticness

Miščević distinguished three broad categories through which this desideratum can be met (i) intrinsic realism, which deals with the internal coherence of the thought experiment world and (ii) extrinsic realism which explains the relationship of a thought experiment with the actual world. And finally the stability of arrangement. Based on the intrinsic requirement, Miščević makes the following claims about the nature of thought experiments in political philosophy: (a) "can thought experiments in political philosophy hold in its own possible, thought experimental world, given the human nature with its most general truth?" (see Miščević, 2013a, 519). (b) That the thought experiments must also have an internal coherence that is also possible and could be accessible in the real world. The idea is that this desideratum proposes that the imagined scenarios should have a useful coherence that enables thought experiments to combat real-world problems. Therefore, successful thought experiments in political theory must be realistic enough to support action guidance.

Furthermore, on being extrinsic, Miščević, claims that the proposed political arrangement in a thought experiment must also be realistic, such, that the conclusion of the thought experimental argument would be implemented in political matters in the real world. In buttressing this point, Miščević argues that the imagined scenarios through which political thought experiments are produced must hold possible such that they are successful "if the model advanced by the thought experiment can be implemented in the actual world starting from the present situation" (Miščević, 2013a, 519). Miščevic derived three other terms (motivational, institutional and accessibility) from the work of Cohen (2010) to distinguish three forms of possibilities: (i) '*motivational possibility'* and (ii) '*institutional possibility' and (iii)'accessibility possibility'*. He believes that through these possibilities the extrinsic requirements of PTEs would be legitimately fulfilled (see Miščević, 2013a, 519). Via **motivational possibility**, Miščevic thinks that political thought experiments are only useful if they are motivationally realistic meaning "people must be such that in principle they can be motivated for such a project as proposed by the imagined scenarios" (Miščević, 2013a, 519). Similarly, Miščević used the i**nstitutional possibility** to respond to concerns about whether

political thought experiments can be used to address issues relating to the political arrangement in our society. Miščević seeks to know whether an ongoing society would meet the condition described by the proposed arrangement. In buttressing this point, he asked the following question "Can we here in Europe, or in Asia or Africa, or somewhere on the rest of the continents, create a just society?" (Miščević, 2013a, 519). He thinks that "the proposed political arrangement has to be such that it can be implemented in the actual world starting from the present condition of our society" Meaning such arrangement must be at least in "principle implementable starting from our world as it is" (Miščević, 2013a, 520). Concerning the case of **accessibility possibility**, Miščević wants to know whether current institutions can be changed into imagined ones. In buttressing this, using Rousseau as a case study, Miščević asked the following question, is there any route leading from current circumstances to the society of the general will?" (Miščević, 2013a, 520). The last consideration of this desideratum is that Miščević wants political thought experiments to be stable (see Miščević, 2013a, 519). The claim here is that in successful PTEs if the hypothetical political institution is made actual then that institution should also be stable. The idea is that for thought experiments to be categorized as successful, the model supported by the imagined scenario should in practice be accessible, stable, and motivating to the audience. It must contain information leading to the present condition of society then such political thought experiments can be considered extrinsically or intrinsically realistic.

According to Miščević, there are worries about whether many political thought experiments would satisfy this desideratum; he thinks that some thought experiments in political philosophy do not meet the standard requirements for the desideratum of realisticity. In buttressing this point, Miščević raises the following question: to what extent can Rawls's social contract be considered as being realistic? However, he responds that some commentators on the *Social Contract* like Cohen (2010) claim that the social contract has "a genuine human possibility, that is compatible with human complexities, and with the demand of social cooperation" (Cited in Miščević, 2013a, 521).

2.4.3 The Desideratum of Detailed Coverage

In the "desideratum of detailed coverage", Miščević argues that for political thought experiments to be successful, all relevant information has to be taken into consideration. In buttressing this point, Miščević contends that "we should take into consideration all aspects of political arrangement; the width of coverage; including the richness of detail in the scenario

imagined" for such thought experiments to be considered successful (see Miščević, 2013a, 519).

Miščević thinks that some well-known political thought experiments fail to consider all relevant information and fail to meet this desideratum. For example, Miščevic believes that Rousseau's social contract does not fulfil this desideratum for "lack of detailed information concerning the contract itself" (Miščević, 2013a, 520-23). Similarly, T. Irvin (2008) also questions the deficiencies of Rousseau's social contract. For instance, Irvin raises the following questions*:* "why is self-protection the only aim that deserves to be considered? Why should we insist that each person remain free as he was before?" (Cited in Miščević, 2013a, 520). Both Miščević (2013a) and Irvin (2008) think that Rousseau's social contract is notorious for not providing enough information about the general terms of the contract itself. In buttressing his claims, Miščević raises the following question concerning the details that are lacking in Rousseau's social contract: "what matter is covered by the hypothetical contract; is it a just politics, or perhaps also the economy, and what about gender matters, what is the size of the community created by the contract and so on " (see Miščević, 2013a, 520). The general worry here is that the basic information of Rousseau's social contract lacks detailed coverage. His submission is that political thought experiments could be problematic if they opt for minimal coverage or lack of detailed information or in any way do not take into account all central factors as part of the thought experimental claims.

2.5 Examining the Challenges to Miščević's Desiderata of Political Thought Experiments

In this section, I intend to identify some of the challenges of Miščević's Desiderata of political thought experiments, the aim is to provide a philosophical insight into how these challenges affect the general idea of this thesis. This section should not be mistaken as a direct response to the issues raised as it will only propose how the challenges will be addressed in our subsequent chapters.

Many critics might say that proposing a list of desiderata for successful "political thought experiments" does not suggest that all conditions can be met during thought experimentation. Similarly, Matulovic (2017) acknowledges these concerns while stating that "Miščević's standard for successful thought experiments is so high such that most famous thought experiments could not satisfy" (Matulovic, 2017, 366). However, Miščević has discussed some of these difficulties, as presented above. The thought-provoking question is, do these desiderata

offer a reliable assessment of political thought experiments? How can they be used to justify the relevance of hypotheticals in political philosophy? In other words, in what way would these desiderata be defended as a reliable means for normative truth in political philosophy?

In the sections below, I draw my attention to the various challenges and shall critically examine the extent to which the requirements of these desiderata would affect our study and how they can be met in our subsequent discussions by some thought experiments in political philosophy.

2.5.1 Examining the Desideratum of Relevance

In this section, I claim that the desideratum of relevance can be met in political thought experiments. Meanwhile, Miščević asked the following questions concerning meeting the requirement of this desideratum: have all important aspects been taken into account in the construction of a given PTE? Are the variables or features that have been taken into account really relevant? (Miščević, 2013a 520). Miščević also claims that in the social contract, practical matters are also lacking, like "getting people ready for the contract, the issue of lying, the legislators inventing myths and so on" (see Miščević, 2018 161). By implication, Miščević thinks that in political philosophy the thought experiments scenario must include all relevant details to support its conclusion.

The question is, do we need practical matters to demonstrate the relevance or legitimacy of political thought experiments? Or does Rawls or Rousseau need a demonstration of the contract itself before an intuition would be evoked via thought experiments? The answer is 'NO' because political thought experiments are arguments. I take it that a realistic illustration of a contract might not be important for a valid argument to be reached via thought experimentation, particularly in Rawls's case. Meanwhile, in this thesis, I will be arguing that political thought experiments contain disguised arguments; this understanding offers a plausible explanation of why a practical contract might not offer a relevant factor for thought experimenting in political philosophy. The idea is that if we understand the relationship between arguments and thought experiments in the field, therefore we do not need to worry about having a practical contract via political thought experiments. As arguments, I claim that imagined scenarios that form the basis of the social contract should be seen as premises that are used to justify a hypothetical contract during thought experiments, and this would only be accessed via mental reflection and not via physical demonstration of a contract. Therefore, in the social contract thought experiments, practical matters, or any form of practical contract or

premises leading to practical demonstration should be considered irrelevant conditions for such a project. In addition, this thesis will also defend the use of 'reflective equilibrium' as a method that can be used to show how the conclusion of a good PTE can be relevant to practical matters.

2.5.2 Examining the Desideratum of Detailed Coverage

In this section, I shall be defending the reliability of political thought experiments and will be stating why (in some genuine cases) the limitation of this desideratum should not be used against thought experiments in political philosophy. Meanwhile, Miščević believes that the desideratum of detailed coverage has not been fully met in Rawls's thought experiment. He argues this while claiming that "Rawls is very clear about obedience to the law in ideal theory; but very vague about the character of basic structure". He claims that "we do not know if it has to be statist or not" (Miščević, 2013a, 521). This thesis believes that there is no need to worry about this concern because to understand political thought experiments we must pay attention to the kind of argument they represent. In other words, successful political thought experiments have the actual political matters they want to defend or jettison by means of argumentation. The issue of the basic structure might not need to be captured directly by Rawls's idea of fairness as Miščević claims warrants. However, it is possible that such issues would be subsumed under the rational element between parties in the *Original Position (OP)* as acknowledged by Rawls. One would defend Rawls here by stating that any rational, free, and disinterested persons at the OP should not find it difficult in making just (fair) decisions concerning the character of the basic structure of society if need be, in so far as *the veil of ignorance* blinds them. The idea is that Rawls's *original position* as a political thought experiment is very clear about the issue of fairness in the OP. By implication, according to Rawls, fair decisions for political institutions, economic policies, and all individuals that form the character of basic structure would also be reached when the parties are under *the veil of ignorance.*

Miščević also thinks that "we should take into consideration all aspects of political arrangement; the width of coverage; including the richness of detail in the scenario imagined" (see Miščević, 2013a, 519). I contend that moral and political thought experiments generally need not address all moral or political concerns within a single thought experiment. Therefore, for a thought experiment to be successful, there must be what I would describe as *"intuitive boundaries"*, to provide the limit through which the intuitions from political thought experiments are supposed to operate. The idea is that thought experiments generally have intuitive boundaries, in other words, thought experimenters ought to provide few details or

might need to concentrate on a limited aspect of political arrangement in order to establish their claims. For instance, in the *Original Position*, it might not be necessary for Rawls to capture every detail about political arrangement generally in one thought experiment to establish his claims about just arrangement or his definition of justice for all society. The idea is that some details are left out during thought experimentation to avoid complexities or letting the audience into areas that might distract their intuitive understanding of the political arrangement in question, hence not all details might be relevant to the proposed arrangement. It is my view that Rawls thought experiment captures all the needed details for the notion of justice in the original position. These claims shall be made manifest in our subsequent chapter on political thought experiments as arguments, where I will examine the details of Rawls's argument on the issue of justice.

2.5.3 Examining the Desideratum of Realisticity

This desideratum requires that all forms of "imagined scenarios must hold possible in its own thought experimental world and that the proposed political arrangement has to be such that it can be implemented in the actual world starting from the present situation" (Miščević, 2013a, 519). Miščević thinks that this would be possible if political thought experiments become motivational and/or institutionally possible. Miščević also wants political thought experiments to be stable (see Miščević, 2013a, 519). The claim here is that in successful PTEs if the hypothetical political institution is made actual then that institution should also be stable.

The thought-provoking question is to what extent would a thought experiment in the discipline satisfy this desideratum? First and foremost, I claim that moral and political thought experiments should not be rejected based on this desideratum because: (i) the conclusion of some successful PTEs can be implemented in the actual world; (ii) some political thought experiments can indeed be motivationally and institutionally possible; (iii) they can likewise be stable if implemented. For instance, Rawls's thought experiment is capable of evoking both motivational and institutional possibilities as also required by this desideratum. In this thesis, I will also be discussing how Rawls thought experiment can be hoped to motivate its audience based on its narrative ingredients. I will also be examining how to redesign future political thought experiments to facilitate 'narrative transport' as the required condition for motivation. Secondly, on the institutional possibility of political thought experiments, I will also be defending how PTEs can be institutionally possible in providing actionable policies to political institutions. The thesis will show how some relevant factors such as collective action, rule of

law, and freedom of speech can be deduced from Rawls's thought experiment, thus, these factors will be used to discuss the institutional possibility of PTEs. Using Nigeria as a case study, ideas from this discussion will show how the government ought to implement policies based on the logic of the collective action approach in the fight against corruption. In addition, I will show how the freedom of speech and the rule of law principles can be implemented to check corruption starting from the present situation in Nigeria as suggested by Rawls's thought experiment. The discussions of this desideratum are captured in chapters eight and nine of this thesis.

Similarly, this desideratum requires that any political arrangement that is proposed during thought experimentation must be stable and lasting for an extended period. The thesis believes that the proposed arrangement from most PTEs can be considered a long-term structure, meaning it can be carried out for a long time. While discussing the institutional possibility of Rawls's thought experiment, the thesis will demonstrate this view in chapter nine to show how ideas from PTEs can be sustained for a long time. For instance, as deduced from Rawls's thought experiment, I will show how relevant factors from the *original position* can be considered as a formidable tool or as a key instrument for anti-corruption reforms in Nigeria, and I believe that if implemented as policy guidance it can aid the fight against corruption for a long time. I will show how factors from Rawls's thought experiment can provide a sustainable approach in the fight against political corruption in Nigeria, which I believe could offer a possible solution to the problem of corruption in the country for a long time.

2.6 Summary and Conclusion

Miščević's account of hypotheticals has indeed shown significant breakthroughs concerning the usefulness of thought experiments in political philosophy. His definition of thought experiments in political philosophy is indeed useful, particularly on what should count as political thought experiments which also contributes to the main focus of this thesis. His desiderata for successful hypotheticals in political philosophy has also provided useful insight into the general scope of political thought experiments. As such, Miščević's work is considered as the pivoting point through which the primary purpose of this thesis shall be pursued. From Miščević's account, there are essential points I consider useful that need highlighting to defend PTEs in this thesis. (i) It can be deduced that for an imagined scenario to be considered as a political thought experiment, such a situation must be dealing with hypothetical cases that explain facts concerning a particular or general political arrangement. (ii) Hypothetical

scenarios will be relevant in political philosophy, if and only if there is an established direct concern with issues to do with justice, right, liberty and other relevant concepts for the development of our political life. Meanwhile, I wish to acknowledge important points of agreement on Miščević's account of PTEs, especially some of his claims on the desiderata of PTEs. I agree with his claims on desiderata of relevance, realisticity, especially his claims on both institutional and motivational possibilities including the claim on the stability of arrangement via PTEs. While agreeing with the above desiderata from Miščević, I wish to disagree with the claims on detailed coverage as I think it should not be considered as a necessary requirement for successful PTEs (see reasons as discussed in 2.5.2). Similarly, his claims on the accessibility of PTEs should be jettisoned. In the case of **accessibility possibility**, Miščević wants to know whether current institutions can be changed into imagined ones. The question is what relevance will this bring to the general purpose of PTEs and of course our political situation?

In this thesis, I will be responding to this concern using the institutional possibility of Rawl's thought experiment to show how features from a typical political thought experiment can be used to describe a better way of organising society. Discussions surrounding this concern shall be explained in chapter nine, to demonstrate how it could be possible to change our current political society, in particular how the features from Rawl's thought experiment can be used to explore how PTEs might be useful in bringing about actual political change.

To date, Miščević's project on political thought experiments has faced several criticisms that I have not mentioned in this chapter. In the next chapter, I will draw attention to some of these concerns that could be counted as serious problems against his desiderata of political thought experiments and I will also offer some responses to them in other chapters as I proceed.

CHAPTER THREE

3. CRITICISMS OF MORAL AND POLITICAL THOUGHT EXPERIMENTS

3.1 Introduction

Intuitions elicited via moral and political thought experiments (PTEs) play a crucial role in philosophical reasoning, especially for both political theorizing and normative debates. However, much remains unclear about the nature of the intuitions they evoke and their usefulness, in particular, their usage for normative evidence has been subject to serious scrutiny by mostly empirically-minded scholars (see Thaler, 2016; Klampfer, 2017; Tobia, 2016; Machery, 2011). To date, there are controversies regarding whether intuitions elicited by unrealistic hypotheticals in PTEs can ever be used to guide action, in particular, in responding to political issues in the real world as we know it. Critics have claimed that thought experiments in political philosophy are problematic because intuitions from these disciplines are untrustworthy and that the extension of thought experiments to political philosophy must be rejected. This chapter explores the criticisms and problematics levelled against moral and political thought experiments, in a bid to chart a course for which they can be seen as arguments, necessary for normative debates or as tools for the development of political theories.

Briefly, Thaler (2016), in his criticism of political thought experiments, argues against the use of unrealistic hypotheticals in political theory. He thinks that intuitions from unrealistic hypotheticals could not be substantiated in the real world. He argued against the use of abstraction and claimed that unrealistic hypotheticals in political philosophy abstract us away from real political issues. As such, political thought experiments cannot be used in responding to serious political issues and debates in the real world as we know it. For Thaler, thought experiments must be considered untrustworthy and stand to be rejected in political philosophy.

Similarly, Klampfer (2017), in his works has also criticised the use of thought experiments in political philosophy. Klampfer, on several points, has demonstrated the weaknesses of intuitions elicited via PTEs. He thinks that conclusions evoked via political thought experiments are unstable. For him, political thought experiments produce indeterminate results, whose conclusions are often controversial and divisive among readers, meaning such intuition evokes varying responses from different people. Klampfer also argues that thought experiments in political philosophy are prone to different extrapolation and generalisation; perhaps readers of such thought experiments do find it very difficult to understand what exactly to intuit from

them. He uses these claims to argue that intuitions from PTEs could not offer a reliable means for normative truth as such thought experiments in moral and political philosophy should not be used as a legitimate means of responding to moral and political problems.

This thesis argues that both critics misconceived the purpose of thought experiments in political philosophy, and shall also suggest a reliable way through which readers of hypotheticals would assess the legitimate roles of thought experiments in political philosophy. The remaining chapter shall proceed as follows: section 3.2 presents Thaler's criticisms of PTEs and shall demonstrate the step-by-step arguments why Thaler thinks thought experiments in political theory are untrustworthy. Similarly, in section 3.3, I shall discuss the concerns from Klampfer's views and will outline his claims on why intuitions from moral and political thought experiments should be considered problematic. Section 3.4 shall summarise and conclude the significant findings of the chapter and how it contributes to the general understanding of this thesis, in particular, my defence for the use of thought experiments in political philosophy. The section shall also provide some suggested guidelines through which the criticisms of political thought experiments would be addressed.

3.2 Thaler's Account of Realistic Hypotheticals in Political Theory

Thaler (2016), in his criticism of political thought experiments, presents novel ideas against the use of unrealistic hypotheticals in political theory. He claims that the ultimate purpose of political theory is to provide an action-guiding role by telling us how to respond to various political issues in society and providing a meaningful explanation of how we should behave to live a better life. Thaler thinks that for thought experiments to be useful in political theory, they must be physically plausible, meaning their hypothetical scenarios must contain real and recurring features in the real world that is possible for us, here and now. He believes that this will enable political thought experiments to create moral/political intuitions that are possible for us, here and now such that they would offer action guidance on how we should live. Thaler claims that our moral intuitions are only trustworthy if we consider those hypotheticals that are very realistic in guiding our real-life actions (see Thaler, 2016, 1123).

Thaler presents his arguments in two ways. He first argues that abstracting away real-life details during thought experiments in political philosophy can be morally dubious. He explains this via a consideration of Susan Sontag's (2004), and Judith Butler's (2007) works on war photography. Sontag (2004) argues that "photographs have an insuperable power to determine what we recall of events, and for decades, photographs have laid down the tracks of

how important conflicts are judged and remembered" (Sontag, 2004, 1). Both Susan Sontag and Judith Butler's major contention is that war photographers decide which image to capture and which one can be left out, and that such 'framings' can manipulate the feelings of observers in problematic ways. For example, Butler pointed out that, while the American soldiers that died during the American war in Iraq and Afghanistan were mourned on the front pages of newspapers, the lives of the vanquished and the oppressed (citizens of Iraq and Afghanistan) were left out and considered 'ungrievable' (Butler, 2007, 951). Susan Sontag and Judith Butler's work is used by Thaler to argue that decisions to abstract away real-life details of any scenario, through *framing* or through *idealisation* in any thought experiments are not morally neutral. He thinks that facts that deal with human vulnerabilities and suffering should not be abstracted away. For instance, the *ticking time bomb* thought experiment which is used to argue that the prohibition of torture should not be absolute. Thaler observes that this scenario involves abstraction, by implication, the thought experiment does not allow the physical permissibility of torture. He argues that abstracted details in the ticking bomb case are morally significant in telling us how to address core societal issues as it affects us in reality, which means they should not be abstracted away. Thus, many thought experiments in political philosophy risk this.

The second way Thaler presents his argument is by assessing how thought experiments are supposed to provide us with action-guiding moral norms. He argues against those he describes as 'clarificationists' defenders of thought experiments who believe that hypothetical cases do not need to be realistic in order to be relevant (McMahan, 2003; Steinhoff, 2013). For example, McMahan (2003) believes in the use of abstraction during thought experiments because he thinks that the primary purpose of hypotheticals is to filter out irrelevant details that might distract the reader from gaining the right intuition. McMahan, for example, uses the '*Ticking Bomb Case'* to explain the moral permissibility of torture as against the anti-torture absolutist claims. He thinks it can be better to consider thought experiments than real-life cases, considering such issues as real-life cases always include distracting empirical details that can be abstracted away during thought experimentation. For example, in the *Ticking Bomb Case*, the empirical details such as the fear of torturing the wrong victim, the physical objects that can be used for torture, the right person to torture a victim, and other physical consequences one might incur as a result of torture must be abstracted away during thought experimentation. Thus, by abstraction readers can set aside all irrelevant empirical details which can enable the audience to intuit correctly to reach a sound conclusion. Similarly, Steinhoff (2013) has also defended the use of unrealistic thought experiments in political philosophy. He also uses an

unrealistic thought experiment to argue against antitorture absolutism, which suggests that torture can be justified, particularly in some circumscribed cases of self-defence. He argues this using '*the Innocent Jenny Case*: an unrealistic scenario that involves torturing an aggressor in the form of rape for the purpose of self-defence. He uses this scenario to show how a "self-defensive rape" can be morally permissible even though the thought experimental scenario is unrealistic (Steinhoff, 2013, 149). Steinhoff believes that the major role of unrealistic hypotheticals in *the Innocent Jenny Case* is to evoke a moral intuition that can be compared to real-world cases of self-defence. With such existing parallels, Steinhoff shows how a circumscribed case of torture can be morally permissible, which means the anti-torture absolutist's thesis is flawed. What is common with the 'clarificationists' is that thought experiments are used to develop moral/rational intuitions that can be logically compared with our moral beliefs or principles that are permissible in the real world. This method can be described as reflective equilibrium: it is used generally in philosophy for the justification of claims if and only if there is that coherence between the moral intuitions and our moral beliefs or other acceptable normative principles in the real world (this method shall be discussed in details in section 5.6 of this thesis).

Conversely, Thaler uses both Steinhoff and McMahan's claims to argue against the 'clarificationists' thesis. He thinks that the imaginary cases involving both '*self-defensive rape'* and '*the Ticking Bomb'* thought experiments contain hypotheticals that share no similarity with the real world as we know it. They are not neutral and can be subject to scrutiny, in particular when we consider how practically they could occur in the real world. For example, *Innocent Jenny Case* is practically implausible because, for him, there are no real cases from which they imaginatively depart. Although the Jenny Case is not a real and recurring case, as such, Thaler thinks that it is problematically unrealistic because such a case is very unusual to be used as a justified means of self-defence. Thaler argues that such thought experiments only "summon us to imagine a moral universe that is different from the one we currently inhabit" (Thaler, 2016, 1132). This implies that such defensive rape cannot be relevant to the real world as we know it. To be relevant, Thaler thinks that hypotheticals that are used in political theory must engage the audience with practical guiding scenarios that could be recognisable as instantiations of real-world cases. This for him will help to direct the audience in solving political issues that are affecting them in the real world as we know it. His primary worry is that thought experiments in political theory present unrealistic scenarios that are quite imaginable, but their situations lack real and recurring features to explain how human actions could be guided in the

real world. As such, "its practical value for judging real-world cases is negligible" (see Thaler, 2016, 1134). Thaler thinks that 'real' and 'recurring' features are important in political thought experiments because his basic idea is that policies and laws need to be shaped to deal with circumstances that are fairly common rather than exceptional. Therefore, Thaler claims that it is through physically plausible thought experiments: possibly with exact instantiation in the real world that we could evoke the needed moral intuition that could guide the audience practically in solving political issues. He finally argues that in the current literature on political philosophy, there are no thought experiments that can fit this demand, which implies that unrealistic thought experiments are not useful in political philosophy.

3.3 Klampfer's Criticisms and the Problems of Intuition in Moral and Political Thought Experiments

Apart from Thaler's concerns on the nature of hypotheticals in political theory, intuitions elicited via moral and political thought experiments have continued to face serious criticisms from notable critics, in particular, from some empirically minded scholars. Starting from Klampfer's (2017) criticism of hypotheticals, this section provides some examination of the various ways intuitions from philosophical thought experiments have been criticised, especially, in political philosophy.

Klampfer (2017) has questioned the philosophical usage of intuitions, in particular, the ones that are evoked by means of thought experimentation in moral and political philosophy. He believes that any experiments, whether conducted in the laboratory or reader's mind, are supposed to yield intuitive results that are stable and should not be doubted. His concern is that intuitions generated through moral and political thought experiments produce indeterminate results, whose conclusions are often controversial and divisive; meaning such intuition evokes varying responses from different people.

To support his claims, Klampfer considers *The Violinist's Case* and *the Standard Trolley case* as examples of thought experiments that provide varying intuitive responses with indeterminate conclusions. In his view, the popular Judith Thomson's *Violinist case* and the *Standard Trolley Case* create controversies especially, when determining what should be considered as the right intuitive judgment from these thought experiments. For example, in *the Violinist case,* you are asked to imagine waking up in the morning to see yourself in the hospital, connected to a famous violinist. The Doctor tells you that the Music Appreciation Society has kidnapped you. On the impending death of the Violinist, you were hooked up to

him. If you are hooked up for nine months, he will be cured and you will not suffer any harm at the end of nine months. Thus, no one can save him except you. If you unplug yourself, you kill the innocent Violinist. The moral question is "do you have an obligation to stay connected?" (see Thompson, 1971, 47 for this argument). In assessing the varying responses such thought experiment might evoke, Klampfer evaluates a relatively recent BBC online survey posted by D. Sokol (2006). According to Sokol, while one in four of the 58,000 participants opted for 'YES', three in four opted for 'NO' (Sokol, 2006,1). This shows the divided opinion in people's responses. To strengthen his claims, Klampfer also evaluates the *Standard Trolley Case* to justify his argument on people's intuitive reactions. In the *Trolley Case*, we are asked to imagine as follows: "In the path of a runaway trolley car are five people who will be killed unless you flip a switch which will divert it onto another track, where it will kill one person" (Klampfer, 2017, 338). The question is: will you flip the switch? Klampfer claims that such a case posts an ethical dilemma because it evokes intuitive results that vary among different people. According to Klampfer, in another evaluation of *the Trolley Case* through surveyed data, responses from the non-philosophers show that "77 per cent of the total 65,000 responded 'YES' to flipping the switch while 23 per cent opted for 'NO' do not flip the switch" (Klampfer, 2017, 338). Also, using the same *Trolley Case,* the response from professional philosophers varies as follows: while 1,972 philosophers participated, 68.2% said 'YES' and 7.6% voted 'NO DON'T FLIP' and 24.2% remained either agnostic or undecided. Thus, while two-thirds agree that it is permissible to flip the switch in the Trolley case, only a minority departs from that. Klampfer is worried about "what ratio of 'YES' to 'NO' answer would be enough to validate such a conclusion?" (Klampfer, 2017, 338). He uses these two thought experiments to argue that political thought experiment being a sub-species of ethical thought experiments also shares the same fate. Klampfer thinks that because different people have differing intuitions regarding such cases they cannot be used as a reliable means of solving normative disputes.

Klampfer also identified other reasons why he thinks that thought experiments in political philosophy are problematic. For example, he claims that political thought experiments are prone to be affected by (i) multiple generalisations and extrapolation, (ii) undetected effects of bias and prejudice. Briefly, I will begin to examine these concerns in detail in the captions below:

3.3.1 The Issue of Generalisation and Extrapolation in Political Thought Experiments

Klampfer is also concerned that it can be unclear what generalisations readers should draw from moral and political thought experiments. Klampfer argues this using the "Pond case" by Peter Singer. For example, in *the Pond and the Envelope Case* by Peter Singer, you are asked to imagine that you are a professor, dressed up in new clothes and new shoes. You have a lecture in ten minutes, on your way to the lecture you are walking by a lonely pond where you spot a drowning child. Without thinking, you jump inside the pond and save the child because the child's life is more important than your new clothes or having a successful lecture. In another case, imagine when you return home to receive an envelope from the United Nations International Children's Emergency Fund (UNICEF) - a charity organization asking you to donate a small amount of money that you can easily afford in order to save starving children in Africa. It is clear that you have to save a drowning child, but do you have to donate to UNICEF? (See Singer, 1972, 230-232). Singer uses both cases to establish some kind of parallel claims in order to present his argument that promotes charity for the purpose of saving lives no matter the distance. Klampfer thinks that drawing a conclusion from *the Pond Case* thought experiment could be problematic because sometimes the audience finds it very difficult to understand what exactly to intuit from it. He thinks that three intuitive generalisations are possible (i) that "I ought to save the child drowning in front of me, (ii) that, in general, everyone in a position to do so ought to save children from drowning (iii) that one ought to prevent something bad from happening as long as she can do so without sacrificing anything of comparable value" (see Klampfer, 2017, 339). Klampfer claims that all that we clearly intuit in the Pond case is that we ought to pull the drowning child out of the pond since nobody else is around to help and we can rescue the child at an insignificant cost. Everything else is extrapolation and generalization and insofar questionable (Klampfer, 2017, 339). Klampfer uses this claim to argue that thought experiments in political philosophy are prone to such generalisation. He contends that the fact that possibly there are multiple generalisations in moral and political thought experiments means they can mislead the audience over the right moral or political action to follow. Therefore, thought experiments in moral and political philosophy must be rejected.

3.3.2 The Effects of Bias and Prejudice of Intuitions in Political Thought Experiments

Klampfer (2017) also believes that there are indications that different factors about human nature could influence our intuitive judgments during thought experimentation. For instance, he thinks that we are not completely free from bias and prejudice and that factors such as age, our cultural and demographic background, language, the order of presenting a thought experiment, framing effects, physical and social environment, gender, race, the training and the personality of the reader could influence our decisions during thought experimentation. He contends that our biased nature contributes to our political actions; and could as well distort our moral behaviour and political decisions during thought experimentation in political philosophy. The idea is that since we could become manipulated by these factors during moral and political thought experiments; therefore, our intuitions could also be manipulated; perhaps such intuitions cannot be considered as a reliable indicator for normative truth. Klampfer uses these claims to argue that the various evaluations derived from moral and political thought experiments, raise the concern of whether reliable and credible intuitions can ever be evoked by means of moral and political thought experiments.

3.4 Other Criticisms of Intuitions from Philosophical Thought Experiments

While concerns about the ability of intuition via thought experimentation continue to post serious epistemic damage to the usefulness of thought experiments in political philosophy, it has also received wider attention from different critics. Different empirically minded scholars have also joined voices with Klampfer to provide other reasons why the use of intuitions for philosophical analysis is generally problematic. They have supported their claims by highlighting different factors to show why intuitions cannot be considered as a reliable indicator of philosophical truth via thought experiments. These factors include our cultural and demographic background, language, the order of presenting a thought experiment, framing effects, physical and social environment, gender, race, age, the training and the personality of the reader, and so on (see Tobia, 2018). They believe that our intuitions can easily be influenced by these factors which place philosophical intuitions as a controversial phenomenon and perhaps an unreliable means for normative judgment.

In identifying the controversial nature of intuitions, Tobia et al. (2018), claim intuitions are unreliable because "they can be influenced by other factors like the cultural background of the reader even though they are obviously irrelevant to the truth of the content of the intuition"

(see Tobia et al., 2018, 372). Tobia buttresses this point by citing Weinberg's contentions on philosophical intuitions. Weinberg (2001) pointed out that intuitive outcomes from American students with East Asian cultural backgrounds have significant differences from other students with European cultural backgrounds. This finding shows that our cultural background can affect our intuitions, especially during a philosophical argument. Using *the Trolley case* as an example, Alhenius et al. (2012) also extended the argument on how culture can influence our intuitions during thought experiments to show that there is a cross-cultural difference in how individuals may react to such a complex ethical case. In surveying the moral judgment and moral disagreement of both the Chinese and the Americans, Alhenuis noted that the Chinese are less utilitarian than the Americans. From his finding, he concluded that the "Chinese are more prone to accepting some form of non-consequentialist ethics than the Americans. Thus, these cultural differences affect their intuitive judgment in *the Trolley Case"* (see Alhlenius, 2012, 196).

Friedman (2012) has offered another reason why we need to be sceptical about the usefulness of intuition in philosophical analysis. He claims that the epistemic intuitions of philosophers differ from non-philosophers. He contends that intuitions are prone to be influenced by one's academic discipline because those academic philosophers who are more conversant with philosophical training respond differently from how nonphilosophers could respond due to their training in academic philosophy. To reinforce this point, Tobia (2016) has reported that in the philosophy of religion, the epistemic intuitions of atheists vary from those of theists. He contends that "personal beliefs are inappropriately shaping argument analyses; they are subject to a pernicious influence of religious belief and by default 'infect' our philosophical analysis" (Tobia, 2016, 57).

There are also claims that age (a 45-year-old adult versus a 12-year-old child) can influence our intuition during philosophical thought experiments. Stich et al. (2004) report that older people have more experience than young ones and their cognitive abilities are likely not going to be the same. He stated this while assessing people's intuition on the traditional justified-true-belief account of knowledge. Using the '*fake-barn cases',* he submits that "an older person will be the more cautious one in attributing knowledge when potential defeaters are salient". The *fake-barn case* is used by Stich to argue that the "younger people seem to have little or no problem counting a fake-barn case as a genuine case of knowledge, regardless of the number of potential defeaters. Older people, on the other hand, are substantially less inclined to judge that these cases count as knowledge" (Stich et al, 2004, 200). I agree that age

can sometimes determine the way we think, as it is impossible for a twelve-year child to think the same way as a forty years old man. But these differences I can say are merely biological and might depend on one's cognitive abilities and mental development. I subscribe to the view that age should not limit us from gaining an objective or stable intuition from thought experiments if we follow the right argument through which thought experiments are developed. However, in the future, the younger person will always grow to have the same intuition as others in so far as the premises of the thought experiments remain the same.

Language is another factor that has been identified as capable of affecting our intuition during philosophical thought experiments. Costa et al (2014) have also reported that the epistemic intuition elicited by philosophers differs depending on their native language. Costa et al, (2014) claim that moral intuition evoked by different people varies as a result of the language used in presenting a thought experiment. Costa has a way of interpreting *the Trolley case* in various languages to examine how readers respond to it. From his findings, he claims that "assessing thought experiments with different languages reduce the emotional response elicited by the reader of the foreign language, consequently reducing the impact of intuitive emotional concerns" (see Costa et al., 2014, e94842).

Machery (2011) is also sceptical about intuition from philosophical thought experiments, in particular, the importance of the epistemic status of the judgments they elicit. Machery first of all acknowledges that thought experiments are supposed to be considered to play a distinctive role in contemporary philosophy because many influential arguments from thought experiments rest on premises supported by judgments elicited. He thinks that it is unclear whether the judgement elicited by thought experiments can provide evidence for the premises of philosophical arguments. Citing the Trolley case, the Gettier case, Mary the neuro-scientist and the Society of Music Lovers in ethics, he thinks that they are all unreliable because these thought experiments fail to provide evidence for the premises of philosophical argument (Machery, 2011, 191). His main worry is that the judgments elicited by these philosophical thought experiments are made from illusionary claims as they do not have any counterparts in everyday life, as such intuitions from these thought experiments in philosophy are untrustworthy.

In showing that intuitions can be shaped by extraneous factors, the various claims made by Tobia (2018), Costa (2014), Friedman (2012), Stich (2004), Costa (2014), Machery (2011), and Klampfer (2017) provide some serious reasons for concern. However, in this thesis, I will be arguing that the concern that different people have different intuitive responses to thought

experiments can be overcome and that political thought experiments can be reliable even though sometimes people report different intuitive responses to them. This thesis shall show how some moral and political thought experiments can be demonstrated as arguments, such that their scenarios can be reconstructed as premises that give rise to a logical conclusion. As valid arguments, the thesis believes that the judgment that is elicited can be considered trustworthy, which means their role in philosophical debates can never be underestimated. I will explain how I think such concerns can be overcome in chapter six.

3.5 Summary and Conclusion

In this chapter, I have shown the various criticism from Thaler (2016) and Klampfer (2016). Both critics have presented some thought-provoking issues against the usefulness of thought experiments in moral and political philosophy. They have raised the concern claiming that intuitions elicited via moral and political thought experiments are problematic. Although both arguments show clearly how thought experiments can fail I will argue in this thesis that their concerns can be overcome. While both criticisms of moral and political thought experiments remain thought-provoking, there are reasons to believe that some usage of hypotheticals for normative thinking should be considered trustworthy. In the next chapters (four, five, and six), I will begin to respond to some of these criticisms. In chapter four, I will defend the claim that some thought experiments in moral and political philosophy are arguments. As arguments, imagined scenarios used during thought experimentation are a representation of logical claims. A good political thought experiment can be considered equivalent to a good logical argument. Considering thought experiments as arguments will allow me to respond to the concerns discussed in this chapter. In chapter five, I will show how Thaler and Klampfer's concerns can be overcome. Chapter five will respond to Thaler's rejection of the use of unrealistic hypotheticals, in particular on the legitimate role of abstraction in political thought experiments. It will also offer a response to Machery's claims and other criticisms by demonstrating how the judgments elicited by successful political thought experiments could be linked to their counterparts in everyday life using the reflective equilibrium method. Chapter six will use *the argument view* to respond to Klampfer and other concerns against the nature of intuitions in PTEs.

CHAPTER FOUR

4. POLITICAL THOUGHT EXPERIMENTS AS ARGUMENTS

4.1 Introduction

In this chapter, I will be applying John Norton's claim that thought experiments are arguments to the field of political thought experiments and will also be defending the view that successful thought experiments in moral and political philosophy are identical to valid arguments. This will be used to develop the thesis that successful political thought experiments can also be reconstructed as valid arguments without any epistemic loss. As discussed in chapter two, thought experiments have played pivotal roles in various political debates, such that, via thought experimentation, their intuitions can be used as evidence for normative claims in political and moral philosophy (See Miščević, 2012; 2018). To date, the use of thought experiments in political philosophy remains controversial because of the various scepticisms and debates concerning their usefulness in political philosophy. For instance, as already demonstrated in chapter three above, Klampfer (2017), Thaler (2016), and other empirically-minded scholars have argued that intuitions from PTEs cannot be used as reliable indicators for normative truth. The claim that political thought experiments are arguments (as discussed in this chapter) will be used in chapter six as a theoretical background in responding to the various criticisms of political thought experiments as captured in chapter three. It will be used as a framework to defend the efficacy of PTEs in responding to normative issues and those concerns surrounding the unstable nature of intuitions elicited via thought experiments in political philosophy.

To defend my claims that political thought experiments are arguments, section 4.2 will focus on John Norton's *argument view*. The *argument view* is important in this chapter because it offers a theoretical background through which political thought experiments shall also be considered as arguments. In 4.3, I try to respond to the question of whether it is possible or not to consider a political thought experiment as an argument. I shall argue in this section that it is quite legitimate to examine thought experiments in moral and political philosophy as arguments because they contain imagined premises and conclusion that are similar to a standard argument, which means PTEs consists of hypotheticals that can be translated to premises and conclusions. Section 4.4 will present a complete defence of these claims by explaining how moral and political thought experiments can be seen as arguments that are ruled by a system of logical reasoning. Using *the Reconstruction Thesis* of John Norton, I will also try to reinforce

my claims by examining in detail how John Rawls thought experiment on *the Original Position* could be reconstructed as an explicit argument. In section 4.4.2 I summarize Rawls's thought experiment as an argument based on the reconstruction strategy. Section 4.5 provides critical comments on the possible implication of *the argument view* in the field of political thought experiments.

4.2 An Exposition of John Norton's 'Argument View'.

There has been some considerable debate about whether thought experiments should be considered as arguments (see Norton, 1991, 2004; Beck and de Wijze, 2015). John Norton is one of the philosophers who has defended the central thesis that thought experiments are nothing more than "dressed-up arguments" (Norton, 1991, 2004). In an attempt to respond to the epistemological problems of thought experiments, particularly how scientific thought experiments generate new knowledge, and how to validate them, John Norton presents a novel account of thought experimentation, a view that claims that the answer to the epistemological problems lies in the argumentative structure of thought experiments (see, Norton, 1991, 1996, 2004). Norton believes that only pure thought via thought experimentation cannot transform into new knowledge aside from logical truth that is derived through valid arguments. In buttressing this point Norton (2002) argues that pure thought can only transform what we already know; and if thought experiments can enable us to gain new knowledge, they must do so via argumentation. Norton's major contention lies in the fact that "the epistemic reach of thought experiments turns out to coincide with that of argumentation and that this coincidence is best explained by the simple view that thought experiments are just arguments" (Norton, 2002, 2). It implies that thought experiments can be presented to the reader in form of premises and conclusions. Some of these sentences are premise-like, which try to convince the readers of a given conclusion, just like formal arguments do.

What is an argument? To proceed, we must be able to understand what Norton means by "argument". Formally speaking, an argument consists of a set of premises and a conclusion, where the premises and the conclusion are inferentially related by some kind of logical reasoning. An argument can be said to be valid if its conclusion is derived from its premises. Also, an invalid argument is an argument in which the premises do not provide conclusive reasons for the conclusion. However, Norton adopts any inference that is explained by abductive, informal logic, inductive, and deductive premises as arguments (Norton, 2004a, 64). He believes that thought experiments are just picturesque arguments, but arguments that posit

some features; for example, he states that "they are arguments which posit hypothetical or counterfactual states of affairs, and which invoke particulars irrelevant to the generality of the conclusion" (Norton, 2004; Norton, 1991, 129). He believes that thought experiments provide us with information about the physical world and that this information does not come from new empirical data. According to Norton, "they are elicited from the information we already have by an identifiable argument. Such knowledge deduced from thought experiments is a result of transforming the tacit premises of an argument via logical reasoning. Meaning thought experiments are just cleverly disguised inductive or deductive arguments" that are governed by a system of logic (Norton, 1991, 129). It is important to note that since Norton's view is tied to a system of logic as explained above, what exactly does Norton mean by Logic? By logic, Norton implies a kind of systematisation that is in tandem with the distinction between form and structure. Because of such systematization, Norton has claimed that good thought experiments are those whose argument forms agree with their logic (Norton, 2004b, 1140). It, therefore, implies that to understand what hypotheticals must do in a thought experiment, all we need to know is to identify them as premises of an argument from where new knowledge can be established during thought experimentation.

Norton believes that it is only through the *argument view* that we can validate the internal structure of various scientific thought experiments. However, *the argument view* supports the idea that the knowledge we gain from thought experiments is already implicit in the premises of the arguments. This is why Norton claims that "a good thought experiment is a good argument, and a bad thought experiment is a bad argument" (Norton, 1991, 131). However, it is believed that for thought experiments to fail the argument must be invalid, which means that the thought experiments possess maybe some incoherent, unjustified, or inconsistent premises. This means they fail when their conclusion becomes illogical or when the premises also fail to support the conclusion. The implication is that a thought experiment must be evaluated as if it is a standard argument in order to justify its validity. Hence the claims that successful thought experiments must have valid arguments and fail if supported with illogical inference. In summarising his account, Norton makes five key claims:

1. ***The Identity Thesis:*** According to Norton, the identity thesis suggests that "thought experiments are mere picturesque arguments" (Norton, 2004b, 1139). Thus, a good thought experiment should be categorized as a good argument; and a bad thought experiment should also be seen as a bad argument. And this explains the simplest view that all thought experiments are identical to arguments and can be evaluated the same way as a standard argument.

2. ***The Reconstruction Thesis***: The reconstruction thesis implies that "all thought experiments can be reconstructed as arguments based on tacit or explicit assumptions. Belief in the outcome-conclusion of the thought experiment is justified only insofar as the reconstructed argument can justify the conclusion" (Norton, 2004b, 1142; Norton, 1996, 339). Norton shows that the epistemological problem of thought experiments can be overcome by reducing all thought experiments into some form of argument through logical reconstruction. The reconstruction thesis encompasses two fundamental theses, which are:

(i) *The Reliability Thesis:* The thesis, according to Norton, shows that "If thought experiments can be used reliably epistemically, then they must be arguments (construed very broadly) that justify their outcomes or are constructible as such arguments" (Norton 2004a, 52). Norton uses this thesis to demonstrate that a thought experiment is a reliable mode of inquiry if after being reconstructed as an argument justifies its conclusion.

(ii) *The Elimination Thesis:* Norton uses this thesis to explain the fact that any "thought experiment can be replaced by an argument without the character of a thought experiment" (Norton 1996, 336).

3. ***The Epistemic Thesis:*** This thesis explains that thought experiments have the same epistemic significance or epistemic reach as non-thought experimental arguments. In other words, thought experiments are not epistemically superior or inferior to their corresponding non-thought experimental arguments. They both share the same epistemological status, meaning after reconstructing thought experiments as an argument; both the argument and the thought experiment share the same epistemic impacts.
4. ***The Empirical Psychological Thesis:*** This thesis, according to Norton, presupposes the fact that "the actual conduct of a thought experiment consists of the execution of an argument" (Norton, 2004a, 50). This implies that to understand a thought experiment; we must have in mind that every thought experiment has the prospect of being considered as an argument.
5. ***The Empiricist Thesis:*** Norton uses this thesis to support the fact that "the result of a thought experiment must be the reformation of experience by a process that preserves truth or its probability, that is, by deductive or inductive argument" (Norton, 2004b,1142).

We should note that in his work Norton limits his account to scientific thought experiments. He defends his approach in several articles, by demonstrating how some of Einstein's thought experiments can be reconstructed as arguments. In this thesis, I will argue that Norton's approach can be extended to thought experiments in political philosophy. In the next section, I

will begin to look at some of the criticisms of *the argument view* and how Norton has responded to them.

4.3 Criticisms of Norton's Account Of Thought Experiments

The *argument view* presents a very broad understanding of thought experiments and arguments, but it has also been seriously criticized. According to critics of this view, there are several reasons to worry about Norton's claims. James Brown is one of the most persistent critics of *the argument view*. He argues that not all thought experiments in science should be described as arguments. Brown (1991) argues this by identifying some thought experiments he considered as "Platonic thought experiments". According to Brown, this set of thought experiments provides a priori knowledge by enabling "us to see with our mind's eye to grasp abstract entities in a non-inferential means which cannot be subjected to mere argumentation" (See Brown, 1991,77). As cited by Brown (1991), Galileo thought experiment on *the falling bodies* is the most beautiful example of platonic thought experiments. The *falling body case* involves the following:

> Imagine a heavy ball L and a small ball S, all made of the same material falling from the same height. In the Aristotelian view, L must fall faster than S because L is heavier than S. Assuming they are both joined together: since S is slower than L, it will reduce the speed of L, which implies that the speed of L+S altogether will be slower than the speed of L. On the other hand, L+S is heavier than L alone. If that is the case L+S must fall faster than L alone. (see Brown, 1991, 2).

Galileo uses this scenario to point out the contradiction in the Aristotelian theory; the contradiction is that the tied masses cannot fall simultaneously faster and slower than the large mass falling alone. Brown thinks that this 'destructive' part of the thought experiment can be understood as an argument. Galileo then goes on to use this thought experiment to generate a new idea for the motion of bodies, the idea that all bodies fall at the same rate. Brown thinks that this 'constructive' part of the thought experiment enables us to gain new knowledge, but this new knowledge is not generated as a result of an argument. He claims that "it is a priori knowledge of nature, since there are no new data involved, nor is the conclusion derived from old data" (Brown, 2004b, 1129). Brown uses this illustration to show how Galileo was able to rationally justify his claims via apriori evidence, which means there was no new empirical data involved in arriving at a conclusion and also claimed that "it is not a logical truth either"

(Brown, 2004b, 1129). Brown claims that Galileo's thought experiment was not raised from arguments, therefore, Norton's claims are incorrect.

Responding to these claims, Norton has argued that the only nonmiraculous means to get new knowledge about the mind-independent world without gaining data from experience is by logical manipulation of the existing knowledge that is presented in the premises of the thought experiments. Therefore, the only way to accept the "epistemic magic" via thought experiments is by "accepting the view that thought experiments are arguments" (Norton, 2004a, 45). His view is that we must agree that thought experiments are arguments because it is the only empiricist-friendly way of reasoning from old premises to gain new knowledge. Using his reconstruction thesis, Norton has gone further to reconstruct Brown's Platonic thought experiment as follows:

1. The assumption for reductio proof: the speed of balls of bodies in a given medium is proportionate to their weights.
2. From 1: if a large stone falls with 8 degrees of speed, a smaller stone half its weight will fall with 4 degrees of speed.
3. Assumption: if a slower falling stone is connected to a faster falling stone, the slower will retard the faster, and the faster speed, the slower.
4. From 3: if the two stones of 2 are connected, their composite will fall slower than 8 degrees of speed.
5. Assumption: The composite of the two weights has a greater weight than the larger.
6. From 1 and 5: The composite will fall faster than 8 degrees.
7. Conclusions 4 and 6 contradict.
8. Therefore, we must reject Assumption 1. (see Norton, 1996, 342-3).

To arrive, at the general claim that "all stones fall alike", Norton claims that further assumptions must be made. He went further to create the following assumptions, which he thinks, are implicit in the scenario.

8a. Assumption: The speed of fall of bodies depends only on their weights.

8b. Assumption: The speed of fall of bodies is some arbitrary monotonic increasing function of their weights.

8c. From 3, 5. If the function is anywhere strictly increasing, then we can find a composite body whose speed of fall is intermediate between the speed of fall of its lighter components.

8d. The consequence of 8c contradicts 8b.

9. From 8d. The function is constant. All Stones fall alike (see Norton, 1996, 342-3).

From the above reconstruction, Norton has argued and defended the view that Galileo's thought experiment is not above the interpretation of *the argument view*, thus there is no leap into the platonic world during thought experiments as claimed by Brown. However Brown seems not to have given up on his claims; he has maintained that even though some thought experiments can be reconstructed as arguments, he thinks that they are not essentially arguments (Brown, 2004, 1133). The *argument view* has also attracted criticism from other scholars. Some scholars have argued against the *identity thesis.* Haggqvist (2009), for instance, claims that thought experiments and arguments vary ontologically. He argues this by claiming that arguments are made up of certain determined propositions while thought experiments are not. One of his contentions is that thought experiments are made up of "some psychological process inside the thought experimenter's skull," which might not be equivalent to arguments. Haggqvist wants us to know that it might be very realistic to describe an argument as valid or invalid, while this may not be applicable to thought experiments (see Haggqvist, 2009, 60-1).

Sorensen (1992) has also put forward another thought-provoking question against the *identity thesis*. He believes that "thought experiments should be studied as if they were experiments". His major contention is that "thought experiments are arguments if and only if experiments are arguments" (Sorensen, 1992, 214-6). Sorensen first of all claims that it is possible to understand philosophical thought experiments by concentrating on their resemblance to real experiments. He thinks that thought experiments share many important features with real experiments. For him thought experiments are not just any kind of hypothetical reasoning about imaginary scenarios. For instance, like real experiments, thought experiments investigate what happens if we change some factors in a specific and controlled way within a given situation. Like real experiments, thought experiments are also based on background assumptions and background theories. Given the similarities, Sorenson believes that these are good reasons to treat thought experiments and real experiments equally. His main contention is that thought experiments are "a subset of unexecuted experiments" (Sorenson, 1992, 213).

I think that Sorensen's claim that thought experiments should be considered alongside real experiments is wrong. I worry that despite some well-established relationship between real experiments and thought experiments, they can never be the same as one involves counterfactual reasoning while the other does not. Sorensen's claims look more problematic because real experiments are done in the laboratory with laboratory equipment while thought experiments are not done in the laboratory, and as such do not need scientific equipment. Thought experiments are purely an experiment in thinking, an imaginative kind of investigating the universe, a very different approach compared to real experiments. A real experiment requires some laboratory apparatus like the pipettes, bunsen burner, test tube holder, test tube brush, ebulliometer, etc. while thought experiments involve no such apparatus, they are like an experiment (in thinking) using the mind.

Rachel Cooper (2005) has also offered her criticism against Norton's view. Her major concern is that some thought experiments only involve the reader's imaginative abilities as against being argumentative. She argues this using Hume's *Missing Shade of Blue* scenario as a typical example of thought experiments that can never be considered as an argument. In Hume's *Missing Shade of Blue*, "Hume asks us to imagine "what a missing shade of blue looked like without having seen it" (Hume, 1978, 6). Cooper uses this to argue that Hume's thought experiment does not sound like an argument because one can only consider what it is like to see blue via imagination and not by the process of argumentation. Therefore, the argument view must be rejected. One could say that Hume's missing shade is an introspective experiment, such that we use our own minds to see what is possible for human minds. Norton's focus is on scientific thought experiments and not introspective arguments as applicable in Hume's case. In Norton's view, all scientific thought experiments can be reconstructed as arguments, meaning all scientific thought experiments are not beyond the reach of the reconstruction thesis.

Micheal Bishop is another renowned critic of *the argument view*. He thinks that by examining the historical trajectory of certain thought experiments in science, they cannot be described as arguments. He claims that those cases where various thinkers "disagree over the outcome of a thought experiment" mean the argument view can be problematic (Bishop, 1999,535). In presenting his case, Bishop, while assessing the famous *Einstein clock-in-the-box thought experiment*, noted that both Bohr and Einstein use the same thought experiment to present two conclusions, which means their arguments were different. Bishop uses this view to argue that some thought experiments are prone to presenting two different episodes of arguments. Bishop wants us to know that "given the situation of having two arguments from

one thought experiment, implies that thought experiments cannot be an argument" (Bishop, 1999, 540). In response to Bishop's criticisms of Norton, Brendel (2018) offers a more comprehensive review of the merits of the argument view in particular, on the identity thesis. She reported that Edmund Gettier's thought experiments against the traditional analysis of knowledge have different versions, but they still obey the same logical structure. She argues that "the identity thesis would only be undermined if, for example, two tokens of one thought experiment type corresponded to two tokens of different argument type" (Brendel, 2018, 285). In responding to Bishop's criticism of *the identity thesis*, Norton (2004) has a convincing response, he claims that Bohr and Einstein ought to arrive at different conclusions "because of the differences in the premises pertaining to spacetime setting" (Norton, 2004, 63-4). He argues that both Bohr and Einstein use different scenarios involved in two thought experiments. While Einstein uses scenarios that involve classical spacetime, Bohr's scenario involves relativistic space-time. However, for the fact that different thought experiments can be narrated with different scenarios, Norton argues that it does not undermine the fact that they can't be reconstructed as arguments.

While the debates between James Brown, John Norton, and other critics of the 'argument view' are still ongoing; *the reconstruction thesis* continues to offer the strongest reason and explanation on why we must consider thought experiments as arguments especially, how their scenarios can be used to evoke new knowledge without new data from experience. Brown (1992) in the following rejoinder has admitted why Norton's account is hard to argue against "Norton says that thought experiments are often disguised, not explicit arguments. So the real claim is that they can be reconstructed along his empiricist lines. Existence claims like this are devilishly difficult to defeat. I doubt that an actual refutation could ever be delivered'' (Brown 1992, 275). Due to the strength of the reconstruction thesis, I shall also consider it in the field of PTEs to justify whether thought experiments in political philosophy can be reconstructed as arguments. But despite all the criticisms, Norton has continued to show the strength of his argument account by claiming that there are no thought experiments in science that cannot be reconstructed as arguments. This is premised on the fact that the account for him remains the best approach that provides a cogent view for testing both the internal and external validity of thought experiments. Meanwhile, even though Norton did not guarantee the use of *the argument view* in other disciplines, I do think that his account is not misleading and can be extended into thought experiments in moral and political philosophy. In the section below, I will try to demonstrate how moral and political thought experiments can be studied as

arguments. This will provide us with a background view for reconstructing some moral and political thought experiments as arguments, in order to justify the reliability of thought experiments in political philosophy.

4.4 A General Background of Moral and Political Thought Experiments as Arguments

Are moral and political thought experiments arguments? In this section, I try to answer this question by showing the various conditions under which all good moral and political thought experiments can be considered as arguments. By extending Norton's *arguments view* to the domain of political philosophy, I will provide a framework that will enable us to answer the above question. I contend that some political thought experiments are indeed a product of argumentation and are ruled by a system of logical reasoning. Although the character of argumentation may be disguised, they can easily be translated into explicit arguments, as suggested by Norton's *reconstruction thesis*. In the next section, I will examine the Rawls thought experiment on *the Original Position* as an example of a reliable political thought experiment within this category to show some of its logical or argumentative features. Using Norton's *reconstruction thesis,* I will show how it can also be translated into an explicit argument.

4.5 Political Thought Experiments as Arguments: Focusing on Rawls's Original Position (OP)

In this section, I apply Norton's reconstruction thesis and reconstruct Rawls's *Original Position* to its logical format. The core of this thesis is that Rawl's thought experiment contains a disguised or picturesque argument that involves tacit premises, assumptions, general principles, and hypotheses. Based on Norton's "reconstruction thesis", Rawls's thought experiment can be reconstructed as an argument. The reconstruction thesis states that a good thought experiment is argumentative and can also be reconstructed or translated into "picturesque arguments" (Norton, 1991, 1996, 2004). I will capture this view in Rawls's thought experiment to justify my claims. Rawls uses the *Original Position* to offer an elaborate set of arguments for justice as fairness. It is a hypothetical situation, used by Rawls to evoke a reliable intuition for a fair decision procedure based on rational choice. The *Original Position* also became a framework for various philosophical claims that formed the theoretical background in promoting his key ideas on the *principle of liberty and the difference principle* for both man and society.

However, it is pertinent to explain the thought experiment again in detail most importantly, to demonstrate the nature and purpose of its argument which will guide me while reconstructing it as an explicit argument. To explain Rawl's thought experiment, the background question is, how can a society be just? In responding to this question, justice for Rawls means a fair distribution of social goods, and he believes that with the help of certain hypothetical situations, one can achieve fairness both at the individual and the societal levels. He employs the use of a hypothetical situation called the *Original Position* to explain how a fair distribution of social goods can be met based on his two principles. In this case, Rawls asked us to imagine some group of men and women who find themselves in *the original position* to decide on what principles are to govern the distribution of justice in the new society. For Rawls, these are rational men and women with different personalities such as cultures, races, sex, income level, gender, religion, and personal choices. In the *Original Position*, Rawls claims that an egalitarian society can be reached by such an arrangement procedure using a set of knowledge conditions called "*the veil of ignorance"* (see Rawls, 1971, 118). According to Rawls, with the "veil", each party is temporarily ignorant of his/her characteristics. It is assumed that the parties do not know certain kinds of particular facts like their sex, race, or religion. First of all, "no one knows his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities, his intelligence and strength, and the like will all be blind" (Rawls, 1971, 118). It is believed that this will deny them all conditions and egoistic privileges that could enable them to tailor the social goods to their advantage. In this case, Rawls believes that after they decide on those principles, they are supposed to turn into the society they established. Rawls asked us to imagine what would that society turn out to be like. He argues that if those men and women at the OP are rational, a fair and equal distribution can be reached even though they have no idea of what will be their future or what they will become. Based on this uncertainty Rawls believes since they are rational they will be guided by the "maximin principle" while making a rational choice. The maximin rule is to guarantee some reasonable benefits even to the worst-off individuals. Rawls believes that since no party will want to be disadvantaged, they will decide to make an objective decision that makes them avoid losing out completely even when they finally see themselves as the worst off when they come out from the veil of ignorance. Rawls argues that they will consider his two principles to guarantee their equal liberty that is compatible with the liberty of all; and reasonable inequality in power, wealth, income and resources must be for the interest of all. This form of fair distribution of social goods is what Rawls described as justice in the OP.

4.5.1 The Original Position (OP) Reconstructed as an Explicit Argument.

Rawls uses the OP argument to show how political thought experiments can provide us with new knowledge on the issues of justice. It is important to note that, to conjure up new ideas in a thought experiment, using a model, it must involve a valid argument that is formed by premises and a conclusion. I shall justify my claims by reconstructing Rawls's thought experiment above into an explicit argument to show the extent of its validity. To say that Rawls's thought experiment could be reconstructed into an argument warrants that we examine some important aspects more carefully to replace them with assumptions (used also as plausible premises), general moral principles, and hypotheses to justify the argumentative form. To achieve this aim, I believe that the reader should remain patient with my translations to understand the logic of the *OP*. In reconstructing the thought experiment, I will begin by translating the *original position* argument considering the various plausible uses of premises (assumptions), and hypotheses within Rawls's thought experimental claims. From lines 1 down to 13 below, I will demonstrate how a valid (reconstructed) argument on justice can be established at the OP based on a fair distribution of social goods as outlined by Rawls in *the Theory of Justice* (1971).

Briefly, Rawls made it clear that different people are born into different positions; some positions, according to Rawls, have deep inequality that is determined in parts by political arrangement. Thus, one of the deepest inequalities is that institutions sometimes favour a particular starting point for some individuals against others. Justice as fairness for Rawls is to remedy this situation and perhaps an alternative account to classical utilitarian and intuitionist accounts of justice. He claims that "justice is the first virtue of institution, likewise laws and institutions no matter how efficient and well-arranged must be reformed or abolished if they are unjust" (Rawls, 1971, 3). Justice for Rawls is so important that even the entire happiness of the society cannot override it. Rawls thinks that an individual will be able to achieve his/her rational plans if society as a whole is ruled by fair decisions. Rawls criticizes the utilitarian notion of justice because he thinks that most of its policies are not rational and perhaps do not favour the interest of all. He claims that it is the case that sometimes "the sacrifices made by the few are outweighed by the larger sum of advantages enjoyed by many" which means people can be denied their true sense of justice (Rawls,1971, 3). It is against this background that Rawls seeks to answer one fundamental question; how do we attain a principle that is just? To achieve his major aim of establishing a just society, Rawls believes that we can have a fair distribution but only in a fair society. According to Rawls, a fair society is a society that meets

the conditions described by a "hypothetical world" he called *the Original Position* where men and women exist to decide what is to be the foundation charter of their society (Rawls,1971, 10). The *Original Position* is hypothetically created by Rawls to describe where everyone can settle their individual terms of social ownership and other basic terms of association to collectively participate in a new social contract (Rawls, 1971, 15-9). The above explanation can be translated into the following assumption used as part of the premises for Rawls's argument:

1) **Assumption:** The original position (OP) is a hypothetical place where rational men and women meet to decide their future terms for arrangement**.**

However, in the heart of justice as fairness is what Rawls considers as fair distribution in a well-ordered society. According to Rawls, a society is well-ordered when everyone accepts and knows that others accept the same principles too (Rawls, 1971, 4). His main concern is to establish a fair distributive system of social goods within a well-ordered society so that everyone can achieve his/her rational potential and goals. So any principle that is chosen must meet the expectation of all parties no matter who they are and where they will be finding themselves in the new world. What Rawls meant by justice is simply to have new principles as the considered choice to arrive at a fair distribution of social goods. In Rawls's words, "the social goods are the index of those goods which a representative individual can look forward to", such as "wealth", "liberties", "income", "rights", "freedom" (civil and political rights), and "the social bases of self-respect", etc. (Rawls, 1971, 79). Rawls made us believe that *the Original Position* does not have an actual situation that resembles it in the real world. It is to lay the foundation where all the agreements concerning the new principles (of justice) can be reached based on rational choice. However, Rawls considers parties in the original position as both rational and mutually disinterested men and women (Rawls, 1971, 12). Thus, rationality is key to the general successes of the OP, and Rawls believes that parties would have a general knowledge of the various "principles of justice," "principles of economic theories", the "basis of social organization," and the "law of human psychology" (Rawls, 1971, 119). He thinks that, at the original position, "a person's good shall be determined by what is for him the most rational long-term plan of life given reasonably favourable circumstance" (Rawls, 1971, 79). Of course, they are to apply their rational knowledge to directly engage in core decision-making and to manage the ideas surrounding the contents of their plans if they finally settle on new principles.

There are further claims that Rawls wants us to consider in furthering the circumstances of justice in *the Original Position*. Rawls thinks that in *the Original Position*, it is naturally possible and necessary to consider the circumstances of justice as a normal condition under human cooperation. According to Rawls, the society is typically marked by a "conflict" as well as an "identity of interest", where men could be prone to consider larger to a lesser share (Rawls, 1971, 109). The implication is that due to the conflict of interest resulting from limited natural resources, men can naturally pursue their subjective interests to meet their own goals in life. These plans, according to Rawls, lead people to different ends and conceptions of goods. The effect of this is that men will suffer from various shortcomings when making judgments, and men's judgments can likely be distorted by bias because of their feelings and/or desire for their own affairs (Rawls, 1971, 110). Rawls claims that some of these defects are caused by "moral faults, selfishness, negligence, and to a large degree men's natural situation" (Rawls, 1971, 110). Rawls pointed to the fact that "if wealth, position, and influence, and the accolades of social prestige, are a person's final purposes, then surely his conception of goods is egoistic" (Rawls, 1971, 111). Due to these subjective circumstances of men, Rawls believes that parties at the OP can naturally become biased, and of course, such motivations could set in as a constraint for fair distribution (see Rawls, 1971, 109-11). I hereby translate these claims to the following assumption:

2) **Assumption:** The decision procedure at the OP could be influenced by personal choices and self-interests, which causes biases against the plan for a fair distribution of social goods in the OP**.**

Remember also that Rawls considers *the Original Position* as a position that should guarantee justice against biased situations. In other words, the *Original Position* is supposed to set up a pure process that guarantees a good outcome; in this case, justice. Also, Rawls thinks that within the process, constraints of *the Original Position*, which include injustice are irrational and therefore, won't come about. Rawls thinks that it might be possible that selected parties might still be influenced by existing institutions and personal bias, as such a new ethical arrangement is needed to guide all decisions, discussions, and agreements in *the Original Position.* This will help to checkmate particular inclinations and aspirations of a person's conceptions of goods in order not to affect the adopted new principles (Rawls, 1971, 16). However, to avoid any influence from any social institution and to propose principles for rational acceptance, Rawls believes that such a principle must be mutually agreed to and be

free from bias to guarantee fair decisions, respect for people's social status, and rational choice. Rawls claims that if an ethical principle is chosen by those taking part in it, it means, it could be decided by the self-effort of those who are interested, which might result in another conflict in *the Original Position*. Rawls believes that the issues relating to bias in the *Original position* can be ruled out by *the Veil Of Ignorance* (Rawls, 1971, 118). Rawl's "veil of ignorance", happens to be the first hypothesis that will check all ethical concerns that might corrupt the mind of parties or make them become biased. So by it, they will all be blinded by all personal characteristics. These personal characteristics are what Rawls considers as what could lead men to tailor their particular circumstances in favour of their egoistic wants in the *Original Position*. These claims can be translated as follows:

3) **Assumption:** The *Veil Of Ignorance* can make parties at the OP become blinded by biased inclinations, including their personal choice and conception of goods.

What is the *Veil of Ignorance* as mentioned above, and how is it supposed to work at the OP? It is a set of knowledge conditions; hopefully, one that exists only in the hypothetical world as such it does not have an actual situation in the real world. Its main duty is to stand as the only existing general hypothesis upon which every other decision can be objectively assessed. Freeman (2003) also buttresses this point while acknowledging that with *the Veil of ignorance* Rawls believes that "decision on the principle of justice can become ahistorical, which means parties would make decisions strictly impartial with respect to people's social status, natural characteristics and abilities, and even their conception of the good" (see Freeman, 2003, 11). Rawls believes that *the Veil Of Ignorance* is supposed to influence the cognitive thinking of the parties by making them forget their previous status in life which often makes them biased. By that, Rawls believes that parties in the original position can be free from biased decisions. These claims can be logically translated as follows:

4) From 3 and 2, if parties can pass through the veil of ignorance then the parties in *the Original Position* can become blind to personal biases.

As a follow-up to the above claims, Rawls also introduced the idea of social cooperation based on reciprocity as part of his argument at the original position, perhaps the cornerstone for making fair decisions based on mutual interest. Rawls believes that since they are under the "veil", parties will engage in reciprocal exchange; as a result of the fear of becoming disadvantaged when they are out of the veil and that parties will consider exchanging their

conception of goods for mutual benefits. Rawls believes that, of course, reciprocal cooperation based on mutual advantage is morally binding to parties and can be considered as a condition under which parties at the OP are obliged to follow because they are blinded to understanding their characteristics by the veil. Rawls explains this while stating that "parties will not gain at another's expenses since only reciprocal advantages are allowed" (Rawls, 1971, 89). Furthermore, Rawls's thinking is that from the standpoint of the *Original Position*, and because of the veil, "the parties would reject the principle of utility and adopt the more realistic idea of designing the social order on the principle of reciprocal advantage" (Rawls, 1971, 155). It is rational to do so because, with their mutual interest, parties begin to think that all decisions reached must be for the good of themselves. Thus, this can be translated as follows:

5) **Hypothesis 1:** The Principle of reciprocity shall guides parties to protect their self-interest based on fair share and mutual advantage.

6) From 4 and 5, if parties are blinded to their personal biases by the veil of ignorance and they reason from the principle of reciprocity then parties shall pursue their self-interest at the *Original Position* on a mutual advantage that can be considered as fair

Rawls further made it clear that without the principle of reciprocity, his principle of justice could be fragile if not impossible (Rawls, 1971, 433). However, the fact that parties are not biased and are in agreement to work together for the best interest of everyone, then they will now be pursuing a fair arrangement. The implication is that they will agree on the rules of engagement for reaching fair decisions for everyone's best interest. Let us briefly consider how the principle of reciprocity is supposed to work, for instance, if Mr X is to share 100 pounds of which he is also part of the shares with maybe the other four beneficiaries including Mrs A; Mrs B; Mr C; Mr D (the 100 pounds might be their social surplus resulting from their cooperation). Remember, Mr X must be the last to choose after sharing. Given such a level of complete ignorance of what might be the leftover when others have chosen or the fact that unfair shares may affect him or any of the parties. To avoid this, the sharer must work within the principle of rational choice in order to arrive at an equal sharing formula, or any unequal distribution, in this case, must be for the benefit of all. This is the kind of rationally self-interested process that Rawls wants us to think about, especially if we want to design a society that yields impartial judgments once decision-makers abstract away from their particular aims of biases. The intuitive idea is that while assessing situations from an impartial point of view,

Rawls believes that an unequal share from Mr X if any should be for the objective interest of all the parties. However, Rawls believed that by virtue of *the Veil of Ignorance*, it implies that "nothing shall tempt them to exploit social and natural circumstances to their own advantage"; therefore parties can begin to see things objectively from the point of general consideration (Rawls, 1971, 118). Rawls also claims that since the parties are rational, and because of the veil they will not know where they will belong in the new society, it now becomes insignificant to seek special advantages for themselves, or to a particular social status in the society. Rather they will be advancing the principle that will favour everyone, the rich, the poor, and including the least advantaged members of society so that everyone can equally be situated (including the worst-off) when they come out from *the Veil Of Ignorance*. This can be translated as follows:

7) **Assumption:** Parties at the *Original Position* are self-interest individuals and on a basis of a mutual advantage they will be working towards a fair arrangement of social goods for their future engagement in case they become worst-off individuals.

8) From 7, 6 and 1, if parties are self-interested individuals working towards a fair arrangement of social goods, based on the principle of reciprocity and are rational by nature then they will consider an arrangement that will protect their future even if they happen to be among the worst-off persons when they come out of the veil of ignorance.

How then can we justify that decisions that will be reached are impartial, rational, and fair? And how will parties consider a fair arrangement that will protect their future even if they happen to be among the worst-off persons when they come out of the veil? These are some of the challenges that parties faced in situations that are full of uncertainty at the OP. Rawls applies the "maximin principle" as a rule that will guide their choice in making a fair and rational decision on their basic liberties and fair equality of opportunities, in particular, when they have no idea of what the outcome might be as a result of their collective choice. The maximin criterion according to Rawls is generally understood as "a rule for choice under great uncertainty" (Rawls, 1971, 71). However, Rawls assumed that rational men would go for a fair distribution pattern, but the question is, how can we ascertain that such decisions reached in *the Original Position* would be fair? Rawls has provided a further illustration to justify his claims; given the fact that by virtue of *the Veil of Ignorance*, parties will not be able to win a

special advantage for themselves, Rawls argues that this might set in the state of uncertainty in the *Original Position*. Because they lack information about what they might become in the future, particularly when they are out from *the Veil of Ignorance*. Rawls claims that because of such uncertainties and the fact that parties are rational individuals, they will take into consideration what he describes as the "maximim principle" as a rational rule to facilitate what constitutes a fair principle for *the Original Position* (Rawls, 1971, 132, 153). He thinks that parties are meant to "direct their attention to the worst that can happen under any proposed course of action, and to decide in the light of that" (Rawls, 1971, 132, 133). According to Rawls, using the maximin rule the basic liberties and fair equality of opportunities which they seek to achieve can be secured even if they happen to be the worst off. Therefore, to ensure that their undertakings are not in vain, parties will have to consider what he describes as the two principles of justice (basic liberties and fair equality of opportunities) as the possible solutions for future partnerships. Rawls believes that this could help persons in the OP have an (i) equal right to the most extensive basic liberty compatible with a similar liberty for others (ii) social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be everyone's advantage and (b) attached to positions and offices open to all. He thinks that to secure their basic liberties and fair equality of opportunities that will benefit all even in their worst-off circumstances the two principles would be their best choice (Rawls, 1971, 266). It can be observed that without 'the maximin rule' the idea of justice will fail and the two principles of justice could never be possible, meaning the idea of fair distribution of social goods in the original position could never be reached. The above claims by Rawls can be translated as follows:

9) **Hypothesis 2:** The "Maximin rule" makes it possible for all parties to consider the two principles that could secure their basic liberties and fair equality of opportunities that will benefit all even in their worst-off circumstances.

10) From 9, if the two principles are chosen based on the 'maximin rule' (i) each person having an equal right to the most extensive basic liberty compatible with a similar liberty for others (ii) social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be everyone's advantage and (b) attached to positions and offices open to all.

Furthermore, Rawls claims that there comes a point where parties have to choose among various competing alternatives which he believes that as rational individuals the

maximin rule might be the best rational choice to guide them to overcome such challenges. Rawls argues at great length that the use of the maximin rule by parties in *the original position* has three important features: (i) That "the parties have no objective basis for any probability assignments" (Rawls, 1971, 134). (ii) "The person choosing has a conception of the good such that he cares very little, if anything, for what he might gain above the minimum stipend that he can, in fact, be sure of by following the maximin rule. It is not worthwhile for him to take a chance for the sake of further advantage, especially when it may turn out that he loses much that is important to him" (Rawls, 1971, 134). (iii) "other conceptions of justice may lead to institutions that the parties would find intolerable; in other words, the rejected alternatives have outcomes that one can hardly accept. The situation involves grave risks. For example, some choices would subject parties when they come out of *the veil of ignorance* to the risk of being extremely poor or even a slave" (see Rawls, 1971, 135). Rawls contends that having a satisfactory minimum from the maximin rules, "it seems unwise, if not irrational, for parties to take a chance that these conditions are not realized" (see Rawls, 1971, 135). The point is that Rawls argues that the parties may not be able to foresee the possible outcome of the society because of how probable certain conditions of the society might turn out to be in the future. Thus, if rational parties in the original position want to protect themselves against any contingencies or worst-off situations, the maximin rule becomes the plausible solution.

11) From 1, 8 and 9, if parties at the OP are rational and want to protect their worst-off situation when they come out of the veil then the "Maximin rule" is the best approach to guide their decisions for fair distribution plans in the OP.

Rawls argues that given the maximin features, "most extreme disparities in wealth and income are allowed provided that they are necessary to raise the expectation of the least fortunate in the slightest degree. But at the same time, similar inequalities favouring the more advantaged are forbidden when those in the worst position lose by the least amount" (Rawls, 1971, 136). The idea is that parties in *the original position* are obliged to consider a satisfactory minimum that will be in the interest of everyone in the future. These features have enormous philosophical underpinnings, which are believed to have shaped the principle of justice that would provide the best possible outcome in the future. It is believed that the maximin rule provides enough support for parties to believe that they no longer run at any risk because even if they become part of the worst-off, they will always have the chance of gaining fair equality of opportunity. It can be deduced that the maximin rule takes into consideration equal

advantages for everyone in *the Original position* in the face of possible natural inequalities when parties are out of *the veil of ignorance*. Given the above illustration, it is assumed by Rawls that the "maximim rule" is considered as the rational permissible "maxim" in *the original position* upon which "the two principles of justice can be subsumed" (Rawls, 1971, 135). This explanation can be deduced or translated as follows:

12) From 9 and 11, if parties in the OP consider the maximin rule, then the two principles are their best choice because it maximizes the best interest for all, including the worst-off.

Rawls's argument for the adoption of the two principles is a clear invocation of the maximin principle. The first principle is the liberty principle which is used by Rawls to show support for basic freedom and the right of everyone in the OP. It includes political freedom (the right to vote and be voted for), freedom of speech, right to own private property, freedom from arbitrary arrest, liberty of consciousness, religious freedom, and freedom of thought (Rawls, 1971, 53). These are the kind of freedom that is found in liberal democracies, and Rawls believes that parties will all agree to them as their starting point for the new society. In the second principle, Rawls believes that since his first principle guarantees moral equality for everyone, the "difference principle" acknowledges real-world situations where competition from various persons can lead to social and economic inequalities/ indifferences. By implication, if anyone happens to gain greater wealth, or is in a position of power as a result of the agreed principles, such benefit/position must only be justified if it is meant to bring a greater benefit to all. Rawls states that *"*all social values-liberty and opportunity, income and wealth, and the social bases of self-respect- are to be distributed equally unless an unequal distribution of any, or all, of these values, is to everyone's advantage" (See Rawls, 1971, 54). Given the maximin rule, the two acceptable principles can be translated as follows:

13) **Conclusion:** From 12 and 10, based on the two principles chosen (i) each person will have an equal right to the most extensive basic liberty compatible with a similar liberty for others (ii) social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be everyone's advantage and (b) attached to positions and offices open to all.

In line 13 above, it can be deduced that if parties decide to settle for the two principles via an invocation of the maximin principle, then Rawls's advocacy that the original position provides

the basis for fair distribution of social goods for all, including the worst-off persons in the new society can never be denied. The implication is that given the two principles, the original position provides the basis for the fair distribution of social goods for all, including the worst-off persons in the new society. However, in as much as we accept the fact that a fair distribution can be reached at the OP, and the worst-off are also taken into consideration by these fair distribution patterns, intuitively, Rawls claims on the possibility of justice at the original position is certain, meaning there is justice at the OP.

4.5.2 Summary of the Reconstructed Argument

Below, I summarise Rawls thought experiment as an argument. Of course, to understand the nature of its validity readers should take into consideration the various uses of assumed or translated premises, the hypothesis, and the conclusion as reconstructed from Rawls in his *Theory of Justice* (1971). Later in my critical comment, I will further analyze the validity of the original position argument in line with the rules of logic. Based on the above reconstruction, the original position can be summarized as an explicit argument as follows:

Reconstructing the Original Position Case as a Valid Argument

1) **Assumption:** The original position (OP) is a hypothetical place where rational men and women meet to decide their future terms for arrangement**.**

2) **Assumption:** The decision procedure at the OP could be influenced by personal choices and self-interests, which causes biases against the plan for a fair distribution of social goods in the OP**.**

3) **Assumption:** The *Veil Of Ignorance* can make parties at the OP become blinded by biased inclinations, including their personal choice and conception of goods.

4) From 3 and 2, if parties can pass through the veil of ignorance then the parties in *the Original Position* can become blind to personal biases.

5) **Hypothesis 1**: The Principle of reciprocity shall guide parties to protect their self-interest based on fair share and mutual advantage.

6) From 4 and 5, if parties are blinded to their personal biases by the veil of ignorance and they reason from the principle of reciprocity then parties shall pursue their self-interest at the *Original Position* on a mutual advantage that can be considered as fair

7) **Assumption:** Parties at the *Original Position* are self-interest individuals and on a basis of a mutual advantage they will be working towards a fair arrangement of social goods for their future engagement in case they become worst-off individuals.

8) From 7, 6 and 1, if parties are self-interested individuals working towards a fair arrangement of social goods, based on the principle of reciprocity and are rational by nature then they will consider an arrangement that will protect their future even if they happen to be among the worst-off persons when they come out of the veil of ignorance.

9) **Hypothesis 2:** The "Maximin rule" makes it possible for all parties to consider the two principles that could secure their basic liberties and fair equality of opportunities that will benefit all even in their worst-off circumstances.

10) From 9, if the two principles are chosen based on the 'maximin rule' (i) each person having an equal right to the most extensive basic liberty compatible with a similar liberty for others (ii) social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be everyone's advantage and (b) attached to positions and offices open to all.

11) From 1, 8 and 9, if parties at the OP are rational and want to protect their worst-off situation when they come out of the veil then the "Maximin rule" is the best approach to guide their decisions for fair distribution plans in the OP.

12) From 9 and 11, if parties in the OP consider the maximin rule, then the two principles are their best choice because it maximizes the best interest for all, including the worst-off.

13) **Conclusion:** From 12 and 10, based on the two principles chosen (i) each person will have an equal right to the most extensive basic liberty compatible with a similar liberty for others (ii) social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be everyone's advantage and (b) attached to positions and offices open to all.

4.6 Critical Comments

One can see quite clearly that the reconstructed argument above shows that the *original position* thought experiment by Rawls is a typical example of a valid argument; because the

condition for establishing a valid argument is satisfied. To understand its validity, readers must understand the tacit premises and various hypotheses which are at work in the OP as outlined by Rawls. The *Veil Of Ignorance* is the first hypothesis used by Rawls to make parties at the OP become blinded by biased inclinations, including their personal choice and conception of goods. The role of this hypothesis is used by Rawls as an ethical principle that will influence the moral judgment of parties, such that, parties in *the Original Position* can become blind to personal biases when making decisions. Similarly, the maximin rule is considered in the argument as the second hypothesis rational people could consider under such conditions of uncertainty. The maximin rule also provides a distinctive role of offering deductive support why the two principles are their best choice in case they become the worst off when they are out of the veil of ignorance. However, considering the complex nature of the OP, carefully we have translated the various claims by Rawls into different plausible, hypotheses, premises, and a conclusion has been reached based on logical entailment. Besides, looking at the entire argument, it is also clear that there was no epistemic loss which means the intended intuitive claims from the original scenario of the thought experiment are also maintained. By implication, in the conclusion, a just society via a fair distribution of social goods as demanded by Rawls is also intact. The reconstructed argument above, also shows that the original position is coherent and understandable and the two principles of justice do indeed follow via an invocation of the maximin principle. Going by Norton's claim that "a good thought experiment is a good argument, and a bad thought experiment is a bad argument" (Norton, 1991, 131), we can reasonably state that based on the valid argument above, the *Original Position* can be said to be a good and reliable political thought experiment.

However, let us assume that some critics of the *Original Position* might say that *the Veil of Ignorance* (in premise 3) is not physically possible; therefore, Rawls's argument is invalid. I take it that Rawls's argument is valid because the implicit hypothesis and its premises do not need to be physically possible for the thought experiment to be valid. What we mean here is that it is not a necessary condition for *the Veil of Ignorance* to be realistic before it can be used in raising valid arguments in political thought experiments as it is also applicable in most scientific thought experiments. The idea is that we need to distance our thoughts completely away from the actual situation of the *Veil Of Ignorance* to give the right argumentative attention to the scenario according to Rawls. As such, it will enable the reader to also gain the right intuitive support which is needed to understand the logic of Rawls' argument. There are important facts to be deduced from this chapter for our defence of thought

experiments in moral and political philosophy, (i) that some moral and political thought experiments are indeed a product of argumentation and can be reconstructed as valid arguments. (ii) Rawls's case presents a valid argument; as such, it can be said that intuition from it is therefore stable, objective, and uncontroversial. I will elucidate further on these claims as we proceed while responding to criticisms of political thought experiments in the chapter below.

CHAPTER FIVE

5. OBJECTIONS TO THALER'S CRITICISMS OF THOUGHT EXPERIMENTS IN POLITICAL THEORY

5.1 Introduction

As already discussed in chapter three (3.2), Mathias Thaler (2016) is sceptical about the usefulness of unrealistic thought experiments in political theory. He is critical of some thought experiments in political philosophy, in particular, those whose scenarios are completely detached from 'real' and 'recurring' situations as they occur in the actual world. He claims that the role of political theory is very practical: it is to guide us on how to live a good life in the real world. For thought experiments to be productive in political theory, hypotheticals must address practical issues that are connected to real and recurring situations in the world. Thaler also argues that unrealistic hypotheticals are not morally neutral in directing people's moral choices because important empirical details needed for moral considerations in the physical world are often abstracted away. Thaler's ultimate position is that for thought experiments to be useful in political theory, they must be practical enough or very realistic in such a way as to evoke trustworthy moral intuitions that are tied to action guidance. In this chapter, I intend to argue against Thaler's claims while defending the idea that political theorists/philosophers can legitimately use unrealistic hypotheticals to direct us on how to live a better life. Drawing from the method of reflective equilibrium, I aim to show that unrealistic hypotheticals in political thought experiments have a legitimate role to play in the development of political theory regardless of the bizarre nature of some of their scenarios.

5.2 Hypothetical Cases and Thought-Experimentation in Political Theory

Thaler's scepticism on the use of unrealistic thought experiments in political theory provokes the following questions: do we need hypotheticals with 'real and recurring' features (alone) in order to solve political issues via thought experimentation? Or can a trustworthy intuition ever be evoked from unrealistic hypotheticals for solving political issues in the real world? I contend that Thaler misconceived the purpose of thought experiments in political theory: therefore, his rejection of unrealistic hypotheticals in political theory is unjustified.

The concrete proposal of this thesis implies, in a nutshell, that Thaler's view must be rejected because the use of 'unrealistic hypotheticals' in political theory can be legitimately justified given that: (i) the purpose of political thought experiments involves an *experiment in thinking* as against an experiment of the 'real and recurring situations' of the world. (ii) PTEs

are arguments capable of providing stable intuitions that can be used to examine complex phenomena for an improved *understanding* of political issues. (iii) Their empirical details are irrelevant because their imagined premises/conclusion can remain unrealistic, yet the argument is considered valid. Therefore, the use of abstraction is legitimate because realistic hypotheticals are never a precondition for valid arguments during thought experimentation. And finally, (iv) unrealistic thought experiments can also be used to establish a *reflective equilibrium*: whose moral intuitions could be rationally compared to our moral principles to aid action guidance in the real world.

In defending the above views, this chapter will focus on four major goals. To avoid being misunderstood, the first goal of this chapter (5.3) is to elucidate the role of hypotheticals in PTEs using Rawls's thought experiment. This will provide a framework to demonstrate what unrealistic hypotheticals are and why we need them in political theory. In the second part (5.4), I will defend unrealistic thought experiments by cross-examining their relationship with '*arguments*' and *'understanding'* in general. Using Steinhoff's methodology of thought experiments, the thesis will argue that unrealistic thought experiments are quite useful because they create an improved way of understanding complex political issues without the use of experience. The third goal of this chapter (5.5) is to respond to the existing debate on *abstraction,* surrounding the proper use of unrealistic thought experiments and their implications in political theory. I will demonstrate why Thaler's claims on '*abstraction'* as discussed in his criticism should not give us any reason to worry about the reliability of unrealistic hypotheticals in political philosophy. This will be followed by insight into how to evoke a trustworthy intuition via *abstraction.* The last goal of this chapter (5.6) will examine the method of '*reflective equilibrium'* in political thought experiments. This method will be used by this thesis to argue and defend the logic of comparison that involves the use of unrealistic thought experiments in political theory. The method of reflective equilibrium will be used to argue that unrealistic thought experiments are capable of providing moral intuitions that could offer action guidance in the real world, particularly how we can solve political issues that can be possible for us, here and now.

The remaining sections (5.7) will be used to support some of my claims demonstrating other shortcomings of Thaler's arguments including a brief exposition of four conditions that might lead to the success and failures of thought experiments in political theory. The conclusion (5.8) will cover some objections and possible responses that might arise from our claims and, finally, a summary of the major findings in the chapter.

5.3 The Role of Hypotheticals in Thought-Experimentation: Focussing on Rawls' *Original Position*

Before tackling the issues raised by Thaler, I need to explain in detail the various characteristics and features of unrealistic hypotheticals in political theory. Therefore, this section proposes a primer for explaining how hypotheticals should be assessed, to demonstrate the challenges one might face within their usage in PTEs. It will also offer a framework for defending my view on the legitimate role of unrealistic thought experiments in political theory. I aim to show that some unrealistic thought experiments can be used reliably: using Rawls's *Original Position* as a case study, I will provide a philosophical explanation of their usefulness in political theory.

There are several reasons to believe that unrealistic hypotheticals should not be rejected in political theory; one of these reasons has been clearly stated by Miscevic (2018). He has claimed that hypothetical scenarios can be useful in political philosophy "as a source of political theorizing" (Miscevic, 2018, 159). But, the question is, do we need to engage our thoughts with realistic hypotheticals alone in order to evoke trustworthy intuitions in political thought experiments? Philosophically speaking, thought experiments, in general, do not need to be realistic in order to evoke a reliable intuition. Some are unrealistic, yet; they are capable of evoking trustworthy intuitions. For example, as reported by Skaf (2017), the *Demon Case* used by James Maxwell cannot be proven practically by experience, yet its intuition helps in the development of thermodynamics (see Skaf, 2017, 19). Similarly, political thought experiments do not also need to be physically plausible before trustworthy intuitions could be evoked based on the following reasons:

(i) It can be problematic to follow all the empirical details arising from a political thought experiment due to empirical interference. This could mislead the audience from intuiting correctly from thought experiments, in this case, details need to be abstracted away.

(ii) Because PTEs are arguments (as I will show in the next chapter), having real and recurring features as part of its premises in the thought experimental claims can sometimes be irrelevant to the conclusion that a thought experiment might want to offer, which means it can be pointless to examine their actual cases in the real world before their arguments are considered valid.

The listed claims in (i) and (ii) above, explain why thought experiments in political philosophy can be best described as an "experiment in the laboratory of the mind" whose hypotheticals describe "what if" questions mostly in a counterfactual reality as against their actual world or as against an explanation of the real and recurring situations of the physical world. The idea is that during thought experiments, imagination gains prominence over our experience of the real world (See Brown, 1991; Cooper, 2005; Rescher, 2005), which can also be ascribed as arguments in a narrative or imaginative form (See Norton, 2002). The question is, what roles do hypotheticals play in a political theory via thought experimentation? In responding to this question, Rawls' thought experiment below will help us to demonstrate what hypotheticals could do in political theory.

John Rawls, in his *Theory of Justice* (1971) employs a thought experiment to think about how a just society might be organised. Rawls asks us to imagine what people would agree in the 'original position', parallel to the state of nature in social contract theory. In real life we cannot separate people from their circumstances and interests; it seems that people who are given the power to create a fair society, are not fair and equitable in sharing the resources but instead share them between themselves or to their advantage. Rawls believes that people should imagine being in what he describes as *the Veil of Ignorance*. By virtue of *the Veil of Ignorance*, Rawls claims that everyone will know nothing about their true circumstances in the new world. They will not know their gender, religion, race, income, orientation, disabilities, personal values, and other distinguishing factors. It implies that with the 'veil', we become blinded to all facts about ourselves, and this makes it difficult for us to tailor principles to our advantage. He argues that since everyone is similarly situated, lacking such details, even self-interested people must choose principles of justice that are fair. Therefore, the result of this fair sharing is what Rawls equates as a just society. Rawls expects that the agreed principles will provide basic human rights and equal opportunity for all and that the "primary social goods" such as wealth, income, civil and political rights, liberties, the social bases of self-respect, etc. must be shared equally unless an unequal distribution of any is for the benefit of all. (Rawls, 1971, 105).

There are some fundamental reasons for defending the use of unrealistic thought experiments that need highlighting as deduced above. Rawls' thought experiment is used in this thesis to argue that such decisions to reject thought experiments in political theory for lack of real and recurring features or because they are physically implausible are unjustified. Rawls' use of unrealistic hypotheticals can be useful in political theory for various reasons. Let us

consider three significant ways to explain how unrealistic hypotheticals would be useful in political theory, as deduced from Rawls' thought experiment.

i. They are arguments that provide stable intuitions

Unrealistic thought experiments in political philosophy are arguments and can be used to examine complex phenomena for an improved *understanding* of political issues. As an argument, Rawls used an unrealistic thought experiment to provide the reader with no empirical interference. These claims have been well argued in chapter four using Rawls's original position as a case study. I have shown how the argument view explains how to evoke a stable and objective intuition of a just society without the need for the real and recurring example of a just way of life. I have done so by reconstructing some moral and political thought experiments to defend this view, in particular, how to gain stable intuitions which can be used as moral data in guiding our political judgments over what we ought to do, how we ought to live and to understand what the government owes to its citizens. With stable intuition, the audience could also be able to establish a stable reflective equilibrium to determine those conditions that make our society unfair and how we can create a fair one. To evoke a stable intuition, the audience must be able to conceptualize the role of logic in understanding Rawls's thought experiment, and without being emotional through this, a stable intuition about fairness can be established. This has been demonstrated in chapters four and six where I have shown how to translate unrealistic hypothetical cases as premises of arguments.

i. Providing the thought experimenter with leverage to assess difficult or complex real-world issues.

Unrealistic hypotheticals are relevant to political theory because they help to tackle political issues that real and recurring examples could not assess. As demonstrated in Rawls's case, the salient point is that unrealistic cases in political theory can be useful because they permit legitimate means for *framings* to solve complex cases that are difficult, if not impossible, to establish using real or recurring cases. Plato, John Rawls, Albert Einstein, Galileo, and James Maxwell utilised this significant advantage in performing their thought experiments in both science and philosophy, respectively, so that the needed intuition would be established. For example, Einstein claimed to be '*running in the speed of light'*, Rawls imagined '*going back to the state of nature' (the original position*), the assumptions made by Galileo concerning '*massless bodies'* or '*no air resistance'*, James Maxwell using '*the demon'* to demonstrate a demon's role in thermodynamics are all scenarios that were once driven by unrealistic hypotheticals some of which might not provide the relevant result. For instance, one

might not consider a demon to evoke the needed intuition from such a thought experiment to be successful (see Skaf, 2017, 23). These framings offer easier means for new ideas, revealing tensions within a theory, generally used to raise arguments in proving or disproving a theory beyond the confinement of experience. However, one reason to construct unrealistic hypotheticals even when their actual cases might be available is that the actual cases might invite lengthened discussions that are irrelevant to the issues which can engross or distract the reader from the issues at hand. For example, consider Rawls' thought experiment asking us to imagine going behind *the 'veil of ignorance….and so on'*. Because such thought experiments depart significantly from what might be possible here and now, if the reader examines it from their real and recurring examples, the empirical details of what veil of ignorance might mean to a different society, a clear situation of such ignorance might invite distractions that are not relevant to the facts the thought experiment might want to prove. To avoid such distractions, most thought experimenters frame scenarios and consider them as an 'imaginary existence' to evoke the needed intuition. It is therefore believed that in such cases, the reader should abstract away their real-world cases to avoid the distractions from their empirical details to arrive at the desired result.

ii. Helping the thought experimenter in making universal claims without empirical boundaries.

Some of the challenges faced by most thought experimenters involve making universal norms. In Rawls's case, a similar issue can be deduced. For example, how do we attain a just society universally? Or how can the issue of rights, fairness, and equality be addressed universally? Solutions to such questions can be possible using unrealistic hypotheticals. Rawls' use of unrealistic hypotheticals was able to address the issue of rights, equality, and fairness that are required for an egalitarian society within our global world. This implies that no matter our culture, tribe, or race, his principle of justice can be legitimately applied. An application of *the Veil of Ignorance* in the *original position case* suggests that without *self-interest,* every society universally can be just as suggested by Rawls. It can be difficult, if not impossible, to make such universal claims, especially concerning justice using real and recurring examples. Because a typical real and recurring situation of justice can be limited to either a particular culture or race, with defined geographical boundaries and with different political interests. It is on this basis that the thesis submits that real and recurring features in hypotheticals can hinder the idea of the universal applicability of Rawls's conception of justice: because it might result that justice may not mean the same thing to every society. The varying conception of justice,

therefore, creates a serious impediment to the implementation of such claims as a universal norm. For example, in the real world, just marriage may not mean the same thing around the world. For example, while same-sex marriage is just for most European countries, it is considered an unjust marriage, particularly in some African countries (see Songunro, 2014, for issues on same-sex marriages in Nigeria). But using unrealistic cases, the reader does not need to have a real situation of same-sex marriage to evaluate what a just marriage might look like. The reader only needs intuitive evidence to universalize it as a norm. Let's consider the following as deduced from Rawls's case; suppose we are all under *the veil of ignorance* where we might lack knowledge of our sexual orientation in the new world. Given that we might likely become homosexuals, heterosexuals, or bisexuals. As decision-makers, because we might not know what our sexual orientation might be when we come out from the 'veil', we might be forced to make laws to favour all sexual orientations that will also define the permissible kind of marriage. The intuition from this little illustration shows that since *the veil* makes us neutral, there is bound to be impartiality when making just laws for persons of all sexual orientations. However, the idea is that we can evoke universal moral norms when we go beyond empirical boundaries because unrealistic claims help us to situate our claims and moral judgement within the best wider background of possibilities.

Finally, there are important facts to deduce from the above section for our argument. It is unjustified to reject Rawls's thought experiment because of the lack of real and recurring features, as argued in Thaler's claims. Thus, I propose that it is quite legitimate to infer different realities during thought experiments, as demonstrated above. However, unrealistic cases are capable of guiding our intuitions in making rational choices; they help the reader to assess universal moral and political norms and also enable them to evaluate complex cases that are beyond the confinement of experience. In the section below, I will proceed to defend their usefulness in political theory within the context of *understanding*.

5.4 How Unrealistic Thought Experiments Increase Understanding: Focussing on Steinhoff's Methodology of hypotheticals in Thought-Experimentation

Thaler thinks that thought experiments in political theory are untrustworthy because they are unrealistic and therefore irrelevant. This section will argue that they are indeed useful for philosophical elucidation. I will explain their usefulness by exploring Steinhoff's methodological approach to thought experiments. Steinhoff's view will offer a framework for defending how unrealistic thought experiments facilitate '*understanding',* especially on

complex political issues. This will enable the thesis to argue that unrealistic thought experiments can be used as a reliable tool for political reasoning without the need for actual situations in the real world.

To justify how unrealistic thought experiments increase understanding, we will, first of all, explain what we mean by '*understanding'* in thought experimentation. By understanding, we mean to have a mental grasp of thought experiments, or the ability to show comprehension and to have insight or good judgment about normative or moral issues during thought experimentation. Briefly, understanding is created during thought experiments by means of explanation (see, Stuart, 2018; Gilbert, 2000). Through explanation, a thought experiment can use counterfactual scenarios to offer various skills that evoke understanding. This would involve different processes such as argumentation, simulation, and so on. These processes are used to guide the reader to gain new insight into complex situations before deciding whether a particular normative claim is correct or incorrect. By argumentation, thought experiments can be seen as disguised arguments that provide an easy process for the validation of normative claims. In this case, the thought-experimental scenarios can be seen as eliciting disguised premises, ruled by a system of logic that provides justification for its conclusion. In this case, " a good thought experiment is a good argument and can fail when the conclusion is not adequately supported by a legitimate logical inference" (See Norton, 1996, 2004). The idea is that, by argumentation, a thought experiment needs to conform to the rules of logic to justify whether a theory is valid or invalid. Logically, by argumentation, the information on the premises and the conclusion must not contain real and recurring features of the physical world for an argument to be considered valid. All that a valid argument might need is the ability to establish a logical entailment between the premises and the conclusion (as discussed in the argument view of PTEs in chapter four above). Also, by simulation, thought experiments represent mental pictures in the reader's mind. First of all, the thought experimenter provides this mental picture through the use of well-established and codified counterfactual claims that are capable of eliciting intuitions. Such intuitions can be simulated by the reader in his/her mind to establish moral data. This data contains the raw facts needed by the reader to create some parallels with the real world in order to enable an understanding. The reader thereafter uses such understanding to form intuitive evidence to act as an instrument for raising further arguments for or against a moral or political principle. If a thought experiment fulfils these conditions, it can be described as being effective, meaningful, and efficient.

Moreover, I contend that the evaluation of the outcome of a hypothetical case can go beyond the real world during thought experimentation. For example, just like a pilot may use a flight simulator to understand how to control flight operations, and practice skills in flight operations, in the real world the pilot might not need to consider what he/she is doing inside the simulator as actual. The simulator is used to train the pilot, in some cases, to test the pilot with different problems that are artificial to explain different possibilities on how they would overcome similar situations in the real world. Similarly, thought experiments can make use of unrealistic claims to stimulate new ideas that become a mirror to potential issues that might be difficult to imagine using real and recurring examples. Why do I defend unrealistic thought experiments within the context of *understanding*? Thaler (2016) has argued that thought experiments in political philosophy must have some *'real and recurring features'* in order to be relevant in political theory. Thaler thinks we should reject those that lack connection with the real world as we know it, as they are not relevant in the field. This thesis disagrees with Thaler's claims and argues that such calls for the rejection of unrealistic thought experiments in political theory must be ignored because they are indeed useful. They offer the reader an improved way of gaining an *understanding* of complex political issues needed for political action. It is based on this *understanding* role that the thesis argues that unrealistic thought experiments have a place in political theory and, therefore, must not be rejected as claimed by Thaler.

The question is: do we need realistic hypotheticals to increase our understanding of political issues via thought experimentation? To respond to the above question, Steinhoff's methodology of hypotheticals becomes apt for further explanation. Uwe Steinhoff (2013), in his methodology of hypotheticals, uses an unrealistic thought experiment to raise an argument against anti-torture absolutism. He argues this via a consideration of *the Innocent Jenny Case* to demonstrate why the absolutists could reconsider their previous beliefs on the principle of torture in political theory. Anti-torture scholars have argued that torturing an aggressor, even in extreme circumstances, can never be justified. If we allow torture, we cannot avoid allowing it in the form of self-defensive rape. Steinhoff claims that the idea of self-defensive rape can be morally permissible. He argues that we can test its permissibility using unrealistic hypotheticals to know what a justified rape might look like in another possible world. He asked us to imagine a naked woman called *Innocent Jenny*, who is on a vaginal medication, with the side effect of killing any man whose penis is exposed to it by sexual intercourse. In this scenario, Steinhoff asked: using her only available opportunity, should *Jenny*, who is also

attacked by a naked serial killer, save her life by raping the serial killer? Steinhoff argues that such *self-defensive rape* can be justified. Because, for him, in reality, such a "defenceless woman can never be denied of her only means to defend herself against a serial killer" (Steinhoff, 2013, 149). According to Steinhoff's claims, even though such a scenario does not occur in reality: it can be possible to evoke an intuition that could be paralleled to another form of self-defence in the real world to which his claims can be legitimately compared. In this way, thought experiments would increase our understanding of how a *self-defensive rape* can be rationally justified, particularly when there is no rational reason for denying such a defenceless woman the only available means she has to rescue herself from a serial killer.

Steinhoff's claims on hypotheticals are used in this thesis to defend and argue that unrealistic thought experiments can be used legitimately to increase our understanding without having a connection to the real world. This can be demonstrated based on the following findings: From Steinhoff's methodology of hypotheticals, it can be observed that the practical form of defensive rape can be morally dubious. However, the reader does not need such empirical details to gain an understanding of *self-defensive rape*. But, such hypotheticals help to increase our knowledge of political issues via thought experiments and might be needless to have their physical instantiation in the real world. Because the idea of rape is wrong in reality, it implies that the actual world given Thaler's claims might be problematic for thought experiments. However, Steinhoff's argument teaches us that unrealistic hypotheticals guide our understanding of political issues even when the scenario does not need to be actual. It, therefore, means that to gain an understanding via political thought experiments involves an '*experiment in thinking'.* Meaning understanding can be attained even when the supposed scenarios could not occur or are not likely to occur in reality. In this case, the actual world of hypotheticals is not essential for understanding to take place during thought experiments. The imagined scenarios are therefore designed to test our rational intuition by describing "what if…?" and not their actual cases in the real world. As *'experiments in thinking,'* it is correct to submit that unrealistic thought experiments operate only within the mind of the reader to improve their psychological, theoretical, and conceptual understandings of various political/moral views. It also follows that unrealistic thought experiments can legitimately be used to facilitate the reader's cognitive ability towards political issues without the need to consider their empirical details. Following these views, the reader must not rely on the actual nature of the empirical details from the imagined scenarios before an understanding of political

views is evoked via thought experimentation. Therefore, Thaler's view goes contrary to the typical criteria for thought experimentation in political theory.

The second argument to draw from Steinhoff's methodology of hypotheticals for our claims is that unrealistic thought experiments would be used to expose the limitations/dangers of all absolutists' claims on political theory. This means that philosophically speaking, all political views can be subject to scrutiny using thought experiments. For example, the use of unrealistic hypotheticals to test the principle of torture demonstrate that the theoretical background of antitorture absolutism can be flawed. Generally, many thought experiments have been useful because they play vital roles to defend or disproving a particular theory (Norton, 2004; Brown, 1991). For this to be possible, such roles must be supported with understanding by the audience to authenticate the correspondent commitments. Moreover, Steinhoff's view teaches us that unrealistic thought experiments can help the reader positively by increasing their understanding of the dangers of holding to one belief, which might result in dogmatism over political or moral issues. The question is, how will the audience benefit from such understanding? For example, Thaler questioned the veracity of Steinhoff's claim, as he states: "*what should the audience do with such intuitions tested in a situation that shares close to nothing with the real world, such as Innocent Jenny?"* (Thaler, 2016, 1138). By means of response, unrealistic thought experiments are capable of playing an educating role between the reader and the theory. However, going through such a scenario makes the normative claims more readily graspable. This can also help the reader to test the reliability of claims using unfamiliar situations continuously. With such understanding, the reader becomes aware that all forms of absolutism in political theory can be opened to scrutiny via thought experimentation. Therefore, unrealistic thought experiments are used to evoke understanding by providing an in-depth explanation of certain counterfactual scenarios to convince the reader to reject a political or moral claim.

In the above section, I claim that unrealistic thought experiments can be relevant because they enable understanding by means of explanations. Using Steinhoff's view, I argue that such an explanation does not require any consideration of the real and recurring examples from the imagined scenarios. Steinhoff's view is also used to argue that unrealistic thought experiments help to inspire the audience to overcome all epistemological challenges resulting from absolutism and dogmatism in political theory. However, Thaler's claims fail to acknowledge such legitimacy and therefore stand to limit the potential of thought experiments in political theory. In the next section, I will demonstrate how 'abstraction' can be used to

provide sound intuitive judgment, particularly from thought experiments whose hypotheticals are unrealistic.

5.5 Abstraction in Political Thought Experiments: A Case Study of McMahan's Methodology of Hypotheticals.

This section will respond to the existing debate on '*abstraction',* surrounding the proper use of unrealistic hypotheticals and their implications in political theory. The aim is to demonstrate how Thaler's claims on '*abstraction'* could be flawed. I will be arguing for the moral permissibility of abstraction in political thought experiments.

The question is, why do we defend unrealistic thought experiments within the context of abstraction? First of all, by *abstraction,* we mean the ability to focus on ideas alone as against their empirical details during thought experimentation. Shoemaker (2015) reports that Philosophers love thought experiments because they enable them to abstract or "isolate some of our real-world experiences and evoke intuitions about it, and these revealed verdicts enable us to adjust relevant theories in light of what we find" (Shoemaker, 2015, 1). This implies that abstraction can sometimes be needed for the successful use of unrealistic thought experiments in political theory. To date, critics have become sceptical about their usage which has created controversies in PTEs. Critics of political thought experiments have claimed that abstraction can lead to dishonest conclusions. Thaler (2016) in particular, has argued that abstracting away some real-life details from thought experiments can be morally dubious because, for moral consideration, facts to do with human vulnerabilities and suffering must not be abstracted away. He argues that the abstracted details are morally significant and shouldn't be abstracted away. He also thinks that imaginary cases are meant to guide human actions; therefore, *abstraction* provides untrustworthy conclusions that are insignificant to the action-guiding aim of political theory. Thaler's claim provokes the following questions: Can '*abstraction'* be morally permissible? How can we justify what to abstract away from thought experiments? In responding to these questions, I will interrogate my views via a consideration of McMahan's methodology of thought experiments to argue and defend the legitimate role of *abstraction* in political thought experiments.

McMahan (2013), in his methodology of thought experiments, has engaged his views on the issue of abstraction: and he is in support of their moral permissibility. This, for him, enables the audience to focus on those hypothetical cases we might want to use in testing our moral intuition, in order to gain rational knowledge. In presenting his views, McMahan uses

the Ticking Bomb Case to argue that, the use of torture as a means of self-defence can be justified. In this case, we are asked to imagine a terrorist who has been caught by security agencies with strong evidence that the terrorist has information on how to locate a 'ticking-time bomb' that will explode and kill maybe millions of people in a congested city in less than maybe two hours. The question is, should the terrorist be tortured in order to get information on the location of the bomb to save the life of millions of their citizens? Or should the security agencies allow millions to die? The purpose of *the ticking bomb* case by McMahan's argument is to test the morality of torture against the absolutist's thesis that torture should, in all forms, be rejected. Note that McMahan does not believe in the institutionalization or legalization of torture in practice, because he thinks that the legalization of torture might lead to abuse.

Conversely, Albie Sachs (2013) has expressed his concern about McMahan's view by arguing that torture, whether in the actual world or via hypothetical examples such as the ticking bomb, or any form of hypothetical examples should be rejected. He claimed that the idea of using "concocting imaginary situations" to defend the act of torture can be "horrific" and must be avoided (see Sachs, 2013). He argues this judging by his personal experience with torture and the various experiences of people where governments have used torture on them. Such a society, according to Sach, might not welcome any debate on the morality of torture because of the humiliating historical experience they must have faced in the past. Based on personal experience, torture for Sachs is tied to pain, hurt, discomfort, agony, affliction, irritation, humiliation, dehumanization of victims, and sorrow, and therefore, must be rejected. McMahan's response is that given that the actual situation of torture is wrong, but with the use of imaginary cases as an argument, means torture for the purpose of self-defence can be justified (hypothetically), which invalidates the claims of the anti-torture absolutism. Therefore, one could also count on hypothetical arguments rather than simply relying on personal reactions to deliver the desired conclusion about torture (McMahan, 2013, 3). McMahan thinks that such empirical details about the historical examples of torture can be abstracted away when justifying the moral permissibility of torture. Empirical details such as the effectiveness of torture in reaching the truth and the amount of time that it takes are typical examples of what needs to be abstracted away. He further argues as follows: "what hypothetical examples can do that historical examples seldom can is to filter out irrelevant details that can distract or confuse our intuitions, thereby allowing us to focus on precisely those considerations that we wish to test for moral significance" (McMahan, 2013, 2).

McMahan's methodology of hypotheticals is used in this thesis to argue that abstraction gives morally sound conclusions to thought experiments in political theory. It implies that the reader can go through the same physical movement as a 'virtual event' in mind to evoke intuitions as a possible outcome intended to be used to evaluate our moral judgment. Thaler disapproves when we filter out empirical details via abstraction. Sach rejects the use of hypothetical examples to defend torture, while McMahan's view supports the idea of abstraction as a reliable approach to thought experimentation. The next question that this might provoke is how can we reasonably justify their legitimate usage in political thought experiments. In responding to this question, the thesis will demonstrate using *Rawls' Veil of ignorance* to support the view on abstraction. I contend that Rawls's case gives us a genuine reason why abstraction would be justified in political thought experiments.

Why use Rawls' case to defend the concept of 'abstraction' in the discipline? Abstraction in Rawls' case provides a sound conclusion to unrealistic thought experiments given the following four major reasons: (i) because some empirical details are always irrelevant to Rawls' conclusion. (ii) Abstracted details in Rawls's case are morally insignificant because they do not increase or constitute any form of moral vulnerability and suffering of the moral agents as claimed by Thaler. (iii) Abstraction can be morally permissible because it provides the reader with a trustworthy intuition needed for understanding the rational conception of justice in our society. (iv) Finally, abstraction can be reasonably justified because it makes the reader gain access to rational intuitions when making rational choices on justice.

The question is, what details do we have to abstract away, and why? As suggested in Rawls' Case, human beings are naturally biased in one way or the other, which could make us emotional in our actions. These biases/ emotional feelings distract our understanding of justice. Rawls thinks that it is important we imagine seeing ourselves in *the original position* of complete ignorance where we could see reality from an objective point of view through the *veil of ignorance*. The implication is that, for Rawls, having a complete physical setup of *the veil of ignorance* is not relevant to the conclusion, as such provides a justifiable reason why abstraction is important in Rawls's thought experiment. What this means is that any details amounting to the physical possibility of such an imagined scenario can be distracting and, therefore, must be abstracted away in order to gain the right intuition for a better understanding of fairness, according to Rawls.

Critics of political thought experiments like Thaler might argue that, in the actual world, there is no such situation of becoming ignorant of one's gender status, religion, race, income, orientation, disabilities, personal values, etc. as required by Rawls's thought experiment. This thesis argues that such empirical details are irrelevant to Rawls's conclusion. They might distract the reader negatively by posting a wrong intuition and, therefore, must also be abstracted away. However, in order to understand how a nation would be fair or just, Rawls has also noted that it is important that decision-makers stay neutral when making moral decisions that are fair. Critics might say that we can never be neutral based on their personal experience of the world: the reader must abstract away such details from the thought-experimental scenario to enable better judgments. Also, because some political thought experiments are arguments means we can do away with their empirical details while assessing the nature of their validity (see chapter four above for more details on the argument view of political thought experiments).

Furthermore, can abstraction be morally permissible as used in Rawls's case? I argue that it can be morally permissible given that through abstraction, we gain a rational knowledge of justice that would be needed for political development. However, abstraction in Rawls's case enables readers to understand the principle of justice that suits all societies without causing harm to moral agents. What we mean here is that abstraction in Rawls's case can be reasonably justified because it does not dehumanize human beings or contain facts that might increase human vulnerability and suffering as presented in both Thaler's (2017) and Sachs's (2013) arguments. For example, in *the veil of ignorance case*, we are to imagine a situation in which we do not know our real place in the new society, our natural strengths and weaknesses, religion, gender, race, etc. such empirical details are facts to do with our self-interest which should be considered as morally insignificant and should give us no serious concern during abstraction. This is because they are not dehumanizing facts nor will they increase our vulnerability and suffering. I argue that abstraction in such cases only provides a sound conclusion to thought experiments without any supposed harmful connection or victimization of human victims, whether in the hypothetical sense or the real world. Instead, they are facts when abstracted would form the basis of an improved understanding of the social condition of human beings in the real world through thought experimentation. This is possible because it provides intuitive support for fairness where the reader could understand how everyone would be treated equally and fairly, especially, where the social well-being of all components of the

society can be guaranteed as suggested by Rawls. Given this role, one can legitimately state that abstraction in such cases can be morally justified.

However, McMahan's view teaches us that empirical details might make us arrive at the wrong intuition about thought experiments, showing why abstraction should be considered. Similarly, abstraction in Rawls's case also teaches us that it is important that we do not consider the real and recurring features of the imagined scenario because of the distractions that come with empirical details. I argue that for the reader to avoid any compromise over moral standards on the issue of justice, all empirical details resulting from the imagined scenarios must be abstracted away. They could easily influence our 'self-interest' values and, by extension, can distort our ability to gain the desired conclusion. As seen in Rawls's case, abstraction helps the reader to access a rational knowledge of justice, and equality and also makes us objective in our normative evaluations without any form of empirical distractions. Given this understanding, the thesis submits that thought experiments in political theory do not always target the empirical details of imagined scenarios, the target is to offer us the moral/political intuitions to evaluate normative claims, principles, and theories. For example, as argued by both McMahan 2013 and Brun (2018), the ticking bomb is not meant to show the empirical or practical applicability of torture in the actual world. But, to show that "torture is permissible in some circumstances just in case we refute the claim that torture is impermissible under all conditions whatsoever" (see Brun, 2018, 199).

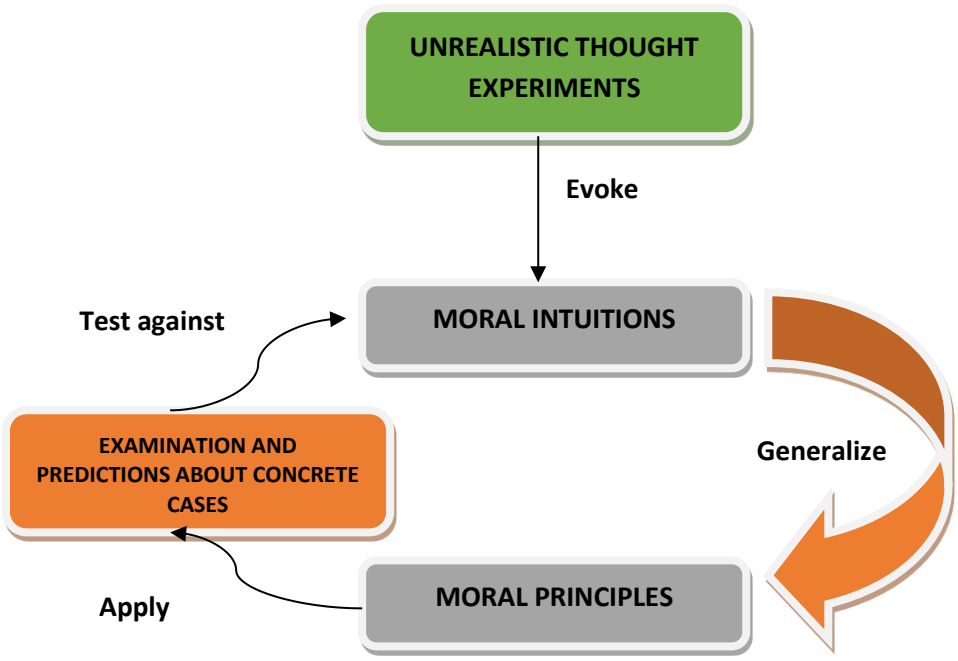
To summarise the major argument of this section: Thaler claims that abstracted details are morally significant and shouldn't be abstracted away. McMahan thinks empirical details can be distracting, meaning without abstraction, we might reach undesired results. Given this dispute, using Rawls's thought experiment as an example, the thesis argues that (i) some forms of abstraction can be morally permissible because it makes us less emotional while addressing moral/political issues. (ii) these abstracted details are not facts to do with human vulnerability and suffering, and therefore (iii) abstraction enables the reader to become neutral and impartial over normative issues, which gives rise to a fair evaluation of justice. Given all these claims, the thesis submits that 'abstraction' can be morally permissible in political thought experiments because some abstracted details are morally insignificant. Moreover, abstraction provides support for trustworthy intuitions. Thus, as deduced from Rawls's case, for the sake of a just society: it is morally sound to adopt the abstraction strategy in a political thought experiment, especially when such empirical details or information are irrelevant to the conclusion that a thought experiment might want to offer.

5.6 Reflective Equilibrium: How Political Thought Experiments can help as Action Guidance

Another way of defending the use of unrealistic thought experiments in political theory is by examining their methods of investigation. This section will examine the method of *reflective equilibrium,* which explains the relationship between unrealistic hypotheticals and the real world (see Rawls, 1971). This method will be used to show how unrealistic thought experiments can be justified, particularly how such justification can be used to enable an understanding of the real world. These findings allow us to argue that unrealistic thought experiments would help to explain facts as they occur in the real world for the purpose of providing action guidance.

Consider the question: What is the purpose of reflective equilibrium? Reflective equilibrium is a method of philosophical analysis used for the justification of particular normative claims using our moral beliefs, principles, laws, and theories during thought experimentation. It involves the logic of deductive/inductive inference: used to enable understanding, stability, objectivity, and justification of normative claims by providing 'intuitive parallels' during thought experimentation. The method is used to provide consistency to our normative judgments by providing that enablement, whereby our moral intuitions cohere with the already-existing laws, principles, or theories, to aid understanding and justification. This method says that during thought-experimentation, we should endeavour to generalize our moral intuitions with a general moral principle to avoid inconsistencies (see Fig. 1.0 below). For example, if a thought experimenter creates a scenario that concludes that killing is good, then he/she must be able to establish such claims as parallels with a general principle that permits the act of killing in the real world. The intuitions from such thought experiments must be able to be reconciled with the respondent's commitments to theories, through which we can then decide which ones need to be changed or given up. It involves "a process of bringing judgments about particular intuitions and general principles of inference into accord with one another" (Nelson, 1985, 71). In other words, to validate such scenario, it is the case that elicited intuitions should cohere with existing theories, in order to produce a concrete result during thought experimentation. Such coherence acts as a source of motivation and serves as a justification for political/moral claims. However, the process of "evaluating the validity of individual cases and general theories in relation to each other" is what I will henceforth describe as an *equilibrium test* (see Stitch, 1988, 416). The diagram below depicts the process of the equilibrium test from a typical example of unrealistic thought experiments.

Fig. 1: A diagram showing the method of Reflective Equilibrium.



As shown in the diagram above, the method of reflective equilibrium can take place in four important steps:

1. The unrealistic case provides imagined scenarios and questions as drawn by the thought experimenter.
2. The thought experimenter goes through the imagined scenario to elicite moral intuitions.
3. The moral intuitions are then compared as parallel with already acceptable moral norms/principles.
4. The thought experiment is therefore morally justified if and only if its moral intuitions have a legitimate parallel with the moral principle to explain how much of each has to be given up or changed to accommodate the other.

However, one might also want to ponder, how is reflective equilibrium important in the defence of unrealistic thought experiments in political theory. In political thought experiments, understanding and justification of normative claims in real-world situations could hardly be possible if the imagined scenario shares no normative inference or similarity with what we already know as the right moral action. In other words, in the absence of deductive/inductive inference from the equilibrium test, an unrealistic thought experiment could be described as being “speculative”, “artificial” or vague and often considered as having nothing to do with

reality (see Thaler, 2016, 1125-9 for this argument). To avoid such an impasse, the method of reflective equilibrium (or equilibrium tests) is significant for the validation of unrealistic thought experiments in the real world. Based on this method, it can be stated that unrealistic thought experiments are justified if our moral intuitions about imagined scenarios cohere with our moral beliefs about the right actions in the real world. By this view, according to Brun (2018), "our judgments and principles are justified if judgments, principles, and background theory are in equilibrium" (see Brun, 2018, 202). This method remains significant in moral philosophy as Scanlon (2002) also states, "it seems to me that this method, properly understood, is, in fact, the best way of making up one's mind about moral matters and about many other subjects. Indeed, it is the only defensible method: apparent alternatives to it are illusory" (Scanlon, 2002, 149). The idea is that before an unrealistic thought experiment could enable understanding or could be considered as evoking justifiable normative claims in political theory, it must pass its equilibrium test. This involves having adequate coherence with our beliefs or moral principles, or else they could be described as misleading, which makes the thought experiments irrelevant. Thus, in such a case, Thaler's claims might be justified if an unrealistic thought experiment is incapable of producing an adequate equilibrium test with real-world situations.

However, unrealistic thought experiments can be defended by using this method to examine the veracity of moral intuitions and their relationship with the real world. It can be argued here that thought experiments in political theory can be justified because of their successful equilibrium test, which I will demonstrate as we proceed. This means they can legitimately evoke normative judgments that are tied to either a deductive or inductive inference, which contains facts about the real world. To make my point a bit clearer, we need to ponder on the question, how does reflective equilibrium facilitate our understanding of political issues in the real world via unrealistic thought experiments?

Various defenders of thought experiments have used this method to show how it can aid our understanding of political issues in the real world. In the Steinhoff argument against antitorture absolutism, the method of reflective equilibrium was established. For example, using the *innocent Jenny case (a case involving torturing an aggressor in the form of rape in order to circumvent imminent death).* In this case, Steinhoff was able to provide an intuitive parallel of the imagined scenario with the concept of self-defence, which happens to be a generally accepted moral norm in the real world. However, the case demonstrates that self-defence covers anyone who rapes an aggressor if that is the only option in saving one's life.

Given that, self-defence in the real world is morally permissible; therefore, Jenny's choice of "self-defensive rape" can be morally permissible. Because the scenario shares the same intuitive parallel with cases of self-defence in the real world as we know it (see Steinhoff, 2013, 146). In establishing such parallels, it means the intuitions from the imagined scenario can be legitimately compared to a certain degree with the principle of self-defence

Similarly, as discussed in the McMahan methodology of hypotheticals in this chapter, McMahan also claimed that torture can also be morally permissible. In *the Ticking Bomb Case*, McMahan argues that since we are allowed to kill for self-defence or in the defence of another whose life is at risk, then torturing a victim to serve the life of one thousand people constitutes an act of self-defence, and is justified. However, at a conceptual level, it is the case that torturing one person is a lesser evil than thousands of people dying, so torture might sometimes be justified. The method of reflective equilibrium helps to provide concrete evidence for both *the Ticking Bomb* and the *Innocent Jenny cases,* which means both cases have a strong inductive link in the real world. The equilibrium test on both cases does not prove that we should change our existing policies/ moral norms on rape or torture, but we need to accept that sometimes it might be right to flout those policies. The implication is that this view helps us to shape our rational knowledge about holding absolutists' claims in political theory.

There are some unrealistic thought experiments that I consider as special cases because of their unique functions. Some cases can provide direct-action guidance in the physical world. They do not have real and recurring features but can evoke a stable normative inference that can lead to practical actions. I classified them as being destructive and deductive: they are destructive because of their capacity to disprove a particular normative claim in the real world. They are deductive because they can also be used to prove real-world claims. Judith Thomson's case provides us with an example of thought experiments of this category, which I consider as one of the most striking and efficient thought experiments in moral and political philosophy. Thompson's thought experiment on *the Violinist Case* shows that the right to life does not always trump all other rights, and given further argument it can be used to disprove the absolutist's argument on abortion, and also used to deduce why abortion should be morally justified in the real world. The idea is that we all consider the foetus is always innocent, they deserve not to be killed, but through abortion, they can be killed. Anti-abortion absolutists believe that since the foetus is innocent, therefore abortion should be banned by all means. The question is: should a democratic government permit abortion? Or should a mother kill a fetus in her womb? Thomson uses unrealistic scenarios to evoke a moral intuition that shows flaws

in the popular anti-abortion argument. Through it, she was able to establish a deductive argument in support of abortion in our society. The *Violinist Case* involves waking up in the morning to discover that you have been kidnapped by the Society of Music Lovers. With the help of a doctor, you have been connected to a famous Violinist with the hope that you will be unplugged after nine months. You were the only person with the same biological match that can save the innocent Violinist from imminent death. Remember that, since the violinist is innocent, it is also morally wrong to kill an innocent person. The question is: will you choose to stay connected in order to save the innocent Violinist or unplug yourself and kill the innocent Violinist? As part of the argument, Thompson has used the Violinist case to block the simple appeal to the right to life as an end to the abortion debate. In buttressing her point, The violinist case was used by Thompson to establish a reflective equilibrium between killing the violinist and killing a foetus as it occurs in the real world in a typical case of abortion. The idea is that if we can unplug ourselves from the Violinist, which is a legitimate thing to do when kidnapped (according to Thompson), then we can also kill a foetus in the womb. However, Thompson uses this logic of comparison to create a deductive argument that supports the view of why abortion should be morally permissible. Remember that the foetus is innocent, while the Violinist connection is unjust. Thompson argues that abortion should be allowed because "the foetus's right should not trump over the pregnant mother's right to self-determination," which parallels the view why the Violinist's right should not overweight the right of his captive (Thompson, 1971, 48). Despite the various criticism of *the Violinist case*, it can be argued here that *the Violinist Case* demonstrates how unrealistic thought experiments can evoke moral intuitions that can be used as reliable data for action guidance in the real world.

However, as deduced from *the Violinist Case* above, I do not intend to argue that the right to life should be outweighed by other rights. But it should be noted that governments should make informed decisions to support the moral permissibility of abortion in the real world as against the view of anti-abortion absolutists, which prevents any act of abortion: as prevalent in most countries of the world that are still having issue in legalizing abortion. Thompson's logic of comparison teaches us that, when a woman is pregnant as a result of rape, then she has all the right to determine the fate of the foetus.

There are essential facts drawn from this section that are vital to our argument against Thaler's criticism of unrealistic thought experiments in political theory. (i) The method of reflective equilibrium, if applied in a cogent way, can improve the relationship between unrealistic thought experiments and the real world. (ii) With the method, it can also be argued

that unrealistic thought experiments do not explain away facts from the real world as we know it. (iii) Given that in some cases, there is always a justifiable parallel between unrealistic thought experiments and the physical world, it shows that they can have the tendency of providing the role of action guidance to real and recurring issues in our society as deduced from Judith Thompson's *Violinist case*. (iv) It can also be argued that the method of reflective equilibrium provides consistency to unrealistic thought experiments, which makes it possible to use unrealistic thought experiments in evoking a stable normative judgment, which can be considered as objective and unspeculative. Given these findings, the thesis submits that unrealistic thought experiments have a legitimate place in political theory, and those who argue that they are not useful are mistaken.

5.7 How to Assess Unrealistic Thought Experiments in Political Theory

My response to Thaler's scepticism on thought experiments in political theory provokes the following question: What would a successful or a failed thought experiment in political theory look like? This section will outline and defend four conditions under which thought experiments in political philosophy could fail or succeed. One of the reasons why thought experiments in political theory can succeed is when we consider them as an *experiment in thinking*. Following this view, the thought experimenters do not see thought experiments as 'real and recurring' situations of the physical world but only as a means of rational reflection, to test intuitions in an environment free of interference. Some hypotheticals that are set up with real and recurring features of the real world are prone to mislead the audience's intuitive judgments due to distractions from their empirical details. For example, in *the ticking bomb case*, as suggested in this chapter, the audience can easily get distracted by irrelevant empirical details of torture like the physical humiliation, psychological damages sustained by it, the fear of torturing the wrong victim, and other physical consequences one might incur as a result of torture. These factors might impede the audience by misleading them to gain a negative intuitive judgment from the thought experiment. However, Beck & de Wijze (2015) have defended the *ticking time bomb scenario* and have suggested that this thought experiment is "supposed to work in philosophical inquiry… as part of an argumentative device to uncover intuition in our search for moral truth" (Beck & de Wijze, 2). I adopt part of their views and will advance in the next chapter using Norton's argument account to demonstrate how political thought experiments can be regarded as arguments. I also suggest that, because of the epistemic importance of unrealistic hypotheticals, thought experiments like the *ticking time bomb scenario* should be encouraged in political theory because it allows the reader to abstract away

irrelevant empirical details that may distort or confuse the audience from revealing their own moral commitments. As argued in this thesis, it is worth abstracting any empirical details that might make the reader become emotional over moral issues because they make the reader prone to biases over moral decisions in a given political scenario, as suggested by Rawls. Thaler's suggestions will make such roles more difficult, if not impossible.

Another factor to consider when developing thought experiments in political theory is to ensure that hypothetical scenarios are capable of persuading the reader to believe what the thought experiment wants to prove. A political thought experiment will be persuasive if it guarantees what Popper describes as "inter-subjectively testing," meaning it must evoke intuitions that can be replicated by many readers (see Popper, 1968, 46). However, intuition forms the foundation of normative claims in a political thought experiment, which makes it an important factor. Any political thought experiment that cannot generate moral/political intuition is bound to fail. To evoke intuition means, hypothetical narratives immediately strike the reader as obvious moral truth or falsity without the need for empirical evidence. It is via intuitions that cogent moral data can be deduced from thought experiments used in critiquing, disproving or proving theories or claims. Yet, intuition is the foundation of political thought experiments as perception is for science; any thought experiment that lacks this capacity to evoke intuition in the reader's mind should be considered ineffective. I suggest that when creating political thought experiments, we should rather consider whether hypotheticals could evoke the needed rational intuition for normative judgment; because without it, a political thought experiment becomes redundant.

The third reason why thought experiments might fail or succeed is if and only if the claims to be proven by thought experiments depend solely on the support of another theory. This will only be the case if the thought experiment was drawn in such a way that it must be true if only there is a particular law that agrees to part of its claims. Scientists do expect that the conclusion of a thought experiment must conform to existing laws in order to create a trustworthy intuition. For example, the '*clock-in-the-box'* thought experiment by Einstein failed because it was not supported by the laws of relativity, which Einstein had thought of as part of his claims. The *clock-in-the-box* needed to be successful if Einstein was able to confirm his claims using relativity theory (See Wilkes, 2003, 12-15 for this argument). In political thought experiments, our moral/political intuitions must conform to our moral beliefs, principles, or theory, in particular, when we have reached reflective equilibrium. As demonstrated in this thesis, the imagined scenario must evoke moral intuitions, which would

be compared to a generally accepted moral norm in a process described as an equilibrium test. The idea is that a political thought experiment is justified if and only if the thought experimenter can establish a legitimate parallel between moral intuition with other acceptable moral principles in the real world. A political thought experiment is bound to fail if such moral intuition does not share any link with the moral belief or principles to which they were both compared. For example, if a thought experimenter develops a moral intuition in support of suicide, and then lacks a moral principle that supports suicide in the real world. Such a thought experiment is bound to fail because there may be no legitimate theory to gain support from by means of reflective equilibrium. Similarly, thought experiments can fail if their equilibrium test misleads the audience into accepting moral principles that encourage actions that are considered harmful in the real world.

Finally, a political thought experiment can also fail if the hypothetical cases are obscure in meaning. That means, imaginatively opaque hypotheticals can lead to unsuccessful thought experiments. A thought experiment, with opaque imagination, poses a 'what if' question that the thought experimenter will not be able to respond which loses the value of the thought experiment (see Cooper, 2005, 328; Wilkes, 1988, 15; Brownlee et al. 2017, 21). The reason is that he/she may not be able to set them up with meaningful hypotheticals that would lead to a useful or valid conclusion. The problem could be that the thought experiments will fail to create the needed intuitive support that is vital for a successful thought experiment. For example, if one is to imagine an animal kingdom as a system of government similar to human beings, one should note that animals are not smart like human beings. For instance, there is no democratic process in the animal kingdom, etc. If a thought experimenter assumes a world like the animal kingdom where the reader cannot know how such a world looks, it becomes a problem for the thought experiments to evoke the needed intuitive knowledge. Imaginary cases that are framed from understandable concepts create clarity and make intuitions easy to evoke than those that are framed from 'wacky' ones (See Brownlee et al 2017, 40). Any form of opaque imagination, whereby the reader does not understand the semantics of the hypothetical cases can make thought experiments to be useless in political theory. Thus, using an opaque imagination in thought experiments also leads to an opaque or wacky thought experiment. Nevertheless, for political thought experiments to be successful, I suggest that such hypotheticals whether unethical, far-fetched or physically impossible must all be well understood by the reader in order to gain the 'imaginative grip' that would enable the reader to attain the needed intuition: this will help the reader to grasp the moral data that the thought experiments want to generate. With the moral data, it will now become possible for the reader

to compare his/her intuitive judgment with other familiar cases in reality in order to make thought experiments relevant to the real world.

5.8 Objections and Possible Response

There are major ways this chapter can be criticized. The first is by those who claim that political theory is purely a practical discipline. Therefore, such thinkers might criticize my view because they might believe that thought experiments can only be useful if they involve practical reasoning alone in order to direct our practical actions in the real world. Thaler might raise similar arguments against my view. Because he might argue for a concrete context within which imagination should engage to avoid being artificial. There are flaws to such criticisms. The question is, do political theories focus only on practical guidance? My quick response to it is that political theory is not limited to solving practical issues alone, there are several cases where their theoretical relevance can be useful in politics (see, Chilton, 2004, 1-4). For example, consider a political theory that claims that the despotic role is illegitimate, such claims do not imply that we might need to go through the real occurrences of despotic rulership in order to depict whether it should be allowed in our society. A political theorist uses his or her theory to draw inferences from historical analysis or some form of utopia in raising some useful deductive claims about why political authority must respect popular will. This process supports the use of counterfactual scenarios to provide useful explanations for political thinking without the need for the actual example of despotism. This process involves unrealistic hypotheticals that are important for political theorizing and action guidance in political philosophy.

However, the scope of this chapter is not to engage the reader in the debate between the practical or the theoretical relevance of political theory. But one might be poised to suggest that the practical dimension of political theory does not explain everything one might need to learn. Therefore, its theoretical dimensions are worth considering especially via thought experimentation, as seen in both Steinhoff and McMahan's methodologies of thought experiment. In defending unrealistic thought experiments within this context, the chapter has shown that even though it might be practically impossible for unrealistic thought experiments to establish a direct link with the real world as Thaler might argue. As an experiment in thinking, we can always draw their real-world example via the method of reflective equilibrium. This suggests that unrealistic thought experiments do not deal with artificial cases;

they can evoke legitimate intuition, which can be compared to the real world for practical guidance.

This chapter can also be criticised by those who think that Rawls` *theory of justice,* in particular, his idea on the *Original Position* is too abstract from the real world; meaning his political theory should be considered unrealistic and cannot be useful in the real world. For example, Robert Paul Wolff (1977) is one of the earliest critics of Rawls's thought experiment: he has claimed that Rawls's political theory "says little or nothing about the concrete fact of social, economic, and political reality" (Wolff, 1977, 195). Wolff agrees with the fact that *the Theory of Justice* is quite impactful, but argued against *the veil of ignorance* in the following statement: "the devices of the bargaining game and *the veil of ignorance*, while preserving the political, psychological, and moral presuppositions of such a doctrine, raise the discussion to so high a level of abstraction that the empirical specificities needed to lend any plausibility to it are drained away" (Wolff, 1977, 195). Wolff's concern is that Rawls's claims in his thought experiment are too unrealistic. By implication, even though we accept the scenario as an argument, Wolff thinks Rawls thought the experiment is not useful. For him, unrealistic claims can be problematic as they cannot be used in responding to the actual problems in the physical world. By way of providing a quick response to his claims and all other similar claims concerning the use of unrealistic thought experiments in political philosophy, I contend that Rawls thought experiment is a disguised argument (see section 4.5). Therefore, it is the case that an argument may contain an unrealistic premise, yet the argument could still be valid. For example, let's consider the following argument whose premises are unrealistic but whose argument is valid in terms of form.

1. All Americans love Coke
2. All People are American
3. Therefore, all people love Coke

The above argument is valid, but the premises are unrealistic because in the real world, it is impossible for all Americans to love drinking coke as stipulated in line 1. Besides, in line 2, all people on earth cannot also be Americans. The idea is that, as an argument, moral and political thought experiments can make use of unrealistic or untrue premises in providing a valid argument, which means the premises of an argument need not be realistic to evoke a valid argument.

Similarly, David Lyons (1975) has also offered another possible criticism against the use of unrealistic political thought experiments. For instance, he has criticised the nature of Rawls's arguments by raising the question, what bearing could a Rawlsian social contract argument, an imaginary agreement, have on us? Using the '*coherence argument',* Lyon has questioned the efficacy of Rawls's argument for lacking moral force. He claims that a moral principle that is useful is supposed to cohere with reality as we know it in order to offer a profitable solution and action guidance to political issues. In other words, according to Lyons, the judgment that we make are to be useful or reasonable if they "express our basic moral conviction" (Lyons, 1975, 145). Such that for them to become moral data, they must have a reflection on an alternative account of such judgments. The point that Lyons is making is that Rawls's claims "do not rely upon contingent, empirical considerations, but turns entirely on logical possibilities" (Lyons, 1975, 148). Lyon contends that it is unclear whether the coherence argument favours Rawls's claims. Thus, Rawls's argument is logically possible, but "the coherence argument remains unsubstantiated" because it pays no attention to practical possibilities with a concrete moral case or specific social arrangement (Lyons, 1975, 148). He also thinks that "the more the Rawlsian social contract argument is subordinated to coherence, the less justificatory force it seems to have" (Lyons, 1975, 150). His major claim is that "we will not or cannot find the principles of justice in the natural world; we cannot discover them, for there is nothing to be discovered" (Lyons, 1975, 157).

Lyons's criticism has questioned the legitimacy of Rawls's logical claims, which implies that Rawls's claims are not realistic because it does not completely satisfy the coherence argument via thought experimentation. It can be argued that, from a logical standpoint, the argumentative force of Rawls's *Original Position* is more potent than the coherence theory because it relies more on logical conviction than mere moral conviction. The idea is that intuitive evidence resulting from valid arguments would be considered more reliable compared to the justifications derived from a standard coherence theory. However, given the validity of Rawls's argument, as shown in section 4.5.1 of this thesis, it can be argued that a stable intuition can be evoked from it, which cannot be guaranteed from a coherent argument. However, coherent arguments can be subjected to scrutiny because they are not self-evident, which means they are prone to be affected by different societal views. As shown in section 4.5.1, I have reconstructed *the original position* as an argument, as we proceed in this thesis (in chapters eight and nine), I will demonstrate how Rawls's argument on the Original Position can be possible for us here and now.

Lastly, this chapter can be criticized by those who might claim that the method of reflective equilibrium does not guarantee the stability of moral intuitions from unrealistic thought experiments. Given this concern, they might argue that the method does not lead us to a reliable normative judgment during thought experiments. There are different ways to overcome such concerns as explained in this chapter. We have addressed the fact that, in order to gain a stable intuition via the equilibrium test, the thought experimenter must be able to explore the option of establishing a parallel from realistic principles that could constitute a general moral norm. This means we are obliged to only compare imaginary cases with those moral principles that are already considered as generally true, for example, the principle of self-defence, etc. The benefit will be that using such acceptable moral norms makes a thought experiment evoke a stable and consistent intuition. In a situation whereby we compare the imagined scenario with our individual or personal principles, it may lead to some form of logical and methodological relativism making normative judgments unstable.

4.9 Conclusion

In conclusion, my objections to Thaler's claims are not mistaken because unrealistic hypothetical cases are relevant in creating a rational understanding that helps to generate genuine knowledge about our political situations. Indeed, they can be considered as arguments used also to penetrate complex issues that are practically implausible, particularly complex cases with physical impediments. These are needed to evoke rational intuitions for moral significance. Thus, they render a straightforward approach where we would test our moral beliefs in an area that is not familiar to our experience of the world. For example, just like a trainee in a tailoring school might only be using 'condemned clothing' to test his/her imaginative skills, to develop new and unfamiliar styles in the fashion industry. Similarly, a thought experimenter also possesses such skills by framing physically implausible hypotheticals (mostly unfamiliar to the audience) to test his/her imagination in an environment that is free from interference. Through this process, it becomes easier for thought experimenters to develop thought experiments that can correct societal issues or problems that might challenge our political institutions. This means that rejecting them can lead to serious epistemic negligence. This chapter also argues that to acquire the needed intuitive support from unrealistic hypotheticals, abstraction is essential, especially when the abstracted empirical details are irrelevant to the thought experiments or when the abstracted details are not facts that concern human vulnerability and suffering. As demonstrated using John Rawls's thought experiment, the readers must abstract away empirical details because the abstracted facts do

not dehumanize the moral agent both in the hypothetical world and the real world. It has also been argued that via abstraction, the desired intuition can be reached without distractions. The chapter has also argued that unrealistic thought experiments can create an action-guiding role if and only if their moral intuitions would be used to establish a parallel with existing moral principles/theories via the equilibrium test. Thus, given all these epistemic roles, the chapter submits that unrealistic hypotheticals are useful in political theory and should not be jettisoned as proposed by Thaler. A political thought experimenter is, therefore, free to imagine any scenario he or she finds appropriate because imagination or any form of idealization via thought experimentation should be considered a "limitless resource" (See Rescher, 2005, 147). The implication is that there is no restriction on what kind of hypothetical scenario one should create. As such, it can be realistic or unrealistic, but one must take into consideration the meaningfulness and the intelligibility of its conclusion with the real world. Therefore, an unrealistic thought experiment is not misleading if it can create meaningful and intelligible intuitive conclusions with a viable equilibrium test in the real world. They can also be unsuccessful if such scenarios become meaningless or fail to produce a justifiable reflective equilibrium. The thesis is not against the use of realistic thought experiments entirely but for the fact that realistic cases are prone to trigger biases that can interfere with our judgments on normative issues. The salient point is that while political thought experiments deal with experimenting with *thoughts*, this must be distinguished from other forms of physical experiments in the real world. However, because of the realistic complexities and diversities of physical scenarios, the reader can be distracted from important moral/political facts. Thus, empirical distractions can sometimes limit the potential of thought experiments in political theory. Finally, as argued in this chapter, it should be noted that political thought experiments are arguments and do not involve experimenting with real and recurring situations; they are experiments in thinking. Therefore, those who argue that the use of unrealistic hypotheticals should be replaced with realistic cases are mistaken.

CHAPTER SIX

6. THE 'ARGUMENT VIEW' AND THE PROBLEMS OF INTUITION IN MORAL AND POLITICAL THOUGHT EXPERIMENTS

6.1 Introduction

In this chapter, I begin to apply the *'argument view'* of political thought experiments (PTEs) to respond to some of the criticisms raised by Klampfer and others in chapter three (section 3.3) against the usefulness of intuitions evoked by means of thought experiments in moral and political philosophy. Briefly, as discussed in chapter three, Klampfer, in particular, has contended that the extension of thought experiments to political philosophy should be considered problematic because of the controversial nature of their intuitions. He argues that PTEs evoke different intuitions from different readers. He is worried that since they yield varying results, it means their intuitions will run the danger of being systematically misleading. He is also sceptical of whether their usage in responding to normative claims can be legitimate. For these reasons, Klampfer and some critics think that thought experiments in moral and political philosophy cannot be justified because of the unstable nature of their intuitive claims. Their concern is that unstable intuitions cannot be used as *prima facie* evidence for normative judgments. They use these claims to argue that, the extension of thought experiments to moral and political philosophy should be considered unfounded.

This chapter argues that Klampfer's claims against thought experiments in political philosophy are mistaken and perhaps present a clear-cut misrepresentation of thought experiments in political philosophy. In defending this view, drawing from my claims some that s PTEs are arguments (as demonstrated in the preceding chapter) that have not been acknowledged by these critics, it is reasonable to state that those claims are misleading. Thus, if one understands these connections, then one can better understand the reliability of thought experiments in political philosophy. The preceding chapter maintains that 'argument' is at the very heart of thought experimentation in political philosophy and that all good PTEs are arguments. As valid arguments, they are ruled by a system of logical inference that contains premises capable of evoking objective intuitions in their conclusions. I contend in this chapter that as arguments, some moral and political thought experiments are not controversial as critics might claim. They can yield a logical conclusion which means they can also evoke legitimate intuitions that can be considered as being stable, objective, and reliable and would be used for responding to complex political issues. My major aim is to examine how *the argument view*

can provide a secured ground for the reliability status of intuitions elicited via moral and political thought experiments. To show the prospect of *the argument view* in moral and political thought experiments, I will briefly discuss my claims in section various sections below:

6.2 Responding to the Problem of Indeterminacy / Instability of intuition in moral and political thought experiments.

As already discussed in section 3.3, Klampfer (2017) has argued that intuitions elicited by moral and political thought experiments are unstable because it often leads to an indeterminate conclusion. He is worried that people cannot easily come to an objective conclusion through moral and political thought experiments. The implication is that Klampfer thinks that intuitions elicited by moral and political thought experiments are untrustworthy, which can mislead the audience on the right moral action to follow. Kampfer contends that unstable intuitions cannot be used as a legitimate means of responding to normative claims (see Klampfer, 2017, 338). Klampfer argues this using *the Violinist Case* to justify his claims. He thinks that *the violinist case* provides a divisive conclusion. He is worried that from the responses he got from different surveys, many people are against the act of unplugging oneself from the Violinist while some also consider staying connected to the Violinist. This divisive conclusion is what Klampfer uses to claim that moral and political thought experiments are prone to unstable intuition (Klampfer, 2017, 338). He claims that since the Violinist case evokes different conclusions from different people, it means that intuitions elicited via moral and political thought experiments are controversial and cannot be used as reliable evidence for a normative judgment.

In responding to Klampfer's criticisms, I claim that Klampfer's claims are mistaken. As shown in the chapter above, there is a robust interconnection between political thought experiments and arguments that have not been acknowledged by Klampfer. If one understands these connections, then one can better understand the reliability of *the violinist's case*. The thesis claim is that *the violinist case* is a product of argumentation and must be assessed within the purview of an argument. Meanwhile, Klampfer misconstrues the nature of thought experiments in political philosophy by choosing to rely more on the empirical assessment from different responders, which I think can take us away from the original context of a thought experiment. As argued in chapter five, political thought experiments must be understood within the context of argumentation. Therefore to have a cogent understanding of the nature of their intuitions, we must, first of all, see them as 'arguments' through which we can evaluate their

important premises that will lead us to their logical conclusion and stable intuition afterwards. As demonstrated in the chapter above, the internal properties of PTEs are made up of disguised arguments. Therefore, in assessing their internal validity, the cogent way of assessing them must be by using the argument view. With the '*the argument view'*, it can become easier to determine whether a thought experiment has a valid argument such that its major conclusion must have a logical entailment from its premises. Readers of PTEs that fail to see them as arguments are prone to become subjective or bring in introspective claims outside what is being stated in the scenario, which might give rise to different intuitions. To avoid such controversies in PTEs, *the argument view* provides insight into how objective intuition during thought experimentation could be assessed. I do claim that *the Violinist Case* can lead to an objective intuition if we can evaluate them using *the argument view*. To proceed, I will reconstruct it into a conventional argument in order to argue and defend the view that moral and political thought experiments can evoke a stable conclusion against Klampfer's claims.

First of all, it is important to note that in the *Violinist's Case*, we do not have to think about the situation in the actual world or imagine our reaction in bed with *the Violinist* in order to evoke a stable intuitive judgment. Instead, we have to see the thought experiment as an argument that wants us to determine that abortion is permissible in some circumstances (see Thomson, 1971). The *violinist case* is a deductive argument that is used to serve a destructive purpose. What I mean here is that there is a clear logical inference in the violinist case, where a deductive conclusion can be established to disprove the anti-abortion absolutist claims. This process can only be possible because of the argumentative nature of the scenario, which involves some disguised premises with a clear system of logic. I contend that Thomson's claims can be considered legitimate because she presented a case that has a valid argument, and perhaps poses a counter-claim to the anti-abortion absolutism. I will proceed to show how *the Violinist's Case* can be reconstructed as a valid argument from where we can see how an uncontroversial result can be deduced from it. Let me quickly summarise the thought experiment below before I proceed:

You wake up in the morning and find yourself back and back in bed with an unconscious violinist. The famous violinist has been found to have a fatal kidney ailment, and the Society of Music Lovers has canvassed all the available medical records and found that you alone have the right blood type to help. They have therefore kidnapped you, and last night the violinist's circulatory system was plugged into yours so that your kidneys can be used to extract poisons from his blood as well as your own. The director of the hospital now tells you, look

we are sorry the Society of Music Lovers did this to you-we would never have permitted it if we had known. But still, they did it, and the violinist now is plugged into you. To unplug you would be to kill him. But never mind, it's only nine months. By then, he will have recovered from the ailment and can be unplugged from you. The question is "do you have an obligation to stay connected? (see Thomson, 1971, 48-9). Thomson argues that there is no such obligation to stay connected for nine months in order to save the violinist. Thus, to stay connected is a matter of choice, and disconnecting from rescuing the *Violinist* can as well be justified.

Using Norton's reconstruction thesis, I believe that the above claims consist of tacit premises that can be reconstructed into a conventional argument. The Violinist henceforth will be called Patrick, and the reconstructed argument is as follows:

1) Patrick was kidnapped and taken to the hospital by the Society of Music Lovers (**Premise**).
2) Patrick was connected to the *Violinist* by the Doctor in order to rescue the Violinist's life after nine months (**Premise).**
3) From 1 and 2, if Patrick was kidnapped and connected to the violinist by the Doctor then Patrick will be connected to the violinist for nine months without his consent.
4) **Hypothesis:** Everyone has a right to determine how his/her body is used**.**
5) **Conclusion 1**: From 4 and 3, if everyone has a right to determine how his/her body is used and Patrick will be connected to the violinist without his consent for nine months then Patrick's right to determine how his body is used has been infringed.
6) From 5 and 2, if Patrick's right to determine how his body is used has been infringed then it is not permissible for the Doctor to use any part of Patrick's body to save the violinist without consent.
7) From 6 and 4, if it is not permissible for the Doctor to use any part of Patrick's body to save the violinist without consent and Patrick has a right to determine how his body is used then Patrick's choice of staying connected or disconnecting himself from rescuing the *Violinist* can be justified.
8) From 7, it is not permissible for the Doctor to use any part of Patrick's body to save the violinist without consent
9) **Conclusion 2:** From 8, Patrick can disconnect himself from the violinist.

Given the above illustration, it can be argued that *the Violinist's case* can elicit intuition that should be considered stable and perhaps uncontroversial. This is premised on the fact that moral and political thought experiments with a valid argument should be considered reliable and non-divisive. Because the conclusion logically follows from the premises, it means the above argument is valid because it satisfies the general condition of validity. In the above argument, it is imperative to note that the conclusion is true because the Society of Music Lovers does not have the necessary consent as required by the general moral principle, which should justify their action in connecting Patrick to the Violinist. Because of this violation, Thomson claims that for Patrick to disconnect himself becomes a legitimate moral action and is logically supported by the general moral principle. However, the translated argument above satisfied what Judith Thompson would like to agree as a sound argument in defence of abortion. Thomson uses the above scenario to form an intuitive parallel with real-world issues concerning any woman who becomes pregnant due to rape. Thomson thinks that any woman who becomes pregnant due to rape has the right to terminate an unwanted fetus because, given the act of rape, the foetus can lose protection from the mother. As demonstrated above, it is the case that Thomson's Violinist case is a valid argument, but what of the case of abortion? Is there a valid logic of comparison between the Violinist case and the abortion case?

It is also important to note that, Thomsom's case was also crutial to the debate about abortion, which I believe can also be looked at in this thesis to justify whether the thought experiment can also yield a valid argument on the issue of abortion. To explain the validity of Thomson's argument on the case of abortion, it is imperative that we also translate important aspects of Thomson's claims into a conventional argument. I believe that the reader should remain patient in order to follow the logic of my translations. (Note that, in my translation, I will use the name Rose to designate a fourteen-year-old schoolgirl, that was raped and became pregnant). As deduced from Thomson's claims, she wants to argue against the claim that the foetus has a right to life. So it is the case that a foetus should be killed, and abortion should be permissible (Thomson, 1971, 48). She believes that the foetus's life can be taken away during an abortion, especially in the situation whereby the foetus was conceived as a result of rape. As such, decisions on what happens to a mother's body should outweigh the right to life of the foetus.

Thomson thinks that the case of being kidnapped and plugged to a Violinist which means one didn't volunteer to be plugged parallels intuitively with a scenario whereby a woman is pregnant due to rape and is assumed to be free after nine months after delivering the child. In

defending the case of abortion, Thomson assumed that any woman who is raped by an unknown intruder and got pregnant in the process can decide not to keep the foetus for nine months and can decide to abort the foetus. Thomson argues that if it is the case that the Violinist's life was limited because one didn't volunteer, then it is also the case that the right of the foetus can be cut short in the case of abortion as a result of pregnancy due to rape, thus, affecting the foetus's chance to live (See Thomson, 1971, 49). These claims can be translated into the following tacit premise and plausible assumption as follows:

1 Rose was raped by an unknown intruder **(Premise).**

2. Rose got pregnant due to the rape and will need to allow the foetus to use her body for nine months if it is to live **(Premise).**

Critics of abortion might contend that the foetus is an independent being just like the mother and so deserves not to be killed. As such, abortion should be completely eradicated because if a foetus is an innocent person, killing the foetus makes the killer a murderer. Thomson thinks that such claims can be falsified. The implication is that one will be plugged for nine months with the Violinist, which means that one may probably remain with the violinist for nine months. She noted that the foetus' right to life or not to be killed must be on just ground. Situations of pregnancy due to rape create an unjust relationship between the mother and the foetus. Therefore, the mother does no injustice to the foetus if she chooses not to maintain the foetus's right to life. Given this situation, the woman can, therefore, not guarantee the foetus her body for shelter and food or the right to life for nine months since "the unborn person's existence is due to rape" or what can also be described as unconsented sexual intercourse (Thomson, 1971, 58).

3. From 1 and 2, if Rose got pregnant due to rape then it follows that Rose will be keeping the foetus in her womb for nine months as a result of the unconsented sexual intercourse.

However, Thomson believes that women, in general, deserve better treatment, especially pregnant ones. Meanwhile, she thinks that just like a victim has the right to decide whether to unplug himself from the innocent Violinist, a woman that is pregnant due to rape deserves the right to self-determination. One of these rights includes women's rights over their bodies, which guarantees the decisions on when and how to get pregnant, and of course, staying pregnant for nine-month is also a decision that they have to make. Thomson tries to project

the idea that women should gain control over what goes on in their bodies. She argues this while stating that "the mother has a right to decide what shall happen in and to her body" (Thomson, 1971, 48, 55). I consider this claim as a hypothesis for her claims as it concerns the right of all mothers during and after pregnancy as deduced from Thomson's argument. This can be translated as follows:

4. Hypothesis: Everyone has a right to determine how their body is used**.**

To support her argument, Thomson has a clear description of what should form the definition of "right" for her argument. She explains this using the Violinist's case as follows "the fact for continued life that violinist needs the continued use of your kidneys does not establish that he has a right to be given the continued use of your kidneys. He certainly has no right against you that you should give him continued use of your kidneys. For nobody has any right to use your kidney unless you give him such a right" (Thomson, 1971, 55). The implication is that no one has the right to ask you to stay connected for nine months, though staying connected should be kindness on your part and not something that you own someone. Therefore, it is not our duty to decide or enforce when a mother should get pregnant or allow a foetus to stay in a woman's womb for nine months. As such, if rape victims are forced to keep the foetus, means their right to determine how their body is used has been infringed. The plausible assumption for the above claims can be translated as follows:

5. **Conclusion 1**: From 4 and 3, if everyone has a right to determine how their body is used and Rose will be keeping the foetus for nine months as a result of unconsented sexual intercourse then Rose's right to determine how her body is used has been infringed.

The above illustration shows that women who own themselves deserve to be respected, especially on decisions that involve their health during pregnancy. Thomson also thinks that it may be hard for a third party to enforce abortion for a pregnant mother. Thus, refusing the choice of a woman means we are trying to take away their right. Thomson claims that a pregnant woman can abort the foetus because no matter how innocent the foetus might be, the woman does not have to accommodate the foetus for nine months. She deserves to make decisions that must save her life because her body is not on loan to her, which may take away such rights (see Thomson, 1971, 54). She argues that, if a woman houses a foetus, it should be that she also decides how she houses it in her best interest. Thomson further claims that "the foetus's right to life does not trump the pregnant woman's right to control her own body and its

life-support functions" (Thomson, 1971, 56). As such, if a Woman's right to determine how her body is used has been infringed then it is not permissible for anyone to force her to allow the foetus to use her womb for nine months without her consent. This can be translated as follows:

6. From 5 and 2, if Rose's right to determine how her body is used has been infringed then it is not permissible for anyone to force her to allow the foetus to use her womb for nine months without her consent.

Given the Violinist case, Thomson argues that it will not be right for you to sit passively by and wait for your death (Thomson, 1971, 52). One could say, "this body is my body", and so you have reasons to unplug yourself from the innocent Violinist. Thomson supports this claim and argues as follows: "if you start to unplug yourself having learned that you will otherwise have to spend nine months in bed with him and perhaps die at the end, there is nobody in the world who must try to prevent you" (Thomson, 1971, 55). This implies that your right has been bridged, similarly, in the abortion case, Thomson argues that the right of the woman who is pregnant due to unconsented sexual intercourse (rape) has been violated, most especially, when the foetus poses a great health risk to her. Therefore, such a decision to abort a foetus must be genuinely supported. She argues this with the following conclusion, "a sick and desperately frightened fourteen-year-old schoolgirl, pregnant due to rape, may, of course, choose abortion, and any law which rules this out is an insane law" (Thomson, 1971, 65). Given the above claims, it can be plausibly translated as follows:

7. From 6 and 4, if it is not permissible for anyone to force her to allow the foetus to use her womb for nine months without her consent and Rose has a right to determine how her body is used then Rose's choice of keeping or aborting the foetus can be justified.
8. From 7, it is not permissible for anyone to force Rose to allow the foetus to use her womb for nine months without her consent
9. **Conclusion 2:** From 8, Rose can abort the foetus from her womb.

6.2.1 Summary of Thomson's argument on abortion:

1. Rose was raped by an unknown intruder **(Premise).**
2. Rose got pregnant due to the rape and will need to allow the foetus to use her body for nine months if it is to live **(Premise).**

3. From 1 and 2, if Rose got pregnant due to rape by an unknown intruder then it follows that Rose will be keeping the foetus in her womb for nine months as a result of the unconsented sexual intercourse.
4. **Hypothesis:** Everyone has a right to determine how their body is used**.**
5. **Conclusion 1**: From 4 and 3, if everyone has a right to determine how their body is used and Rose will be keeping the foetus for nine months as a result of unconsented sexual intercourse then Rose's right to determine how her body is used has been infringed.
6. From 5 and 2, if Rose's right to determine how her body is used has been infringed then it is not permissible for anyone to force her to allow the foetus to use her womb for nine months without her consent.
7. From 6 and 4, if it is not permissible for anyone to force her to allow the foetus to use her womb for nine months without her consent and Rose has a right to determine how her body is used then Rose's choice of keeping or aborting the foetus can be justified.
8. From 7, it is not permissible for anyone to force Rose to allow the foetus to use her womb for nine months without her consent
9. **Conclusion 2:** From 8, Rose can abort the foetus from her womb.

It can be argued that the above argument is valid and sound and, in fact, evokes a stable intuition on the issue of abortion. It is sound because all the premises are true, the conclusion is true, and the argument is also valid. The conclusion in line 9 is true because one can say we have the right to decide how our body is used. In the actual world, if someone threatens you with death, the victim has a right to defend his/herself. In the same way, a woman who is pregnant due to rape can be justified in aborting the foetus to save her life due to health complications, particularly when such pregnancy makes her become traumatized, psychologically deformed or could lead to the death of the mother. Thomson argues this by stating that "a woman surely can defend her life against the threat to it posed by the unborn child, even if doing so involves its death" (Thomson, 1971, 53).

The important fact to learn from Thomson's case is not to kill all unborn children, but that extreme views against abortion should be jettisoned. As deduced from Thomson's claims, the fact that some women might choose not to abort the fetus does not mean we should conclude that she has no right to do so. Therefore, the woman does the foetus no injustice if she decides on an abortion (Thomson, 1991, 61). However, another important fact to deduce from

Thompson's claims on abortion is that some moral and political thought experiments contain valid and sound arguments. Therefore the validity of such arguments shows that political thought experiments can also evoke stable logical reasoning. In this case, it is evident that a stable and reliable intuition on the right of women on the issue of abortion can be evoked by means of thought experimentation in moral and political philosophy. Finally, we need to understand both reconstructions as part of the thought experiment, while the violinist scenario explains the thought experiment in its hypothetical illustration, the abortion scenario explains the intuitive equivalence of the real-world case. Also, reconstructing the *Violinist case* as an argument offers a way to Klampfer's concerns about the people having differing intuitions regarding this thought experiment. Therefore, it can be argued that because the argument is sound and valid then anyone who disagrees with the conclusion is just wrong.

6.3 Responding to the Problem of Generalization and Extrapolation in Moral and Political Thought Experiments.

Klampfer has also argued that intuition from the *'Pond and the Envelope'* case is problematic because it is unclear how the conclusion of the thought experiment should be extrapolated. His concern is that moral and political thought experiments can evoke different generalizations that are not related to the issues at stake. For example, in *the Pond and the Envelope Case* by Peter Singer, you are asked to imagine that you are a professor, dressed up in new clothes and new shoes. You have a lecture in ten minutes, on your way to the lecture you are walking by a lonely pond where you spot a drowning child. Without thinking, you jump inside the pond and save the child because the child's life is more important than your new clothes or having a successful lecture. In another case, imagine when you return home to receive an envelope from the United Nations International Children's Emergency Fund (UNICEF) - a charity organization asking you to donate a small amount of money that you can easily afford in order to save starving children in Africa. It is clear that you have to save a drowning child, but do you have to donate to UNICEF? (See Singer, 1972, 230-232). Singer uses both cases to establish some kind of parallel claims in order to present his argument that promotes charity for the purpose of saving lives no matter the distance.

Klampfer argues that *the Pond Case* is problematic because he thinks that sometimes the audience finds it very difficult to understand what exactly to intuit from this case. Possibly, he thinks that three extrapolated generalizations are possible (i) that "I ought to save the child drowning in front of me, (ii) that, in general, everyone in a position to do so ought to save

children from drowning (iii) that one ought to prevent something bad from happening as long as she can do so without sacrificing anything of comparable value" (Klampfer, 2017, 339). Klampfer thinks that the fact that there are possibly three different generalizations in the same scenario means they can mislead the audience over the right moral action to follow.

This thesis argues that Klampfer misconstrued the purpose of thought experiments in moral and political thought experiments. I argue that thought experiments in moral and political philosophy should not be rejected because the *argument view* can be used as a reliable means to address the problem of extrapolation as raised by Klampfer. Meanwhile, it can be argued that Klamper also misconceived Peter Singer's thought experiment. Singer's claims should otherwise be understood within the context of an argument. Peter Singer uses *the Pond and the Envelope case* to evoke an argument in support of "charity" which I believe involves argumentation. Meanwhile, because of the nature of the argument, some authors, like Brownlee et al. (2016) agree that Singer's case is a good thought experiment (Brownlee, 2016, 14). The thesis also agrees with this point and will proceed to argue that *the Pond and the Envelope Case,* is a valid argument that involves a single generalization in its conclusion as against Klampfer's claims. Using the *argument view* of John Norton, let us consider the reconstruction of *Singer's Pond and the Envelope Case* to show the validity of Singer's argument. In my reconstructed version, Patrick is seen as one who wants to save the child from drowning.

1. If Patrick was passing by on a lonely path and sighted a child who was about to be drowned in the pond, and who he could easily rescue, then he should rescue the child **(Premise)**.
2. If we accept any principle of impartiality and universalizability and equality, then we cannot discriminate or refused to rescue someone merely because they are far away from us or we are far away from them **(Premise**).
3. From 1 and 2, if Patrick has a duty to rescue a drowning child in the pond and Patrick accepts any principle of impartiality and universalizability and equality, then Patrick has a duty to rescue other children far away who he can easily rescue because the distance is morally irrelevant.
4. From 3, If Patrick receives an envelope from Unicef requesting a charity donation to prevent children from death far away then Patrick can easily rescue children from like malaria, starvation due to war and provision of shelter because the distance is morally irrelevant (Premise)

5. From 1, 2, and 4, if Patrick can easily save children from preventable death, like malaria, and starvation due to war and provision of shelter by donating money to charity then he should give money to charity through UNICEF.
6. From 5, Patrick can easily save children from preventable death like malaria, starvation due to war and provision of shelter by donating money to charity (**Premise**).
7. **Conclusion:** From 4 and 5, Patrick has a duty to UNICEF.

Singer's thought experiment above satisfies the condition of a valid argument. It can be considered a valid argument because the conclusion in line 7 is entailed by the premises. To understand the validity of the above-reconstructed argument, there are important aspects that need to be explained. The first aspect is to understand *the Pond and the Envelope Case* as one thought experiment emanating from one argument, meaning we cannot ascertain the validity of the argument only from the perspective of *the pond* scenario without looking at the other scenario that describes *the envelope*; second is what I have described as the '*moral principle'*; thus, line 3 is reconstructed to capture this as a moral obligation that Singer thinks we all have to save the life of fellow humans who are in immediate danger, especially when without sacrificing anything of moral significance. He thinks that because of the feelings of empathy and compassion that we have for any child at risk we are obliged to save them without anyone forcing us to. Singer believes that all lives have equal value, meaning we need to treat everyone equally without being partial (Singer, 1972). Based on the principles, as outlined in line 3, Singer believes that it does not make any morally relevant difference between rescuing one child in the pond and rescuing other children far away. He thinks that just as our lives and the well-being of the child in the pond may matter to us, then other lives should matter as well. The rest of the argument including Singer's conclusion is a clear invocation of this principle as outlined in lines 4, 5 and 6. This explains that it is reasonable to accept that our duty to one child in the pond (without sacrificing anything of moral significance) has the same moral equivalence with the same duty of donating the same little money to UNICEF, which might be used to reduce global poverty among children in Africa and other parts of the world. According to Singer, if we accept any principle of impartiality and universalizability and equality, then we cannot discriminate or refused to rescue someone merely because they are far away from us or we are far away from them (Singer, 1972). This goes to apply that if we agree to rescue

one person in the pond at a little cost then there is no morally relevant difference in donating to charity organizations abroad at a little cost.

However, in the account of how political thought experiments contain valid arguments, I have primarily followed John Norton's *argument view* and Singer's explanation of *the Pond and the Envelope Case*. This is used as a framework to argue in this thesis that, if we can assess thought experiments within the purview of their arguments, the challenges of it being polluted by extrapolation or evoking different generalizations can be remedied. However, *the argument view* provides a cogent means of discovering what exactly the thought experimenter wants us to intuit within any given hypothetical scenario. Thus, the reconstruction of Singer's case shows that the audience will find it easier to understand what exactly Singer's case was meant to address, making it much easier for us to evoke a stable intuitive judgment if and only if we can follow the systematic nature of the premises down to its logical conclusion. Finally, reconstructing the *Pond and the envelope case* as an argument above offers a solution to Klampfer's concerns about the generalization/extrapolation regarding this thought experiment. Therefore, it can be argued that because the argument is sound and valid then anyone who disagrees with the conclusion is just wrong

6.4 Responding to the Problem of Bias and Prejudice in Moral and Political Thought Experiments

Part of Klampfer's argument against moral and political thought experiments is based on the issue of undetected and uncorrected effects of bias and prejudice. He thinks that as human beings, we are not free from bias and prejudice. He thinks that this sometimes affects our evaluation of social and political explanation of matters that concerns us and others. He uses these claims to argue that this concern might affect our spontaneous social and political judgment during thought experiments.

In response to this claim, I do think that naturally, human beings can be biased in one way or the other. I also believe that the *argument view* as projected in this thesis can protect us from the effect of bias and prejudice, particularly during thought experiments in moral and political philosophy. There are so many advantages that the argument view brings to the reader that will eradicate such biased decisions during thought experimentation. The argument view, first of all, makes it possible for the reader to remain focused and analyse thought experiments within a logical context. Within, this purview, we are supposed to assess a particular moral and political thought experiment within the context of its argumentation. As an argument, we are

also not allowed to go beyond the logic of the already existing premises, in order to reach a logical conclusion. According to *the argument view*, an argument in a thought experiment may contain deductive, inductive or abductive inference already provided by the thought experimental claims, which means the audience must rule out other thoughts that do not conform to the premise of the argument. The idea is that *the argument view* remains a cogent means of assessing moral and political thought experiments in order to be free from bias and prejudice: because it directs the attention of the reader only to the premises and conclusion of the thought experiment. In conclusion, it can be argued that any reader that goes beyond the argument of PTEs is likely going out of the original context of the scenario which must be avoided to eliminate being biased while providing the intuitive judgement from PTEs.

There are some key points to deduce from this section. Klampfer thinks that we should consider the extension of thought experiments to political philosophy as problematic because he thinks that (i) thought experiments in political philosophy could lead us to an indeterminate conclusion (ii) it is often unclear how we should extrapolate from thought experiments (iii) and can be affected by bias and prejudice. In this chapter, I argue that Klampfer misconstrued the purpose of thought experiments in political philosophy. I contend that we do not need to worry about these problems, because moral and political thought experiments are arguments. Therefore, as arguments, they can be translated into conventional premises where a logical conclusion can be reached. I contend that if the audience can follow the rules of logic as contained in the thought experimental arguments, there will be no such concern of having a biased decision during thought experiments because one will only be tied to the rule of logical inference. And of course, the problem of extrapolation and indeterminacy can all be addressed by virtue of the existing thought experimental premises down to their logical conclusion. This is because the *argument view* gives us a cogent means of studying the internal structure of thought experiments to check these excesses.

6.5 When do Intuitions from Moral and Political Thought Experiments Fail?

In endorsing the *argument view* for the assessment of moral and political thought experiments, I have identified a few important reasons to consider before rejecting or accepting thought experiments in political philosophy. The first reason that moral and political thought experiments can fail is when readers draw their conclusion outside the argumentative contexts of the thought experiments. The idea is that to assess moral and political thought experiments; readers are to focus on the relevant aspect of their scenarios. The implication is that any reader

who fails to focus on the argument that a thought experiment presents will end up evoking intuitions that are not relevant or that undermine the subject through which a thought experiment was initially established. Such intuitions will fail to address the "what if questions" which contravene the thought experimenter's plan. The idea is that readers of thought experiments must, first of all, understand the background of the argument through which a particular moral and political thought experiment is formed in order not to go against the argument of the scenario. Readers who fail to recognize thought experiments as arguments are prone to drawing a conclusion that does not fit the premises of the thought-experimental scenarios. Therefore, such readers could invite a negative intuition that also fails to address the issue upon which the thought experiment was meant to test.

In addition, some thought experiments can fail because some critics may consider the events of a thought experiment as if they were actual. For example, if *the veil of ignorance* in the *original position* case was considered as something that should be actual, the event of the thought experiment will be affected. This of course can invalidate certain claims of the scenario, because, in the real world, it is impossible to locate situations in which we forget all of our personal inclinations when taking actions that will at the same time affect us. To reorder the event by modification can lead to a failed intuition if such modification fails to capture the exact argumentative context of the thought experiment. Rawls sees *the veil of ignorance* as something that has no connection with the real world, so readers of the *original position* must maintain the scenario as posited by Rawls to evoke the right intuition. Critics of moral and political thought experiments sometimes confront the thought-experimental claims with varieties of counter-examples to show their limitations. During this process, they modify the scenarios and sometimes take us away from the original argument, by citing other incoherent claims. In buttressing this point, Pyror (2006) also noted that the process of offering counter-examples can sometimes be difficult because it can lead to *incoherent* or contradictory claims during thought experiments. For instance, if a philosopher during thought experimentation asked us to imagine "Every square has four corners…,", with a counter-example we go on to say "Not so! Imagine a round square,". In this case, we have not in fact described a coherent possibility, and so we haven't succeeded in offering a genuine counter-example to his claim. Given such a situation, it can be difficult to ascertain whether we have described a coherent possibility or not. That is of course a big part of what makes counter-example so difficult during thought experiments (Pyror, 2006). As deduced from this thesis, while modifying a thought experiment, readers of hypotheticals in the discipline must maintain the original logical plans

of PTEs. I contend that we must see PTEs as arguments else it might be possible to modify away the real contents of a particular thought experiment.

The next reason moral and political thought experiments can fail is if they possess invalid arguments. It implies that a thought experiment can be argumentative, but when the argument is not valid, such cases are capable of eliciting untrustworthy intuitions, which might not be useful for good political reasoning. A thought experiment with an invalid argument falls short of the required standard for a reliable intuition. Unreliable intuitions are considered irrelevant because their usage as evidence for normative claims might be problematic. I suggest that for moral and political thought experiments to evoke a reliable intuition, there must be a clear picture of a valid argument whereby a conclusion is established by means of logical entailment. For instance, in this thesis, I consider *the violinist case*, the *Pond and the envelope case*, and Rawls's *Original Position* as thought experiments with valid arguments and perhaps good examples of reliable moral and political thought experiments. Because their arguments are valid, it means they can also evoke stable intuitive judgments, which should be considered as useful tools for responding to normative claims. However, if the arguments of these thought experiments were to be invalid, then they are prone to be jettisoned because their intuitions will be considered unstable and unuseful for normative evidence.

For example, I consider *the Trolley Case*, by Philippa Foot (2002) as an example of a moral and political thought experiment as problematic. In this case, we are asked to imagine a runaway trolley, hurtling toward a track that can kill five railroad workers. The driver can decide to continue on the same track or steer to another track which might result in killing one worker. The question is, "is it permissible to turn the trolley?" (Foot, 2002, 23). Foot's response to the *Trolley Problem* was that the morally justified action would be to steer the trolley to kill the one workman, thus saving a net five lives. To justify the morality of her claims using '*the doctrine of double effect'*, she made a distinction between what she called 'negative duties' and 'positive duties'. This doctrine states that "it is always wrong to do a bad act intentionally in order to bring about good consequences, but that it is sometimes permissible to do a good act despite knowing that it will bring about bad consequences" (Stanford Encyclopedia of Philosophy, 2020). Based on this doctrine steering the trolley to kill one is what Foot wants us to agree with because it is a positive action.

My first concern for Foot's argument is that the doctrine of double effect is not a generally acceptable moral principle, meaning no society permits the killing of one innocent person in order to protect five persons. This is why the thought experiment becomes prone to different

forms of speculation, in other words, different intuitive judgments can be evoked depending on the background of the reader. Meanwhile, the Trolley problem has led to a matter of controversy between utilitarian and non-utilitarian audiences, in particular, when it comes to the question of whether doing so produces a greater good for others. An audience with a utilitarian background can claim that it is morally permissible to save five workers than save one innocent worker which supports Foot's claims. Conversely, a non-utilitarian audience might claim that we should allow five to die because aiding the killing of one innocent worker could be wrong. My view in this thesis is that moral and political thought experiments should be able to have a generally accepted moral principle as part of the thought experimental claims for it to be valid. The idea is that the *Trolley Case* can be controversial because the thought experiment lacks a generally accepted moral principle or hypothesis in the scenario that can provide intuitive support for a valid argument, thus, readers are prone to arriving at an unstable intuitive judgment. As such, it can only invite speculative ideas from different readers as a result of having unstable intuition. Therefore, reconstructing this argument without a known universally 'acceptable' moral principle in the scenario that permits what number to kill or save given this situation makes it difficult to establish a valid argument from such moral and political thought experiments. In summary, the Trolley problems have been used as a measure of utilitarianism, but their usefulness for such purposes has been widely criticized and to date, the problem has been the subject of many surveys leading to divisive responses.

6.6 Critical comments

From the foregoing, there are important facts to note about the nature of political thought experiments. As a result of our discussions above PTEs can be considered as uncontroversial if they can elicit a consistent intuition, such that its conclusion can be replicated by another, in a sense Karl Popper described as "inter-subjectively testable" (see Popper, 1968, 46). To achieve this aim is to ensure that political thought experiments produce the same intuitive judgments to all readers no matter the discipline, class, gender, race, or cultural background of the reader during thought argumentation. Although political thought experiments are entirely carried out without any new data from experience, many critics of political thought experiments think that the intuitive outcome from political thought experiments might be prone to mere speculation. One way of addressing this concern is that some complex thought experiments in the discipline are often justified by valid arguments. Although the validity of arguments does not depend on having false or true premises or conclusions, by mere logical entailment. However, it is important to note that there are some

moral and political thought experiments that I consider reliable because they are products of valid arguments. In other words, some moral and political thought experiments can present a stable intuition if and only if their scenarios contain assumptions based on a general principle where valid arguments can be deduced through a well-established logical inference. I have a special interest in those thought experiments because they contain universally acceptable moral truth that provides intuitive support for those sets of thought experiments. These thought experiments in moral and political philosophy are often considered legitimate because their general principles share a logical relationship with the conclusion, which means they can also provide a valid argument. To make this point a bit clearer, my position is that any political thought experiments that can generate a hypothesis or a generally accepted moral principle within a given scenario can be reconstructed into an explicit argument. What this means is that the general principle and its assumptions enable deductive and inductive flow used as inference through which a valid argument can be formed during thought experimentation. Given this understanding, a political thought experimenter does not need new empirical data for their argument but depends solely on the hypothesis or general principle or universally acceptable principle and the premises to evoke a valid argument. This is important if thought experiments must play the role of disproving or proving a normative claim.

There are some possible criticisms in this chapter that critics of my views might raise in this chapter especially when it comes to my claims that thought experiments are arguments disguised and their relationship with their embedded arguments. For instance, in the thesis, Thomson's Violinist thought experiment occurs as part of an argument for the permissibility of abortion under certain circumstances. They might argue that the argument on abortion is only one small part (as pointed out above) making the point that respondents do not themselves believe that someone's right to life always trumps their rights to their own body. The equivalent occurs with Singer's Pond case as well. Their major claim might be that thought experiments will always be embedded in larger arguments, but that is neither a reason for seeing them as being those larger arguments nor themselves being disguised arguments. In responding to these concerns, the thesis adopts the method of reflective equilibrium as captured in section 5.6. This method allows thought experimenters to establish the intuitive equivalences of various thought experiments to connect the imaginary world with real-world issues. However, the thesis acknowledges that these thought experiments in question are arguments disguised, but can also be used to establish their intuitive equivalence using the method of reflective equilibrium. Intuitive equivalence is used in most successful thought experiments to connect the imaginary

case with real-world larger arguments or situations. To remedy these concerns, the thesis has provided the reconstructed versions of both the main thought experiments as arguments in disguise and the reconstruction of the other versions with intuitive equivalence. Both the Singer case and the violinist case are typical examples of thought experiments within this category that allows the use of intuitive equivalence to show their veracity in responding to issues in the real world. It can be argued in this thesis that reconstructing the version considered as intuitive equivalence provides evidence for the fact that most thought experiments in the discipline also have a place in a larger argument that can use to respond to real-world issues. Therefore, considering thought experiments as an argument in disguise does not deny the fact that they can also be used to achieve any purpose within larger argumentative claims.

6.7 Summary and Conclusion

The major facts to deduce from this chapter are that PTEs are arguments and that some good ones can also be reconstructed as valid arguments without epistemic loss. Because they are arguments, it means that they can evoke a stable intuition if their arguments are valid. As shown in the chapter, some moral and political thought experiments are made of valid arguments, which means they are capable of producing reliable intuitions. Findings from this chapter also show that political thought experiments should not be rejected as claimed by critics because of their clarity and ability to evoke a reliable and stable intuition for various moral and political reasoning. However, as suggested in chapters five and six with *the argument view*, PTEs can be objectively assessed. This enables the reader to understand the metrics of their intuitions. While being objective, we must also be able to apply a true sense of logical skills in following their features as it pertains to the thought-experimental claims. However, introspective mindsets while accessing political thought experiments should not be welcome to avoid bias and prejudice that might pollute our intuitions during the philosophical analysis of thought experiments. As such, the way to evaluate whether a PTEs work is to reconstruct it as a logical argument. To maintain a sustainable relationship between the reader and the given PTEs, it is believed that *the argument view* provides the parameters that guide our general understanding and framework upon which a stable intuition can be evoked, including when PTEs go wrong. With the argument view, there is clarity, robustness, and a constructive engagement upon which everyone can be brought to one context of argumentation and perhaps to one definite conclusion if the premises are objectively followed.

In the remaining chapters (seven and eight), I begin to shift my attention to concrete political issues, to examine how PTEs would be used in responding to the problem of political corruption in Nigeria. While chapter eight will explain the relationship between PTEs and narratives, to demonstrate how political intuitions can shape our behaviours via their motivational possibility. Similarly, in chapter nine, I will examine the nature and effects of political corruption in Nigeria as a case study. I will use those findings in responding to the question of whether political thought experiments can also be institutionally possible. Ideas from the institutional possibility of PTEs will be used in the same chapter to address the problem of political corruption in Nigeria, especially as it concerns policy implementation against corrupt behaviour in the country. And finally, chapter nine will bring us to the end of this thesis by providing the summary and conclusion of the entire thesis.

CHAPTER SEVEN

7. THE MOTIVATIONAL POSSIBILITY: CAN THE USE OF IMAGINED SITUATIONS IN POLITICAL THEORY SHAPE THE FEELINGS AND POLITICAL BEHAVIOUR OF ITS AUDIENCE?

7.1 Introduction

In chapter two I have discussed one of the fundamental problems of thought experiments that pose the question: can PTEs possibly provide motivation? This question relates to the concern of whether PTEs can motivate their audience, especially in changing their moral and political behaviour. Nenad Miščević in his desiderata of PTEs suggests that imagined situations in political philosophy must have this possibility as part of the requirements for thought experiments to be considered as successful in political philosophy. In Miščević's '*desiderata of realisticity'*, he pointed to the fact that PTEs should have motivational possibilities so that they can become practically useful to their audiences. He argues that PTEs are supposed to be motivationally possible, for instance, while reading a thought experiment "people must be such that in principle they can be motivated for such a project" (Miščević, 2013a, 519). Miščević's major claims pointed to the fact that in the discipline, hardly any PTEs meet this desideratum (see, Miščević, 2013a, 519-24).

This chapter aims to consider whether PTEs can reasonably be hoped to motivate readers to change their behaviour. Before I proceed, we must admit the difficulty in getting motivated by means of thought experimentation. However, is there something so special about some PTEs that can be hoped to motivate readers to change their actual moral or political behaviour during thought experimentation more than other types of non-thought experimental arguments? This chapter argues as follows, that some PTEs are both arguments and at the same time narratives. As narratives, readers are prone to getting more psychologically involved in the thought experimental claims, such that there is reason to think that some narratives in a thought experiment are capable of shaping our beliefs which might result in changing behaviour during thought experimentation. To substantiate this claim, this chapter evaluates McVey's empirical work to unravel the importance of narrative and its effects on human behaviour. Drawing from Mcvey's (2020) study titled *'The moral importance of narrative: a philosophical analysis of narrative transport',* the work argues that when readers are presented with narratives, they act in some ways that they usually will not which leads to changing of behaviour via its motivational powers. Findings from this study make us believe some PTEs would reasonably motivate readers if they are redesigned to include such distinctive features.

In discussing the primary objectives of this chapter, I shall focus my attention on the following questions: *(i) what is it about narratives that motivate actions?* McVey's study suggests that it is narrative transport. According to the study, narrative transport is "the phenomenon of losing oneself in a narrative… it is also the mechanism responsible for changes in belief, attitude, intention, and affect in individuals" while reading a narrative case or a story (McVey, 2020, 78). Perhaps it is the phenomenon that causes the swaying of the reader's beliefs that motivates action, hence leading to changing behaviour. *(ii) Can political thought experiments be expected to facilitate narrative transport?* And if so how can they be redesigned to better provoke narrative transport as suggested by Mcvey's studies? Responding to these questions shall provide us with an informed decision on what we need to do for the use of imagined situations in political theory to motivate action, in particular, in shaping the feelings and political behaviour of its audience.

To proceed with the above questions, section 7.2, reminds us of one of the fundamental claims of this thesis, which states that political thought experiments (PTEs) are arguments. It will also demonstrate how PTEs can as well be considered as narratives. The idea is that some PTEs are a combination of narratives and arguments as seen in Rawl's *original position* thought experiment. In section 7.3, I examine some empirical findings by Christopher McVey (2020), demonstrating the motivational power of narrative scenarios and philosophical arguments. Using Christopher McVey's (2020) findings, I will demonstrate how some narrative scenarios produce narrative transport; and how the combination of argument and narrative provides motivation. Findings from this study give us sound reason to hope that some political thought experiments if they have the right narrative features in combination with argument might help to motivate behaviour. It can be deduced that for imagined situations in political theory to shape the feelings and political behaviour of its audience, stories from thought experiments might need to have the right narrative features to produce narrative transport that can be expected to motivate action. Section 7.4 responds to the thought-provoking question of whether political thought experiments can facilitate narrative transport. The section also examines how we can redesign political thought experiments to better provoke narrative transport, in particular, one that will meet the various elements set out for narrative transport as described by McVey's 2020 study. I will discuss this in relation to Rawls's thought experiment, first by identifying how Rawls thought experiment meets some of the conditions for narrative transport. This will be followed by what we need to do to meet other expected conditions so that political thought experiments could be capable of changing the political

behaviour of their audience in the real world as we know it. Section 7.5 summarizes the general findings of the chapter and shall provide some recommendations for the future development of the narrative component of PTEs.

7.2 Political Thought Experiments as both Arguments and Narratives: An Insight from Rawls's Original Position.

In chapter four, I showed that political thought experiments can be thought of as arguments. I have also shown how Rawls's *original position* can be used to argue for justice, based on a fair redistribution of social goods. To justify this claim I have as well reconstructed Rawls's thought experiment as a valid argument. The question is can valid philosophical arguments do so much in changing or altering the actual behaviour of their audience? To address this question, below I will begin to examine a valid argument on *smoking* to see if it can alter our behaviour in the real world as we know it.

A valid argument against smoking includes as follows:

1. Smoking increases a person's chances of dying young.
2. If we want to avoid early mortality we should stop smoking
3. Patrick wants to avoid early mortality

Therefore, Patrick should stop smoking.

The premise of the argument above simply confirms the fact that to avoid early mortality we should stop smoking. However, we can assume that Patrick knows that smoking is dangerous and can intellectually accept the argument. Thus, we can logically consider that the above argument is valid and sound, but the question is, in the actual world can we rightly justify that Patrick will act as the argument suggest he should? Or can smokers stop smoking based on the validity of the argument above? However, despite the validity and soundness of the argument, in the real world, we ignore what the argument suggests, and many people are still in the habit of smoking despite the risk. Because the argument may not do so much in altering the behaviour of those involved in smoking, it is reasonable to state that some philosophical arguments might not do so much in changing our behaviour in the real world. The concern is that people may not be motivated by such an argument even with the serious warnings by the World Health Organization (WHO) concerning the dangers and health implications of smoking. In the real world, both young and old are still involved in this habit, and there is no indication that people will stop smoking soon. For instance, a survey by the

Office of National Statistics (ONS) in the UK, found that, in 2019, "14.1% of people aged 18 years and above-smoked cigarettes, which equates to around 6.9 million people in the population, based on the estimate from the Annual Population Survey (APS). The survey also states that 15.9% of men smoked compared with 12.5% of women, while those aged 25 to 34 years had the highest proportion of current smokers which is 19.0%" (see Office of National Statistics, 2021). The above example shows that there are traditional philosophical arguments we accept as sound, that is we accept all premises as true, likewise the conclusion, but where we often do not feel any inclination to act as the arguments suggest we should. Based on this, it can be argued that in general, policies emanating from valid arguments might not alter the reader's behaviour, because we may not act according to the intelligibility of valid arguments. It also sounds reasonable to state that the nature of validity of these arguments can only be justified by mere logical arrangement and not by our actual way of life.

Given these concerns, it is reasonable to doubt that traditional philosophical arguments will alter people's behaviour. In other words, this thesis believes that there is reason to think that in general, presenting people with valid and sound arguments does not do much to alter their actual behaviour. Perhaps, it also raises the question of can political thought experiments (as earlier reconstructed as arguments) be hoped to change the behaviour of their audience in the real world as we know it. In other words, can Rawls's thought experiment (argument) be able to change our behaviour or motivate us concerning the issues of justice? I will be arguing that some political thought experiments are not just arguments, but also at the same time narratives. Perhaps, if redesigned with the right narrative features and valid arguments, they could give us a sound reason to hope that some PTEs are capable of motivating their readers. Because they are sometimes presented as stories, they can also persuade people's rationality, particularly, in shaping our minds both logically and motivating us along with the thought experimental claims and of course might result in behavioural changes.

Meanwhile, some literature supports the view that thought experiments would come in different narrative styles. For instance, Brendel (2018) agrees that "narratives are the means through which many thought experimental situations are conveyed" (see Brendel, 2018, 313). Similarly, Norton (2004) also supports the view that thought experiments are arguments disguised in a pictorial or narrative form (Norton, 2004). These views support the idea that a thought experiment can be expressed in both argumentative and narrative forms. As already argued in chapter four of this thesis that Rawls's thought experiment is an argument, thus, can be legitimately considered as a valid argument. It is also important to note that even though the

thought experiments are arguments as argued in chapters 5 and 6, there are other pertinent questions that might evoke, for instance, do thought experiments also have narrative features that can be demonstrated as stories?

In responding to the above question, it will be interesting to first ponder on other similar questions, such as, what is a narrative? And what makes a group of sentences be considered as a narrative rather than something else like a report, a list, an argument, or a description? Similarly, what makes a speech or address move from an exposition to a story? Although making an exact definition of narrative is sometimes difficult, I will adopt a general classification of what can be considered as narratives from McVey's (2020) work. McVey considers narrative as "the presentation of a temporal ordering of incidents laid out in a particular manner with the express purpose of representing some larger truth than the mere telling of the events themselves" (McVey, 2020, 74). He further states that "a narrative possesses a sort of metaphoric element, one that utilizes the temporal organization of events and characterization to achieve its deeper meaning" (McVey, 2020, 74). He thinks that while "science presents things as they are, and in this sense, they are eternal, possessing all middle with no beginning or end", narratives possess a beginning, middle, and end. Generally, he also wants us to believe that narrative "arranges events in a way that metaphorically connects us to the temporal wholeness of life itself, and in doing so alludes to a deeper, more philosophical whole than the sequencing of temporal events in exposition, reports, or science" (see McVey, 2020, 75).

Meanwhile, the *original position* by John Rawls is one of those PTEs that relies heavily on a narrative device that provides a cogent argument for a just society. In other words, the narrative as presented in *the original position* uses a story and is presented in a logical order aimed at persuading the audience in believing in the possibility of justice. For instance, the *original position* is like a story that tells us what justice looks like and how it can be attained in the hypothetical world; it is a scenario that involved a group of people who want to set up a new society. These people are confronted with the story that warrants deciding what principles of justice to be considered when designing a new society; to do so they have to put themselves under the veil of ignorance… (for more details see chapter four). This story provides the reader with some 'imaginary grips', which means, the thought experiment can effectively gain control of the reader's imagination of the events that are taking place in the thought experimental world. By implication, readers could become mentally involved in the scenarios of the thought experiment to produce a behavioural response and can gain an understanding of certain facts

not previously known. With the presence of an imaginary grip, the reader is persuaded to arrive at a certain conclusion brought by the story of the thought experiment, something that was not thought of *ab initio*. Furthermore, it can be deduced that narratives with persuasive power and one with a valid argument emanating from a thought experiment can also evoke changes in the mind of the reader, capable of causing a sway of belief during thought experimentation. It is important to note that changes in the psychological/mental state of the reader are certain due to the conclusion and the level of persuasion that the thought experiment brings to the mind of their readers. Although the level of persuasion might be difficult to fathom, perhaps, those with valid argument presents the logical justification for why changes in beliefs could be plausible due to the narrative impact of the scenario.

The argument I am raising is that *the original position, t*he violinist case and the ticking bomb are all typical examples of presenting a narrative of moral and political thought experiments that are presented as both narrative and argument, capable of creating an 'imaginary grip' on its audience. By implication, some thought experiments can shape our psychological/political beliefs. For instance, taking the Original position case, the conclusion on how resources could be fairly distributed which is what Rawls considers as justice could make any thinker who believes in the impossibility of justice, in particular, those who have an opposing view of what the principles of distributive justice are; as a result of the narrative and argumentative effects could change their political views and ideas. Drawing from the valid argument and the satisfactory conception of justice, there might be hope that Rawls's case can motivate actors to change their psychological behaviour, and this would become a good source of social, moral, and political development. For instance, the fact that parties in the *original position* would agree on unbiased principles of justice based on the simple rules of not being partial means that its readers would also have such feelings in their heads as an ideal knowledge to motivate their moral/political actions. It implies also that such knowledge of impartiality remains as an image in the reader's brain capable of informing the audience how a just political government would be formed in the real world in fighting corrupt behaviour. It is my claim that such knowledge in the brain of the reader can become a good source of political knowledge, hopefully, one that can also motivate us, and perhaps can be considered useful as a significant factor that drives the behaviours of people in the real world into a just way of politicking.

However, the above claims of mine could, of course, lead to more empirical questions, for instance, do narratives and argumentative cases actually work in shaping the behaviour

of their readers in the real world as we know it? There are more reasons to believe that narratives and arguments can impact the behaviour and the thought system of their readers. They can, because stories or narratives that produce '*narrative transport'* can provide the needed instrument for psychological, social, moral, and political motivation which can lead to behavioural changes in the real world. My concern is can a narrative or argumentative case act as action guidance? Finally, what sort of narratives produce narrative transport and how can this bring about moral/political motivation via thought experimentation? In the next section, McVey's studies on narratives and philosophical arguments will be discussed and used to provide a theoretical and empirical framework for responding to these questions.

7.3 McVey's Findings on Narratives and Philosophical Arguments vis-a-vis the Motivational Possibility of Political Thought Experiments

To unveil the motivational possibility of political thought experiments (PTEs), this section outlines McVey's (2020) studies on the relationship/impacts of narratives and philosophical arguments in behavioural changes. Findings from this study are aimed to provide some theoretical/empirical justification of what potential political thought experiments of this kind could look like in order to motivate or change the behaviour of their audience.

Design:

In Christopher McVey's (2020) work, he acknowledges how philosophical arguments and narrative scenarios motivate people differently. He seeks to find out whether philosophical arguments or persuasive narratives are most likely to motivate people to make charitable donations. The study hypothesised that "individuals who experience higher degrees of narrative transport will self-report more positive attitudes towards charitable contributions" (McVey, 2020). He uses empirical findings given different conditions (like the argument, narrative and control conditions) to investigate whether presenting readers with these conditions can best be expected to motivate behavioural change. The study is primarily designed to develop useful points of data in making the case that narrative transport is effective in swaying moral attitudes, intentions, and behaviour. 782 participants were recruited using an online platform called SurveyMonkey. They were divided into four conditions (i) narrative only, (ii) argument only, (iii) narrative + argument, and (iv) the control condition. In each condition, participants were asked to read a text and then after that, they were asked a series of questions designed to

measure how likely it was that they would give money to charity. These conditions are explained below.

For ***the argument*** condition, the study chooses Peter Singer's argument stated as follows:

> *"1. A great deal of extreme poverty exists, which involves suffering and death from hunger, lack of shelter, and lack of medical care. Roughly a third of human deaths (some 50,000 daily) are due to poverty-related causes.*
>
> *2. If you can prevent something bad from happening, without sacrificing anything nearly as important, you ought to do so and it is wrong not to do so.*
>
> *3. By donating money to trustworthy and effective aid agencies that combat poverty, you can help prevent suffering and death from lack of food, shelter, and medical care, without sacrificing anything nearly as important.*
>
> *4. Countries in the world are increasingly interdependent: you can improve the lives of people thousands of miles away with little effort.*
>
> *5. Your geographical distance from poverty does not lessen your duty to help. Factors like distance and citizenship do not lessen your moral duty.*
>
> *6. The fact that a great many people are in the same position as you with respect to poverty does not lessen your duty to help. Regardless of whether you are the only person who can help or whether there are millions of people who could help, this does not lessen your moral duty.*
>
> *7. Therefore, you have a moral duty to donate money to trustworthy and effective aid agencies that combat poverty, and it is morally wrong not to do so" (see McVey, 2020, 139-140).*

For ***the narrative*** condition, McVey (2020) also chose an emotional and true-life story of a girl called Mamtha (a narrative one), about the same word length as the argument, about a family rescued from slavery by a charitable donation. The narrative is as follows:

> *"Mamtha's dreams were simple—the same sweet musings of any 10-year-old girl around the world. But her life was unlike many other girls her age: She had no friends and no time to draw. She was not allowed to attend a school or even play.*

Mamtha was a slave. For two years, her every day was spent under the control of a harsh man who cared little for her family's health or happiness. Mamtha's father, Ramesh, had been farming his small plot of land in Tamil Nadu until a drought dried his crops and left him deeply in debt. Around that time, a broker from another state offered an advance to cover his debts in exchange for work on a farm several hours away.

Leaving their home village would mean uprooting the family and pulling Mamtha from school, but Ramesh had little choice. They needed the work to survive.

Once the family moved, however, they learned that much of the arrangement was a lie: They were brought to a sand mine, not a farm, and the small advance soon ballooned with ever-growing interest they couldn't possibly repay. This was bonded labour slavery.

Every day, Ramesh, his wife, and the other slaves rose before sunrise to begin working in the mine. For 16 hours a day, they hauled mud and filtered the sand in putrid sewage water. The conditions left them constantly sick and exhausted, but they were never allowed to take breaks or leave for medical care.

When Ramesh tried to ask about their low wages, the owner scolded and beat him badly. When he begged for his family to be released, again he was beaten and abused. Ramesh knew the owner was wealthy and well-connected in the community, so escape was not an option. There was nothing he could do.

Mamtha's family withered from malnutrition before her eyes in the sand mine. Every morning at 5 a.m., she watched with deep sadness as her parents left for another day of hard labour—and spent her day in fear this would soon become her fate. She was left to watch her baby sister, Anjali, and other younger children to keep them out of the way. Her carefree childhood was taken over by responsibility, hard work and crushed dreams.

Everything changed for Mamtha's family on December 20, 2013, when the International Justice Mission, a charitable aid organization funded largely by donations from everyday people, worked with a local government team on a rescue operation at the sand mine. Seven adults and five children were brought out of the facility, and government officials filed paperwork to totally shut down the illegal mine. After a lengthy police investigation, the owner will now face charges for deceiving and enslaving these families.

The next day, the government granted release certificates to all of the labourers. These certificates officially absolve the false debts, document the slaves' freedom, and help provide protection from the owner. The International Justice Mission aftercare staff helped take the released families back to their home villages to begin their new lives in freedom.

For Mamtha, starting over in her home village meant making those daydreams come true: She was enrolled back in school and could once again have a normal childhood. She's got big plans for her future—dreams that never would have been possible if the rescue had not come. She says confidently, Today, I still want to be a doctor. Now that I am back in school, I know I can achieve my dream" (see McVey, 2020, 132-135).

In the controlled condition, participants read some text from a middle school science textbook on Energy of roughly the same length as the narrative and argument conditions.

Measures for Motivation.

Participants were asked to respond to six hypothetical questions to act as a prompt after reading the text stimulus. They include as follows:

"1. It is morally good to give money to charities that help those in extreme poverty.

2. People like me should give money to charities that help people in extreme poverty.

3. Currently I feel motivated to give money to a charity that helps people in extreme poverty.

4. Right now, I have no desire to give money to a charity that helps people in extreme poverty.

5. I think it is important to support charities that help those struggling with extreme poverty.

6. Hypothetically, suppose we gave you an additional $10 for participation in this study, along with the option to donate some portion of it to one of six well-known charities that have been shown to effectively fight suffering due to extreme poverty. In this hypothetical case, how much of your additional $10 do you think you would donate?" (See, McVey, 2020, 145).

As part of the measurements/stimuli, from questions, one to five above, participants were asked to rate their agreement or disagreement with each statement on a 7-point Likert scale. In the 7-point Likert, prompts one and two were anchored by, "I completely disagree" and "I completely agree" with "I neither agree nor disagree" at the centre point. Prompt 3 was anchored by "Not likely at all" and "Extremely likely" with "Unsure" at the centre point. In the sixth question, the study poses "a hypothetical question to participants and asks them to indicate the amount they would be willing to donate on an 11-point scale from $0 to $10 at full $1 intervals. In general, this will help the study to determine how much participants would donate to charity organisations.

Transportation Scale:

To measure narrative transport, participants were given a modified version of Green and Brock's transportation scale. The transportation scale presented to participants includes:

"1. While I was reading the text, I could easily picture the events in it taking place.

2. While I was reading the text, activity going on in the room around me was on my mind.

3. I could picture myself in the scene of the events described in the text.

4. I was mentally involved in the text while reading it.

5. After finishing the text, I found it easy to put out of my mind.

6. I wanted to learn how the text ended.

7. The text affected me emotionally.

8. I found myself thinking of ways the text could have turned out differently.

9. I found my mind wandering while reading the text.

10. The events in the text are relevant to my everyday life.

11. The events in the text have changed my life" (See, McVey, 2020, 148-9).

Given the scale consisting of 11 prompts above, participants were asked to rate their agreement or disagreement with each statement on a 7-point Likert scale, the same as previously used in the measurement for motivation above.

Results

Upon completion, the study presents the following as the overall transportation scores against all conditions; and the results are as follows: 53.835 (Argument), 51.610 (control) 56.174 (narrative), 56.352 (narrative+argument). The result from transportation scores can be demonstrated in the chart below:

Fig 5: Transportation score by all conditions (see McVey, 2020, 161).

The result from the six prompt questions about giving to charity is as follows: starting with question one, asking whether 'It is morally good to give money to charities ', the mean results on the 7-point scale Likert scale include- 5.986 (argument) 6.067 (control) 6.341 (narrative), and 6.350 (narrative+argument). For question two which states 'people like me should give to charities', the mean results on the 7-point Likert scale were as follows: 5.138(argument) 5.310. (control) 5.783 (narrative), and 5.799 (narrative+argument). For question three, which states 'currently, I feel motivated to give to charity', the mean results on a 7-point Likert scale were as follows: 4.681 (argument) 4.676 (control) 5.226 (narrative), and 5.210 (narrative+argument). For question four, which states 'right now, I have no desire to give to charity', the mean results on a 7-point Likert scale were as follows: 3.612 (argument) 3.805 (control) 3.313 (narrative), and 3.402 (narrative+argument) For question five, which states 'I think it is important to support charities', the mean results on a 7-point Likert scale were as follows: 5.612 (argument) 5.757 (control) 6.065 (narrative), and 6.070 (narrative+argument). Finally, for question six, which answers 'how much money would you donate if given $10 now', the mean results on a 7-point Likert scale were as follows: **$**4.17(argument) **$**3.97 (control) **$**4.70 (narrative), and **$**5.04 (narrative+argument).

In addition, all individual results from the six questions were calculated to arrive at the following mean motivation score: argument (26.725) control (26.898) Narrative (28.840), and narrative+argument (28.487). A bar chart summary clearly shows the differences between all groups as shown below:

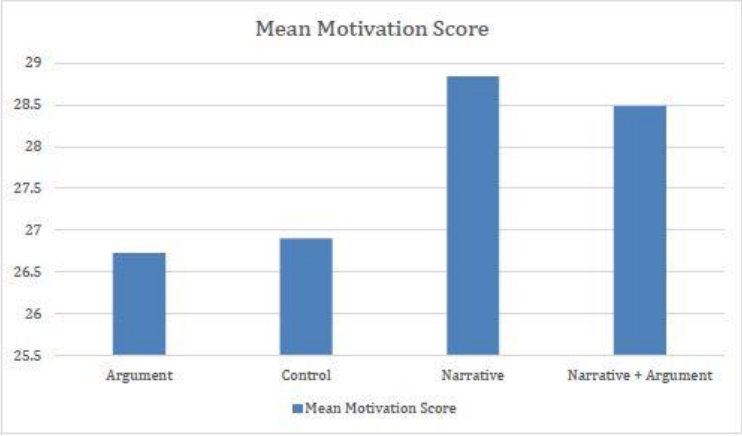


Fig 6: Mean motivation score by all conditions (see McVey, 2020, 200).

Considering the general result and the above differences, the study posits that argument and control are significantly different from narrative and narrative + argument and that there is no significant difference between argument and control or between narrative and narrative + argument. McVey acknowledges that the differences here are small but argues that they are nonetheless significant. He thinks that "they do show promise and support the hypothesis that narrative outperforms argument with regard to changing participants' views about the morality of donating to charities that combat extreme poverty" (see McVey, 2020, 201).

7.3.1 Conclusion/ Major Findings

In conclusion, the study also argues that the results "show promise and support the hypothesis that narrative outperforms argument with regard to changing participants' views about the morality of donating to charities that combat extreme poverty" (see McVey, 2020, 156). The study also predicts that given the above results "If we compare the mean donation rate for participants in the narrative condition and compare it to the mean for the control condition, we see a roughly 22% increase in donations. Viewed in this way, we can easily see the importance of such an effect if the dollar values are increased. After all, a 22% increase could turn a possible $1,000,000 donation into a $1,220,000 donation. That additional $220,000, of course, could make a monumental difference in saving real human lives" (see

McVey, 2020, 182). He uses these findings to also confirm that "higher levels of narrative transportation predict higher degrees of story-consistent belief changes" (see McVey, 2020, 207). The implication is that as shown in figure 5, scenarios that evoke higher narrative transport can cause changes in behaviour more than those conditions with little or no narrative transport. The findings in figure 5, show how 'narrative+argument' motivate readers more than other possible conditions. These findings are used by McVey (2020) to present prima facie evidence on the level of participants' changes in behaviour in charity donation when they are exposed to different situations. The study confirmed that "reading narratives has resulted in higher levels of narrative transport than reading non-narrative texts and that transport is not correlated with any of the demographic measures collected in the study" (McVey, 2020). Therefore, from the pieces of evidence, as captured in these findings, the control condition whose empirical results fall lower than others may not be effective at convincing research participants, particularly towards giving a surprise bonus award to charity. In contrast, 'transportable' moving narratives do appear to be more effective than others in all scenarios (see McVey, 2020). The study further confirms that "higher levels of narrative transportation predict higher degrees of story-consistent belief changes"; according to the study it can change "aspects of our psychology, moral beliefs, intentions, and behaviours" (see McVey,2020, 207-8).

McVey's (2020) findings can be used in this chapter to argue that in situations when political thought experiments contain both features of narratives + arguments, it therefore, means that, such cases would provide a higher level of narrative transport, likewise the hope for higher motivational power in shaping the moral, political and psychological behaviour of its readers. It can also be deduced that, from McVey's (2020) studies, thought experimenters in the discipline should begin to explore imagined situations that can provide narrative transport so that political thought experiments could produce the needed motivational force as predicted in the study to shape the feelings and political behaviour of its audience, especially in changing their real-world action as we know it. We hope that it is through this means that thought experiments in the discipline can provide a strong motivational force in comparison to non-thought experimental arguments or traditional philosophical arguments if redesigned to meet these features as demonstrated in McVey's study.

7.4 Political Thought Experiments and the Question of Narrative Transportation

Critics of my view might argue that political thought experiments cannot motivate the audience because it lacks the various elements for narrative transport. The thought-provoking question is, can political thought experiments facilitate narrative transport? How can we redesign political thought experiments to better provoke narrative transport, in particular one that will meet the various elements set out for narrative transport as described by McVey's 2020 study in motivating action? I will discuss this in relation to Rawls's thought experiment, first to identify how Rawls's claims meet some of the conditions for narrative transport. Later in this section, I will also outline what we need to do to meet other expected conditions (for narrative transport to happen) so that political thought experiments could be capable of changing the political behaviour of their audience in the real world if those conditions are met.

However, for a story to be conducive to narrative transport in order to be effective in swaying our moral attitudes, intentions, and behaviour, McVey's study believes it must have the following four elements to motivate the reader: (i) Relatable characters-these are the characters with whom the story receiver can identify (ii) Vivid imagery- the story must contain a plot that the 'story receiver' can imagine. In other words, the story should invite the readers to imagine what they must be feeling as the events of the story unfold. (iii) Easy to understand- the event of the story must be clear to all readers including realistic and unrealistic stories, for instance, fictional and non-fictional (narratives) that are understandable are all conducive for narrative transport. The study believes that there is no appreciable difference in the level of narrative transport that can be triggered by both types of narratives. (iv) Emotional arc: In this case, the story must be emotional, appealing, persuasive, and end on a positive note and appropriately placed climax. Perhaps, it must have a beginning, middle, and end, such that, in each step of the story we can easily imagine the situations as they happen (McVey, 2020, 136-145). Thus, a narrative with the above ingredients evokes transportation (meaning readers of such stories can lose themselves or become sympathetic to the story, which means they are prone to taking action while reading such stories). McVey's study posits that Mamtha's story as used in the study or any other narratives with these elements or features "can shape our very souls and they do so by altering our beliefs in some ways", which then persuades readers to donate to charity organisations (McVey, 2020).

The thought-provoking question is, can any political thought experiment meet these conditions set out by McVey's study for narrative transport? Or can our beliefs be altered while

reading Rawls's thought experiment? I argue that Rawls's thought experiment falls within the logic of narrative, and it does meet some of the conditions, which means it might be hoped to motivate some actors to change their behaviour while reading the scenario. As argued earlier, *the original position* is a narrative and at the same time an argument that can persuade the reader to understand the possibility of absolute justice, in particular, where the distribution of social goods can be justly shared (distributive justice). It is our belief that the thought experiment can facilitate narrative transport, in particular, one that can bring changes in behaviour concerning the desire for fairness. However, narrative transport in thought experimenting generally suggests that when we lose ourselves in a story that comes with the imagined scenarios, our behaviour and actions are supposed to change to reflect the thought experimental claims. For instance, when reading Rawls thought experiment, we are supposed to change our psychological behaviour first, by thinking that distributive justice is possible, second that the features that are deduced from it (such as freedom of speech, rule of law, and collective action principles) can become instrumental for our political development and to address the issue of corruption in the real world as we know it (I will be discussing the efficacy of these features in the next chapter). Below, I will begin to explain how Rawls's thought experiment could meet some of the above ingredients for narrative transport, and later, I will demonstrate what we need to do to meet other conditions that might be lacking which thought experimenters need to consider when designing future PTEs: (i) On relatable character- We can rightly say that Rawls thought experiment consists of a connected group of stories that involves rational agents. Rawls buttressed this by stating that "the complexity of problems of justice is up to the '*persons'* in the original position to decide" (Rawls, 1971, 454). These persons according to Rawls are rational men and women who are representative of the people in need of just arrangement in the new society. Perhaps, they are characters that readers can easily identify with to understand the nature of justice in the original position. Generally, most successful thought experiments are capable of using the reader as a relatable character. For instance, a thought experiment can prompt the reader with (a what-if statement….), such statements are often asking the reader to imagine seeing herself in a particular hypothetical world and then ask her what action s/he could take (for example, *the trolley case, the pond case,* and many others are typical examples of thought experiments within this category). By implication, the reader goes through the scenarios to experience the event of the story as it unfolds in her mind, and with the picture of the scenario in her head being processed, the reader can then provide a response that is determined by the events of the thought experiment. (ii) Easy to understand- Rawls thought

experiment is also easy to understand, by the virtue of its valid argument readers need to follow the disguised premises, which are sequentially arranged in reaching the conclusion of the thought experiment (for more details see chapter four for my reconstructed version of *the original position* thought experiment as an argument). (iii) Vivid imagery- Rawls thought experiment contains this element because the scenario is imaginary and readers can go through the scenario in their head to see how the story unfolds, in particular, how *the veil of ignorance* takes us away from things that could affect a fair choice and finally how justice can be reached in its hypothetical world. (iv) Emotional arc: Rawls's case also provides a satisfactory conclusion on how men and women in the state of nature could reach fair distribution of social goods, meaning it ends on a positive note on the possibility of justice in the original position.

However, Rawls's case might not produce the emotional feelings that might be required to persuade people in taking action immediately, as is the case in the story of McVey's study. For instance, the story in McVey's study provides adequate emotional feelings that can easily persuade people to donate to charity. According to the study, "the story begins with an emotional hook: the loss of Mamtha's childhood innocence. We then continue the emotional ride as we learn how Mamtha's family went from innocent farmers to force slave labourers. In the end, we experience joy and relief when the family is rescued and experience hope that Mamtha will now get to fulfil her dreams. Much of this uptick in emotion is centred around the rescue which is facilitated by the charitable organization" (McVey, 2020, 137). Of course, emotional hooks do not exist in Rawls's case. Readers are most likely not going to be too emotional when reading the original position narrative compared to the story of Mamtha as used in the study. The implication is that, although both are narratives, they are most likely going to persuade people differently and their level of narrative transport will differ. Based on these findings, it can be argued that by meeting some of the standard requirements for effective narrative transport, some PTEs can produce narrative transport, which means they are prone to motivate the audience in a way that can change their political behaviour while reading the imaginary case in the real world. It is also the case that we need to start engaging in emotionally moving scenarios when developing future political thought experiments to gain a high narrative transport. From McVey's study, emotionally moving scenarios can guarantee a quick swaying of belief and changes in behaviour, which is what we need when developing future thought experiments in political philosophy.

7.5 Critical Comments

It is important to note that my discussion of narratives and their persuasive power in this chapter suggests that more detail should be added to make the thought experiments more persuasive. This of course needs further clarification to understand what exactly needs to be redesigned and how. Many critics might confuse this claim to mean bringing new empirical details to a scenario by readers which are not often part of the thought experiments. The thesis claimed that within the context of the discussion of McMahan, such details are not important because of their ability to lead readers to a wrong intuition, hence they are to be abstracted away. By *abstraction,* readers should not consider imaginary scenarios as if they were actual, adopting the abstraction strategy here means the ability to focus on ideas alone as against their empirical details during thought experimentation because they are not relevant to the existing thought experiments (see section 5.5 for this explanation). However, what the thesis meant (in this chapter) by redesigning political thought experiments to increase the persuasive power does not necessarily invite adding empirical details to current existing thought experimental claims. What is important here is that in the future due to the prospect of narrative and argumentative components of a scenario as demonstrated in McVey's study, thought experimenters should consider making narratives and arguments an important part of their scenarios when creating new thought experiments in the discipline.

7.6 Summary and Conclusion

This chapter reminds us of the fact that some political thought experiments are valid arguments because they are capable of providing a conclusion that necessitates from their premises. Although they can be considered as arguments, the chapter outlines some concerns with valid/sound philosophical arguments, in particular, the fact that in general presenting people with valid/sound arguments does not do so much in altering their actual behaviour. To defend political thought experiments vis-a-vis this concern, I argue that some PTEs can both be considered as arguments and narratives. Using Rawls's thought experiment as a case study, I present the narrative component of *the original position* case, in particular, its possibility of facilitating narrative transport. It suggests that political thought experiments can be hoped to be motivationally possible, especially when their imagined scenarios are capable of providing narrative transport. With narrative transport in thought experiments, it suggests that imaginary cases via thought experimentation can be hoped to alter our behaviour as deduced from McVey's (2020) findings. Findings from McVey's (2020) work suggest that for thought

experiments to shape the political belief of their audience in the actual world, such scenarios must contain an identifiable character, vivid imagery, easy understanding, and an emotional arc. These are some of the identified *factors* that can provoke narrative transport in a story, needed to persuade the feelings of its audience and guide our political motivation in the real world. I believe that for thought experiments to shape the political behaviour of its readers which might look impossible with other traditional philosophical arguments, this chapter suggests that thought experiments must be conjectured within these conditions in addition to the argumentative structure, in particular, one that has an emotional hook. In other words, with the right feature of narrative transport, some thought experiments can be seen as a formidable tool for psychological development that tells us how we ought to behave in the hypothetical world, which can help change our behaviour and political beliefs in the real world as we know it.

To conclude, Rawls's thought experiment might be hoped to motivate actors to change behaviour, for instance, *the original position* presents a narrative and valid argument that readers can consider useful in taking the right and logical political action. Of course, it has some of the right features for narrative transport for motivational effect on people which means it can produce some level of narrative transport that is required for real changes. It is pertinent to note that PTEs need to produce some emotional feeling in the mind of their reader to produce a high level of transportation that is required to bring higher motivational changes. McVey's (2020) study above provides empirical proof to justify these claims, in particular, how narrative transport via emotional stories can evoke changes, feelings, emotions, and behavioural of its audiences. As seen in Rawls's case, *the original position* might not evoke the required emotional feelings like the narrative case in the study, but the validity of its narrative argument means the case is promising for narrative transport. However, political thought experimenters must also begin to think of emotionally moving narratives as part of their thought experimental claims, so that PTEs can provide higher narrative transport, perhaps one with strong persuasive power. I believe that thought experiments with the right narrative in the discipline can evoke the needed emotional feelings to sway our political beliefs. When these conditions are truly met, I believe that imagined situations in political theory can shape the feelings and political behaviour of its audience in such a way that can enable changing of reader's political behaviour while reading through the thought experimental claims.

Employing the argument of this chapter, I have also claimed that readers of Rawls's thought experiment can now have a feeling and understanding of justice that could provide a

rational justification for how to address both moral and political issues in society. For instance, the nature of justice in Rawls's thought experiment provides rational features that do tell us how everyone can live a good life. The features of collective action, rule of law, and freedom of speech can be drawn from it to develop a satisfactory conception of justice, one that can help in the eradication of corruption in any society, especially in Nigeria. For instance, using the logic of collective action, rule of law, and freedom of speech (which are some of the features that Rawls's idea of justice brings to the mind of its readers during the narrative plus argumentative exercise), individuals and nations can benefits from these features in fighting corruption, especially when they are implemented in governance or applied in principle in our political situations. These claims shall be advanced further in the chapter below when I will be discussing the institutional possibility of political thought experiments.

CHAPTER EIGHT

8. INSTITUTIONAL POSSIBILITY: A CASE OF HOW RAWLS'S THOUGHT EXPERIMENT COULD OFFER AN ALTERNATIVE APPROACH TO THE PROBLEM OF POLITICAL CORRUPTION IN NIGERIA.

8.1 Introduction

Corruption is one of the critical political issues that has hindered the social and political development of different countries all over the world. In Nigeria, political corruption has hindered social, and economic growth which is affecting the full implementation of the dividend of democracy in the country. Using Nigeria as a case study, this chapter will examine the nature and effects of political corruption and using the idea of the institutional possibility of PTEs, I shall discuss how Rawls's thought experiment could offer an alternative approach to the problem of political corruption in the country. However, as the effects of corruption continue to rise in different countries of the world, it can be deduced that most governments have also failed to provide the required political will to eradicate corrupt behaviour in politics. This suggests to us that to facilitate the fulfilment of citizens' life plans, and perhaps provide other democratic dividends to everyone, good social and political policies against corruption must be put in place to drive society. Conversely, due to the lack of effective anti-corruption policies, most Nigerian citizens do find some difficulties in achieving their life plans. The thought-provoking question is, can a thought experiment in political theory become relevant to the issues of political corruption in Nigeria? A model analysis of this question using the institutional possibilities of political thought experiments (PTEs) will provide an in-depth analysis of how the problem of political corruption in Nigeria can be addressed by means of thought experimentation.

Meanwhile, the chapter also shows how Rawls's thought experiment could provide a better framework for responding to the problem of corruption in the real world. To achieve the above aim, I intend to show how important features (such as collective action, freedom of speech, and rule of law) deduced from Rawls's *original position* can translate to policy guidance which can provide an alternative approach to responding to the issues of corruption in Nigeria. In other words, Rawls's thought experiment would provide useful features in responding to the issues of political corruption in Nigeria, especially regarding the question of what policy can best be used in responding to the corruption issues in the country. To achieve this aim, I intend to adopt the feature of collective action in Rawls's thought experiment in proposing the collective action framework as an alternative tool for tackling the problem of political corruption in Nigeria. The

section below will begin to examine the problem of corruption and how it has become an issue in Nigeria to date, hence different efforts by the government have produced little or no success in the country. The reason for this failure is that the politicians instead of fighting corruption have decided to pursue personal interests, generally, this has affected the fight against corruption in Africa including Nigeria. Also, corruption in Nigeria has become more entrenched along with the efforts to curb it despite the effort of different political administrations.

8.2 The Nature and Scope of Political Corruption in Nigeria

Before I proceed to the various discussions on how PTEs can be used to respond to corruption issues in Nigeria (using the institutional possibility of PTEs), some empirical analysis of the problem of corruption in Nigeria shall be exposed, as a case study for the thesis. Using the case study, I aim to examine how corruption has affected Nigerian society; findings from this discussion shall be used (in subsequent sections) to examine how PTEs can be used to address political corruption in Nigeria.

Briefly, political corruption is arguably one of the most significant issues in global politics. With a lack of accountability in governance, much has been put into fighting corruption by various governments across the world, but the results are unsuccessful. Existing policies/laws that would fight against political corruption are inadequate. To date, within the corridors of power, continuous misuse of collective resources by government authorities through misappropriation of public funds and bribery is increasing. Perhaps, corruption is the main cause of poor infrastructural development and extreme poverty around the world. According to the United Nations Office on Drugs and Crime (UNODC), political corruption is the bane of any liberal society and the main agent that "stifles entrepreneurship, professionalism and erodes the values of hard work and honesty, and is one of the root causes of under-development in our society" (United Nation, 2017). By implication, every country is expected to fight and eradicate corruption at all levels to improve the social well-being of its citizens. Although different countries have made anti-corruption laws, many others are also faced with challenging thoughts on how it would be eradicated from society.

In Nigeria, political corruption is increasing instead of decreasing because the existing laws and policies are inadequate. For instance, the immunity law, as stipulated in the constitution, encourages corruption because of how it has been abused. In Nigeria, the

immunity law states that the President, Vice president, and all 36 Governors and their Deputies are not placed on trial for any crime in an open court during their political tenure. By it, these politicians gain legal support not to be tried in the open court for their corrupt behaviour and other political crimes (see Fabamise, 2017). It means with immunity laws, the constitution in several ways can be manipulated to fit the greedy desires of these politicians. This has led to a serious increase in political crimes that have become a source of great concern for which PTEs can be used to address. Generally, political corruption has affected both the individual and institutional behaviour of Nigerian society, particularly putting the masses into unnecessary hardship. In addition, the achievement of both institutional goals and the individual's life plan has become impossible. Given this concern, this chapter shall address the negative implications of political corruption in Nigeria and shall further discuss how this has affected both institutional and individual moral and political behaviours. As we proceed, I shall discuss how good moral and political behaviour would be deduced particularly through the study of hypotheticals via thought experimentations in political theory.

What is political corruption (PC)? According to Kramer (1997), PC is "the behaviour of elected public officials which diverges from the formal components-the duty and powers, right and obligations- of a public role to seek private gain" (Kramer, 1997). The United Nations (2017), has reported that political corruption is a common behaviour with many faces, the most common categories are 'bribery', 'fraud', 'embezzlement', 'extortion', and many others (United Nations, 2017). This thesis agrees with the above definitions of political corruption and they shall be used to advance the claims of this chapter. Political corruption involves any political authority, for instance, a president, a governor, a minister, a senior civil servant, and any nominated, elected public officer or anyone who engages in a certain criminal behaviour of tailoring the available social goods that are meant for everyone for personal interest. In Nigeria, the cancer of corruption has been spreading rapidly among these categories of people for several years and is affecting Nigeria's economic growth and causing a serious increase in the poverty rate of the country. As the most populated black nation in the world, with about 200 million citizens, more than 69% are living below the poverty line (see Abbas Jimoh, 2019). Meanwhile, the country is rich in different natural resources, including crude oil, gold, diamonds, and many others. Presently, very little progress has been made to improve the economic and social conditions of its people as a result of the increasing level of political corruption. There is no good road or rail network in the country, and the electricity supply is inadequate because politicians loot the money that is meant for these institutions for their private use. Similarly, due to different forms of corrupt practices, particularly in the

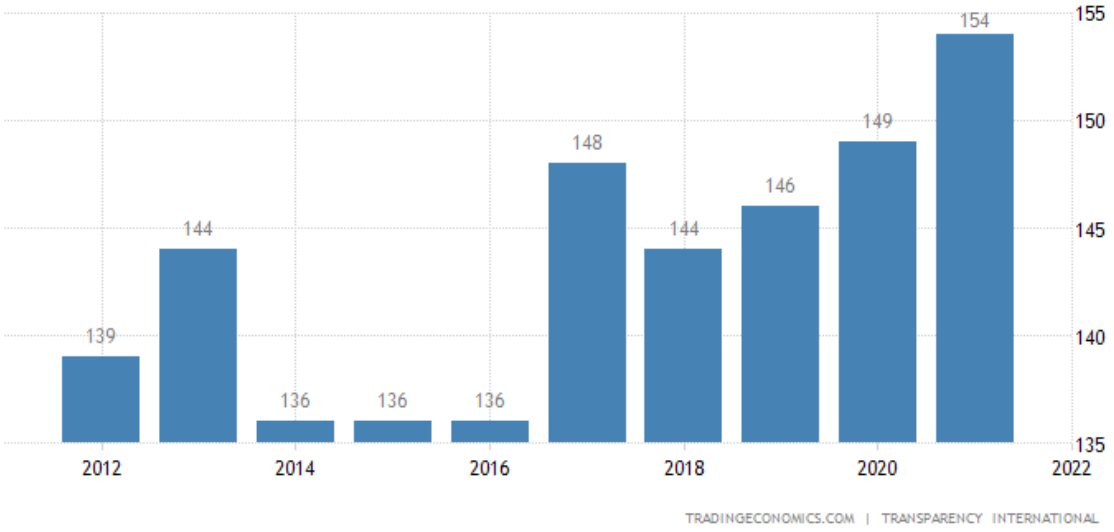
educational, security, and medical sectors, there is currently a poor security network and poor educational and medical infrastructural development. While most existing hospitals lack up-to-date medical facilities, the masses also lack the resources to seek medical attention abroad and are left to suffer. Perhaps, the wealthy class, including politicians, seek medical treatment abroad, hire private security, buy private jets to avoid the bad roads, and even send their relations abroad for a decent education in the face of poor academic infrastructure in the country. The implication is that by this action, Nigerian citizens have been dehumanized, oppressed, exploited, and victimized by corrupt leaders. And for many decades, due to these corrupt behaviours, there is little or no hope for the actualization of the individual's life plans in the face of political corruption in Nigeria. However, the worst scenario is that corrupt practices have come to be seen today as normal, perhaps being considered the right way of doing things by anyone who wants to survive given this precarious situation.

Meanwhile, the fight against corruption by the present government of President Muhammadu Buhari's administration is not sufficient. Critics like Obasanjo (2019), think that the President so far has politicized the fight against corruption. He has been accused of favouring the corrupt political class from his political party, which means he is biased because many corrupt politicians from his region are still walking free without an open trial in court. Obasanjo thinks that the President has also been nepotic in his political behaviour, especially giving most political appointments only to his relatives and close associates from his political zone. Such biases in the fight against corruption or the appointment of public officials should, according to Obasanjo, be considered harmful to democracy and the rule of law (see Obasanjo, 2019).

According to Transparency International (a global coalition against corruption), the *corruption perception index* (2021) in Nigeria is ranked 154, out of 180 countries surveyed with a score of 20% out of 100. The implication is that Nigeria is perceived to be one of the most corrupt nations in the world (Transparency International, 2022). As reported by Transparency International (2022), the overview of the *corruption perception index in Nigeria* has constantly remained on the rise. According to Transparency International (2022), "the Corruption Perceptions Index (CPI) aggregates data from a number of different sources that provide perceptions by business people and country experts of the level of corruption in the public sector. It draws from non-partisan and non-political influence from institutions specializing in governance and business climate analysis covering expert assessment and empirical views of bribery rate in a country" (See, *Transparency International*, 2022, 26 for

more methodological details). In Nigeria, the report from the year 2012 to 2022 justifies the above claims, as shown in the chart below:

Fig. 2: A chart of various corruption perception indexes in Nigeria from 2012 to 2022



From the above findings, it shows that in 2019 Nigeria was 148 in the corruption perception index and the 34th most corrupt country in the world from the 180 that were surveyed, perhaps compare to 2018s relatively high ranking. Therefore, for a decade, corruption is on the increase with the worst record in 2022 as 154 during the administration of President Muhammadu Buhari (Transparency International, 2022). While the existing local and international laws and policies would not eradicate political corruption in Nigeria, it is pertinent to ask one fundamental question; what would be the immediate causes of corruption in Nigeria?

8.3 Causes of Political Corruption in Nigeria

The high rates of corruption in Nigeria can be associated with both individual and institutional failures as the two most common causes of political corruption in Nigeria. There is a high rate of dictatorship (as a result of the immunity clause in governance), lack of institutional accountability, and bribery among political leaders that have caused a serious increase in corruption in the country. For instance, from the years 1993 to 1998, General Sani Abacha, the then-military head of state, embezzled over five billion dollars as a result of his dictatorial behaviours and lack of accountability (See. BBC, 2018). Part of these stolen funds has been recovered by President Mohamed Buhari, but there are serious concerns about how the current administration is spending the funds. The government claims that the recovered money has been shared among poor citizens (See. BBC, 2018). But many critics think that the

same money could be looted for a second time because there is a lack of accountability and transparency on how the government is spending the recovered looted funds. Because the government of the day does not have records of how the Abacha loot is spent, it follows that there were no right constitutional procedures for the disbursement of such public funds (see Malami, 2020). It suggests that the political plans of the present government are that of a dictatorship. The question also is, who are the poor beneficiaries of these funds? How could a government spend money that is meant for the public without following a due process that involves budget approval from the national assembly? These questions are yet to be answered by the government which is becoming a matter of great concern.

Similarly, the individual culture of bribery is another factor that could also impede the fight against political corruption. In Nigeria, bribery has also increased exponentially in recent years, and with a lack of policies to check this evil, the common people in the society are denied the true dividends of democracy. With the culture of bribery, a public officer can easily use some of the embezzled funds to bribe anti-corruption agencies or other institutions of government to evade justice. In the past, efforts to combat embezzlement were compromised by a weak punishment and a flawed legal system that allows corrupt persons to go unpunished. For instance, Mr James Ibori, the former governor of Delta State from 1999 to 2007, was accused and was later freed of all corrupt charges in the Nigerian court. But in 2012, he was caught and persecuted for money laundering/fraud and was sentenced to 13 years in prison for the same offence in the United Kingdom (see, BBC News, 2012). Many think that his freedom in Nigeria is a result of a compromised judiciary and anti-corruption agencies. The general question is, why do politicians choose bad political behaviour to the detriment of the citizens who voted them to power? These concerns are evident in the fact that there is a need for behavioural change in the Nigerian political system. However, in situations whereby certain individual political leaders continue to compromise with the constitutionally acceptable standards, it becomes difficult if not impossible for any government to put an end to the persistent culture of misappropriation and/or embezzlement of public funds (corruption) in our society. However, institutional failure is another reason why political corruption in Nigeria can not be eradicated. This failure can be discussed in two possible ways below:

i) Failed Legislative/Administrative Procedures in Policymaking.

Any country with failed legislative and administrative procedures in policymaking cannot grow in governance, and neither can such a country progress in national development. A nation is also bound to fail when its policies are inadequate in the fight against political corruption.

The idea is that good policy in governance is fundamental to the social, economic, and political growth of any country. In a democratic setting, good policies are vital because they serve as action guidance, particularly in directing our moral and political behaviours. However, in a constitutional democracy, it is the responsibility of the legislatures to make laws and policies for the executive to enforce; the executive implements the laws; while the judiciary interprets the laws. The failure of any of these arms of government could result in bad governance/corruption. By implication, it is pertinent to have checks and balances in place within the three arms so that political power is not concentrated in the hands of individuals or one arm of government. This process, if followed, makes it easier to fight against corruption and check the excesses of the people in power within a political space. Perhaps, the concentration of power in any arms of government can lead to dictatorship, which makes it impossible to fight against political corruption.

There are so many ways the Nigerian government has failed in both legislative and administrative procedures that are hindering the fight against political corruption. Firstly, Nigerian legislatures are prone to making laws that provide an enabling environment for corrupt behaviour among political elites. For instance, the immunity clause (also known as immunity law) that is in the constitution is a typical example of a failed legislature in the Nigerian political system. By immunity law, some politicians are constitutionally granted exemptions from legal proceedings. It is one of such laws that cover governors and their deputies, presidents and their vice presidents, etc. from being placed on trial in a court of law when caught committing crimes when they are in power. The Nigerian 1999 constitution as amended, grants this immunity clause to these principal officers of government and is supported by an act of the legislature. It states that "The immunity clause is meant to provide a shield for the person of the President, Vice President, Governor or Deputy Governor from frivolous or vexatious litigation in respect of personal or criminal proceedings that would distract him/her from the serious business of governance." (Section 308, *Nigeria constitution*, 1999). This law seems anomalous and is raising a lot of concerns; for instance, can the president really get away with anything? And why would a law like this be instituted in modern society? Critics of this view claim that "the purpose of such clause (law) has been manipulated to promote injustice, impunity, and political corruption. Therefore, the clause should be reviewed to checkmate the spate of anomalies being perpetrated under the guise of immunity clause" (see Azeez, 2018). This calls for serious concern because it means that even if they are found embezzling public funds or taking away the civil right of others by killing innocent citizens

they cannot be tried in court. The question is why should the legislature support such laws/ policies that lack moral values in the face of the increasing level of political corruption in the Nigerian political system? This law must be revisited because it lacks fairness, equity, and natural justice. However, by making politicians appear above the law, it denies the masses their right to seek justice and also silences the law against the objective will of all. The idea is that it does not support the fight against corruption, which is a big concern for the growth and development of Nigeria. Therefore, it can be suggested that an act of the legislature that supports the immunity clause in any government manipulates the will of the people, and it is an act against the interest of all, thus, it is tantamount to a serious injustice to the individual members of the society. Perhaps, any law that takes away the moral and political rights of any member of society should be considered draconian and must be revised.

ii) Failure of Foreign Governments and International Institutions to implement policies against Corruption

Failure to establish enforceable mechanisms to cope with corrupt practices by foreign governments and/or international institutions also contributes to the corruption crisis in Nigeria. For instance, some capitalist nations, in particular, provide policies that promote corrupt practices by international investors. With International Anti-Corruption Day set up to remind everyone both locally and internationally about the fight against corruption, very little progress has been made to address cases of corruption globally. There are serious concerns that international institutions are not abiding by the rules guiding the United Nations' mantra for the global anti-corruption campaign.

Similarly, Falana (2019) has accused foreign capitalist countries of providing the enabling atmosphere for corruption to strive in Africa, particularly in Nigeria. He stated this while claiming that:

> "Corruption is nurtured and sustained by capitalism. Any capitalist country, the USA, and Great Britain, among others, are perpetrators of this act. These countries warehouse the bulk of the stolen money from Nigeria. When you are bringing the stolen money, they turn the other eye because they want to use the money for development. Because there is no colonialism again, they encourage criminals all over the world to bring their stolen wealth and keep it in their country to the detriment of our people and generations unborn," (Falana, 2019, 1).

Falana thinks that foreign countries provide a suitable atmosphere for politicians to embezzle and launder money abroad, even when they understand that a certain amount of money can never be owned by an individual who might just be ruling a country for just a few years. By implication, money laundering has become a huge business for the capitalist nations to the detriment of the poor Nigerians who hardly get one dollar per day below the amount that is needed to exit the poverty line. Falana (2019) also argues that if foreign institutions and governments refused to set standard laws to check the inflow of money to their country, they are part of the crime committed by corrupt politicians in Nigeria. For instance, many Nigerian politicians have used proceeds from corruption for investment abroad like buying personal properties (houses and exotic cars) from the proceeds of corruption. Thus, funds from such properties are yet to be repatriated to Nigeria. The fact that stolen wealth from Nigeria has become a source of income for these foreign institutions to the detriment of the poor Nigerians is a source of great concern. Therefore, there is a need for new behavioural changes by foreign governments and institutions to cope with the challenges of political corruption in Nigeria. Hence, he thinks that capitalist nations must do more to protect emerging countries from corruption by making enforceable laws in their countries to discourage political corruption from other less developed countries. The next section shall critically examine how political corruption affects Nigerians as a people.

8.4 The Effects of Political Corruption in Nigeria

The extreme level of political corruption in Nigeria has raised serious concerns, particularly within the domain of human behaviour. The question is, what are the effects of political corruption on human behaviour? There are two possible ways political corruption could affect society, they include:

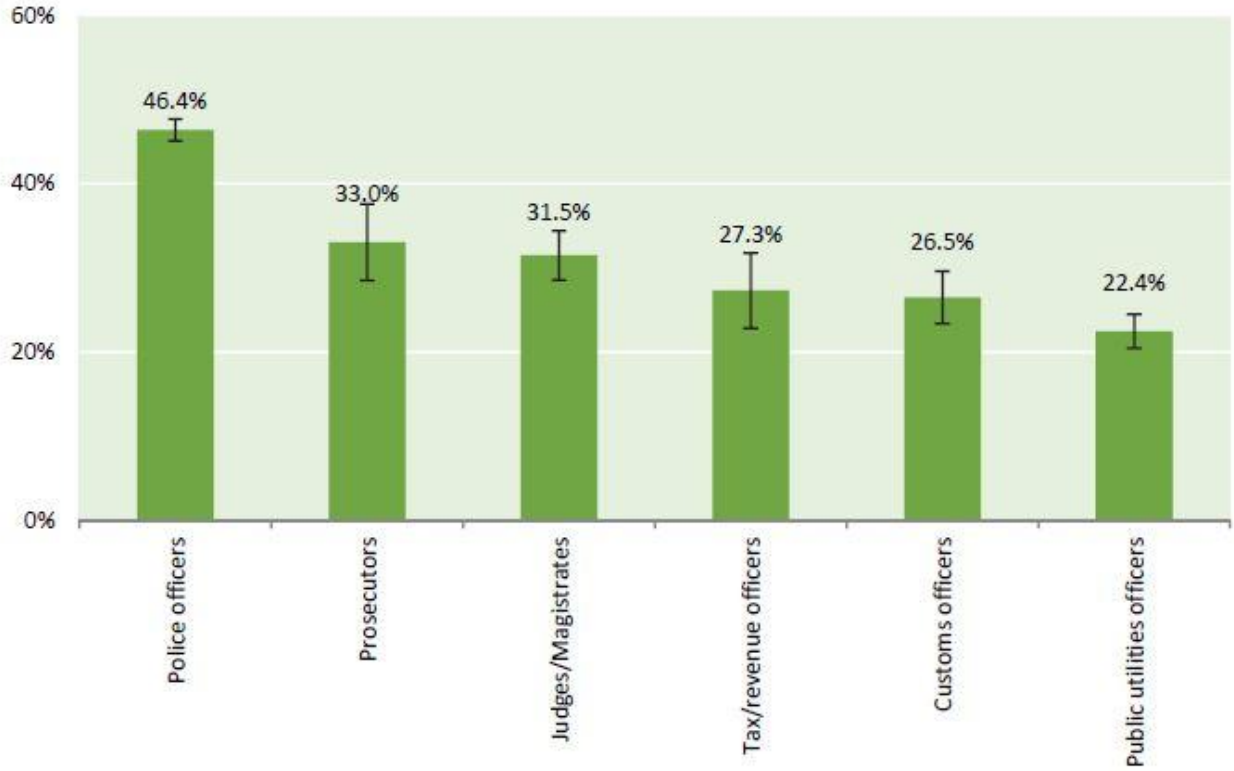
i) The Effects on Public Institutional Behaviour

Political corruption could pose a serious concern to public institutions in Nigeria as a result of the underfunding. The fact that politicians do embezzle public funds means some public institutions could be underfunded. When their staff are underpaid this poses serious risks for society. For instance, underfunded institutions may resort to the collection of bribes from the public in other to meet their institutional needs. According to the United Nations National Bureau of Statistics (2017), it was estimated that "the total amount of bribes paid to public officials amount to $4.6 billion in purchasing power parity terms—the equivalent of 39% of

the country's federal and state budgets for education in Nigeria" (United Nations & National Bureau of Statistics, 2017). The chart below summarises the prevalence of bribery to selected public officials in Nigeria.

Fig. 3: A chart indicating the prevalence of bribery in Public Institutions, in Nigeria

(Source: United Nation and National Bureau of Statistics, 2017)



The figures above (in percentage) are calculations from a survey based on the total number of Nigerian adults who paid bribes to public officials while seeking public services from these institutions. However, from the above report, 46.4% of adults were subject to bribe payments to police officers, and they are the most corrupt public institution in Nigeria, with prosecutors 33.0%, Judges 31%, Tax/ revenue officers 27.3%, Customs officers 26.5% and Public utility officers 22.4%. From this survey, the *United Nations Office of Drugs and Crime* (2017), has noted that "corruption takes place across several sectors of the public administration and that certain public officials have a disproportionate impact on the daily lives of Nigerians", and this calls for serious behavioural concern (see, *United Nations and National Bureau of Statistics*, 2017, 8). In another survey by Transparency International (2019), using the same methodology for data collection, the Nigerian police maintain the highest with 69%, followed by members of parliament with 60%, while Judges have a 51% rate of bribery cases

(Transparency International, 2019). Comparing the 2017 survey with that of 2019 shows that corruption in public institutions is on the increase, which calls for serious concern and, as such, needs a re-evaluation of our institutional moral/political behaviour.

ii The Effect on Individual Behaviour/ High increase in Mortality Rates

Similarly, political corruption can cause serious behavioural issues in society, especially in the individual's moral and political behaviours. For example, due to political corruption in Nigeria, there is a high rate of bribery in public offices, police extortion, port congestion, ghost workers syndrome, and electoral irregularities. The idea is that the individuals in Nigeria have come to accept political corruption as a moral norm because the government has failed to provide serious anti-corruption policies that would address corruption. Therefore, the culture of paying bribes is increasing and has become one of the survival strategies by most Nigerians to access public facilities. Data from the *Global Corruption Barometer* as reported by Transparency International (2019) indicates the increasing rate of bribery by citizens who make use of public services in Nigeria. About 44% per cent of Nigerians surveyed in 2019 report having paid a bribe to access public service, compared to around 43% in 2015 (see Transparency International, 2019). Perhaps bribery has been normalized and seen as the right conduct in every sector of government.

In addition, this endemic culture of corruption poses a serious challenge to the mortality rate in Nigeria. With many struggling to survive, these behaviours generally also lead to slowly decreasing levels of mortality rates, especially among people in the grassroots living below the required poverty line. The implication is that those who could not afford money may not be able to bribe public officials to access medical facilities, thereby having a very little chance to survive their illness. For instance, according to Macro Trends, (2020), "the death rate for Nigeria in 2020 is 11.577 deaths per 1000 people; in 2019 was 11.771 deaths per 1000 people, a 1.62% decline from 2018. The death rate for Nigeria in 2018 was 11.965 deaths per 1000 people, a 2.52% decline from 2017" (see Macro Trends, 2020). The slowly decreasing level of mortality rates as shown above happen because funds that are allocated for medical and research institutions disappear into the thin air without proper auditing and lack of accountability.

Below, I will begin to discuss the institutional possibility of PTEs and will explain how Rawls's thought experiment can be used to respond to political corruption in Nigeria as discussed above.

8.5 The Institutional Possibility of Rawls's Thought Experiment and the Question of Political Corruption in Nigeria.

Before I go further, we have to note that, questions relating to the institutional possibilities of political thought experiments have been thought-provoking. Nenad Miščević in his desiderata of political thought experiments has suggested that imagined situations in political philosophy must be institutionally possible as part of the requirements for thought experiments to be considered successful in political philosophy Meaning, the proposed political arrangement in a thought experiment must also be realistic, such, that it would be implemented in political matters in the real world. For instance, using Rawls's scenario, Miščević (2013a) asked the following question "Can we here in Europe, or in Asia or Africa, or somewhere on the rest of the continents, create a just society?". He thinks that "the proposed political arrangement has to be such that it can be implemented in the actual world starting from the present condition of our society" (Miščević, 2013a, 519). By fulfilling this possibility, Miščević thinks that a political thought experiment can be considered extrinsically realistic. In this chapter, I shall defend political thought experiments, in particular, to explain how Rawls thought experiment would meet this possibility, by so doing I will be providing a framework through which corruption could also be tackled in Nigeria.

This chapter adopts Miščević's definition of institutional possibility to mean those concerns about whether political thought experiments can be used to address issues relating to the political arrangement in our society. As discussed in section 2.4.2, Miščević argues that for thought experiments to be successful in political philosophy "the proposed political arrangement has to be such that it can be implemented in the actual world starting from the present condition of our society (Miščević, 2013a, 519). Given this concern, several questions arising from the institutional possibilities of Rawls's thought experiment are as follows: what can we learn from our discussions (in chapter 8) on the motivational power of Rawls's thought experiment in responding to different political situations in the world? Or to what extent can Rawl's scenario be used as a tool in explaining the potential of political thought experiments to solve the problem of political corruption in Nigeria? This section aims to argue that some political thought experiments can educate their audiences by providing them with certain

features which act as incentives in the fight against corruption in Nigeria. Fundamentally, I shall argue in defence of the following important features which can plausibly be drawn from Rawls's argument on the *original position* as a strategy in fighting corruption in Nigeria, they include as follows: (i) collective action (ii) rule of law and, (iii) press freedom (freedom of speech). I will explain how they can be established in Rawl's thought experiment as we proceed. The question is to what extent can these features from Rawls's thought experiment be used to explain how corruption can be eradicated in the real world? or to what extent can these features help to determine a workable policy for anti-corruption reforms in Nigeria? This thesis will demonstrate how the above features will also help us explain the institutional possibilities of Rawls's thought experiment, in particular, to address how they can be useful as a real-world strategy for policy guidance in the fight against corruption. Below, I will explicate the impacts of these features from Rawls's thought experiment, in particular, why I think they can be considered effective in the response against corruption in Nigeria.

8.5.1 The Effect of Collective Action on Corruption in Nigeria

I will begin with the question, what is the relationship between collective action and corruption? Later, I will discuss how the collective action framework can help to reduce corruption in Nigeria. Using the works of Persson et al (2010), I will explicate the logic of collective action and the principal-agent theories to establish the foundation for our discussion, in particular, relating to the fundamental problems of policies on anti-corruption reforms in Nigeria. I will thereafter cross-examine Rawls's idea of collective action within the logic of collective action theory. Later, I will also demonstrate how this could provide a framework for understanding how political thought experiments could be institutionally possible, in particular, how the collective action framework through the *original position* thought experiment could facilitate anti-corruption reforms through policies that will help to reduce corruption in Nigeria.

What is 'collective action'? Briefly, in this thesis, the idea of *collective action* implies those actions taken together by a set of people (as part of a complementary responsibility) whose motivation is to enhance their common goal that is beneficial to members of that group. Perhaps, we also take a collective action problem to mean whenever any individual within that group fails to work together with others or failed to carry out his/her part of the responsibility to achieve their common goal. In Rawls's explanation, collective action could be likened to those actions that are required in sustaining a collective responsibility/agreement for equal distribution and financing of social goods. However, I aim to prove that the political formation that Rawls thought experiment supports can be institutionally possible because its basic

principles (social goods) are based on collective action, especially one that could be implementable in real-world settings. Rawls acknowledges collective action while stating that "it is the responsibility of every citizen to act collectively and not as isolated individuals in building their state", in particular on issues to do with 'public goods'. He called this an act of 'collective action' (Rawls, 1971, 236). Rawls uses the collective action principle to emphasise certain actions that the state should enforce to bring about cooperation for the distribution of social goods. It is also important to note that Rawls has at some point acknowledged situations where collective action problems might occur, he believes that there are situations where the individual might choose not to be fair because they might think that their effort might not bring about a significant change to the society. For instance, in financing public goods, Rawls talked about the need for taxation. Rawls believes that every citizen must pay their taxes with the belief that others will surely pay them. Thus, where the public is large, including many individuals who are required to pay their taxes, there is always the temptation for each to avoid doing his part (see Rawls, 1971, 236). Rawls believes that the state must enforce some of the binding rules regarding the total compliance of the collective agreement for the society to be considered just. In emphasising what collective action could mean to a state, Rawls in his thought experiment and all through his general discussion of justice believes that "the characteristic features of essential public goods necessitate collective agreement, and firm assurance must be given to all that they will be honoured" (Rawls, 1971, 236). Meaning everyone ought to comply with those collective agreements by playing their part in making society a just one. In doing so, they should believe that others are doing the same so that their interests are all protected without anyone usurping the interest of the other.

However, many critics might think that collective action is irrelevant when responding to the global problems of corruption and in Africa in particular, and therefore, they could think that collective action is not sufficient in combating corruption in Nigeria as discussed above. Some might even ask, why does corruption in Nigeria prevail despite a large number of policies and agencies established by different political administrations/institutions to fight it? It is important to note that one of the reasons that anti-corruption failed in Nigeria is because of the conceptualization of corruption as a 'principal-agent problem'- meaning, they have ignored tackling corruption using the collective action model that suits the peculiarities of the problem of corruption in Nigeria. In responding to the above question, it is imperative to cross-examine our position using these two contending theories of corruption as captured by Persson, A., Rothstein, B., and Jan, T. (2010), to establish the dichotomy between both- to determine the

legitimate way of fighting corruption in Nigeria. The importance of this study is to, first of all, unravel empirical information on why institution fails to address corruption in Africa, and later I will demonstrate how other features which Rawls's thought experiment support could offer legitimate means in the fight against it in Nigeria within the logic of collective action framework. The study by Persson et al (2010) reveals how the principal-agent theory has been used in response to the corruption problems in Africa, perhaps the main reason why various anti-corruption reforms have continued to fail in the continent. It also provides empirical details of why corruption in Africa should be considered as a collective action problem.

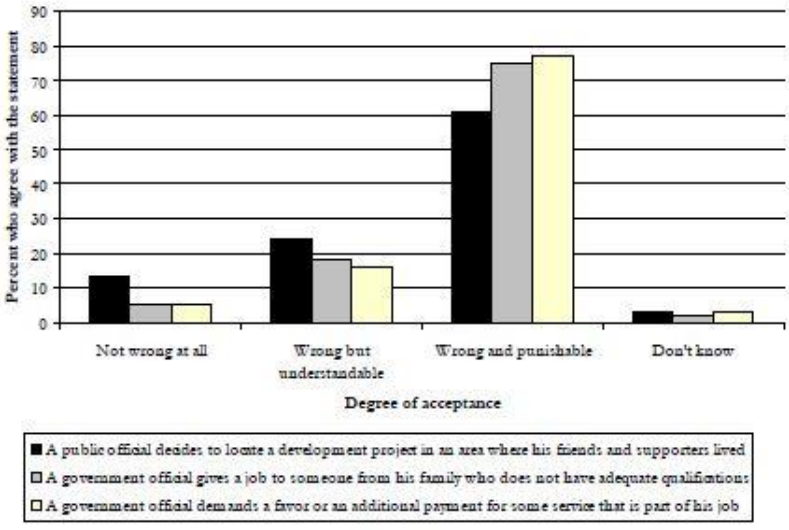
In the study titled "The Failure of Anti-Corruption Policies: A Theoretical Mischaracterization of the Problem*",* Persson et al (2010) argue their claims via a consideration of two theories of corruption: (i) the principal-agent theory and (ii) the collective agent theory-using anticorruption reforms from two African countries to provide empirical insight into why corruption in African prevails despite a large effort to fight it. According to the study, contemporary anticorruption reforms follow the policy implications originating between *the principal and the agent.* According to the study, *an agent* here is seen as someone who is delegated to act on behalf of the principal. Agents are sometimes assumed to act in the principal's interest, and "*the principal* either in the form of a ruler or citizen, typically assumed to embody the public interest in what Klitgaard (1988) observed as "a highly principled principal" (cited in, Persson et al, 2010). However, in the "principal-agent theory, the principal is supposed to hire an agent to represent his day to day actions, this also includes delegation of decision-making authority. According to this theory, *the principal* is to delegate the task of fighting corruption to another collective body of actors; called *the agents* (see, Persson et al, 2010). This theory situates the analysis of corruption in the interaction and interrelations that exist within and without public bodies and is based on two key principles (i) that a goal conflict exists between so-called principals (who are typically assumed to embody the public interest) and agents (who are assumed to have a preference in favour of corrupt transactions insofar the benefits of such transactions outweigh the cost), and, (ii) that agents have more information than the principals, which results to what the researchers described as "an information asymmetry between the two groups of actors" (see, Persson et al, 2010). For instance, the principal might be 'the ruler' or 'citizens' of the country who delegates government tasks to another body (the agent) for the public interest (see, Persson et al, 2010, 2). In this case, the agent might be (i) a tax collector who has all the taxpayers' information and is supposed to collect tax payments for the government, but instead refused to remit 100% of taxes to the

national treasury; (ii) similarly, the agent might also be the corrupt police who are supposed to arrest corrupt officials but instead take bribes thereby compromising the fight against corruption. In this framework, the study argues that "corruption exists when the agent betrays the principal's interest in the pursuit of his/her personal interest". For example, when the agents acquire some information that they are not ready to disclose to the principals or have private motivations other than the goal of performing their delegated task. In this framework, because "the ruler does not perfectly observe which law enforcers behave honestly since they do not have the relevant information that the agents have" this result is a betrayal of trust that leads to what is described as "information asymmetry" between both parties (see, Persson et al, 2010). According to the study, in situations when "the tax collectors are better informed about the revenue potentials of a particular tax base than is the management of a treasury, this opens up opportunities for bribery" (see, Persson et al, 2010, 4), meaning the efficiency of *the principal-agent* model in the fight against corruption can seriously be questioned.

However, to increase the efficiency of this framework, most countries that adopt it sometimes introduce some monitoring mechanism, through freedom of the press, the independent court system, etc, but due to a lack of transparency and accountability between the 'principals' and 'agents', anti-corruption efforts become ineffective. The problem is, within this framework, there seems to be an absence of actors willing to enforce existing laws by reporting and punishing corrupt behaviour and, as such, act as "principals". For instance, as recognised by the United Nations (UN), Nigeria has several agents which include; the Independent Corrupt Practices and Other Related Offences Commission (ICPC), the Economic and Financial Crimes Commission (EFCC), the Code of Conduct Bureau (CCB), the Bureau of Public Procurement (BPP), the Nigerian Extractive Industries Transparency Initiative (NEITI), the Public Complaints Commission, the Office of the Auditor-General of the Federation, and the Technical Unit on Governance and Anti-Corruption Reforms (TUGAR), are all considered as the anti-corruption agencies in the fight against corruption (See United Nation, 2022). Officials from these agencies are comprised of different members from both private and public sectors, including the police force, judiciary, military and paramilitary forces, etc. (as members) who are supposed to arrest, prosecute and jail corrupt officials but instead take bribes to enrich themselves. According to Persson et al (2010), all of these concerns explain why the principal-agent framework has failed in Africa. The study argues that, due to this problem, anti-corruption reforms in Africa fails, and corruption has become more entrenched along with the efforts to curb it. However, in practice, the study acknowledges

researchers who claim that there is always an overall failure of anticorruption reforms due to a lack of proper implementation of laws. For instance, there are situations where the actors are not willing to enforce existing laws by reporting and punishing corrupt behaviour. The reason has been that according to the study, in Africa "no one is willing to enforce existing laws by reporting and punishing corrupt behaviour as such act as principal" (see Lawson, 2009, Svenson, 2005). The study further claims that in some cases when "the media are exposing corruption, the level of reporting and persecution remains low" (see Persson et al, 2010, 2).

Even though the *principal-agent* framework does not have a significant impact on anti-corruption fights in Africa, the study also shows that there is moral disapproval of corruption by most people in most African countries. For instance, the study cited the Afro barometer (2006) survey on some instances of corruption and people's responses. The following question was put across to people to consider corruption as "not wrong at all" "wrong but understandable", "wrong and punishable" if a public official: 1) decides to locate a development project in an area where his friends and supporters live; 2) gives a job to someone from his family who does not have adequate qualifications, and; 3) demands a favour or an additional payment for some service that is part of his job (cited in Persson, et al, 2010). The result of the findings was summarized in the chart below:



Source: Afrobarometer 2006.

The figure above (as demonstrated in the chart) clearly shows that there is decisive disapproval of corruption in most African countries. However, the study further argues that the principal-agent theory is better in theory than practice, because since the supposed 'agents' are

also corrupt and not acting in the interest of the society but instead pursue their narrow self-interest, anticorruption reforms based on this framework will always fail to yield practical results, as it is the case in Africa (Persson, et al, 2010, 3). The effect of this failure is that Africa has become home to the largest number of "losers of corruption" in form of a constant increase in poverty (Persson, et al, 2010, 6).

Conversely, the study further reveals that the African context of corruption is rather that of a collective action problem. Within the collective action theory, the study believes that this framework does not question the potential of effective monitoring and punishment regimes as means of curbing corruption, rather "they question the underlying assumption that every society holds at least one group of actors willing to act as a principal and as such enforce such regimes" (Persson, et al, 2010). Therefore, it is the case that one cannot assume the presence of a 'principled principal.' willing to hold corrupt officials/agents accountable, such as suggested by the principal-agent framework. Rather, it is based on "the rewards and costs of corruption – and hence the existence of actors willing to enforce reform – should be expected to depend critically on how many other individuals in the same society that are expected to be corrupt" (Persson, et al, 2010, 3). In other words, it is expected that people are going to act corruptly, as long as most people are also corrupt, and the cost of being honest is high since that will not create any impact. Even though people might morally disapprove of corruption (as suggested in the 2006 Afrobarometer survey), the study argues that because of the high cost of being honest, and the intended short benefit, they prefer corrupt behaviour to non-corrupt ones. Meaning they serve the corrupt system rather than undermine it (Persson, et al, 2010, 12). The study argues that corruption in the African context shares the characteristics of a collective action problem. To capture the true character of corruption in Africa (as a collective action problem), empirical findings via interviews with more than 60 correspondents from Kenya and Uganda consisting of journalists, NGOs and officials working for anti-corruption agencies were further accessed. From their responses, actors conceived corruption as a collective action problem. For instance, in Africa, everybody is expected to be corrupt because corruption is seen as a moral norm and everyone has something to gain personally from acting corruptly (Persson, et al, 2010). In addition, since everyone is presumably expected to be corrupt, it is the case that nobody is there to receive reports meaning there is automatically a lack of punishment for corrupt behaviour. The correspondent believed that people also refused to report or challenge corruption because of "the fear of repercussion, together with a feeling of being part of a vicious circle of corruption that nobody alone can afford to break out of"

(Persson, et al, 2010, 12). Also, the interviews suggest that leaders are becoming more corrupt due to greed; they cannot also be questioned due to the weak infrastructural power and low degree of democratic legitimacy. Further claims showed that where there are so many institutions meant to stop corruption, they are only used as a means to buy support from ordinary citizens. This, of course, is the main motivation for people not to report and actively challenge corruption in Africa due to weak institutions. The question is does everyone benefit from corruption in Africa? Using the Kenya and Uganda case study, the study reveals that not every group benefit from corruption, to benefit you must be closer to the top of the hierarchy. Meaning the poor and ordinary citizens who are not part of this hierarchy are the ones suffering the effect of corruption. It is also the case that while leaders at the top are seen to benefit from corruption, ordinary citizens are only involved in it passively (for example by paying a bribe) to get public services. Among this group, corruption in Africa is used to avoid unpredictable institutions, for instance, to avoid trouble from the police or courts people are refusing to take action. The study suggests that the cost of playing fairly, and the fact that the cost of challenging corruption becomes too high for the ordinary citizen, means there is zero chance of having real change.

There are important points that can be deduced from this study, in understanding- the reasons why anti-corruption policies fail in Africa- and providing a way forward for future strategies in the fight against corruption, especially in Nigeria. (i) The predominant use of the principal-agent framework to fight corruption in Africa often results in information asymmetry between the 'principal' and 'agent'. (ii) Since the agent act based on self-interest, betrays the 'principal's interest in the pursuit of his or her self-interest and does not act in the interest of the society, any anti-corruption reform based on the assumption relating to *the principal-agent framework* breaks down. (iii) That the overall failure of anticorruption reforms in Africa is partly because all actors are per definition corrupt. (iv) Anti-corruption reforms in the continent ignore the collective action problem character of corruption., hence, the conceptualization of the problem of corruption in Africa as a principal-agent problem contributes to the failure of anti-corruption reforms in the continent. (v) The study also shows that the lack of proper implementation of anti-corruption reforms also brings a major setback to the fight against corruption in Africa, meaning institutional failures contributes to a great extent to why corruption in Africa is still high despite the effort to stop it. (vi) Although some people clearly understand the negative implications of corruption and morally disapprove of corruption in all manifestations, the lack of ability in taking action in reporting and sanctioning due to

institutional failures has become another source of concern. (vii) Finally, because, the problem of corruption in Africa is a collective problem, it means any solution to it should not end at persecuting corrupt officials but changing actors' beliefs about what is expected from everyone (like trust) in having a fair play with other actors.

8.6 An Overview of Corruption in Nigeria as a Collective Action Problem: A Response from the Perspective of Rawls's Thought Experiment.

Below, I will show how the Nigerian government has adopted the principal-agent framework and will explain briefly the reasons why the corruption problem in Nigeria (as already discussed in this chapter) should be assessed within the logic of a collective action framework as suggested by the study. Drawing from the above study, it can be deduced that the present fight against corruption in Nigeria has continued to fail because it is tied to the principal-agent model. Whereas corruption in Nigeria is a collective action problem, any attempt to use the principal-agent model will lead to a fruitless result. Meanwhile, there are great similarities between the activities of corruption in Uganda and Kenya with that of Nigeria vis-a-vis anti-corruption reforms. To date, the anti-corruption reforms in Nigeria are based on the principal-agent framework as part of the contemporary approach. But, despite the effort of different political administrations in adopting the principal-agent framework, for instance, the Nigerian government has empowered the Economic and Financial Crimes Commission (EFCC) as *the agent* to bring corrupt individuals to book. But despite this corruption has become even more entrenched in step with the effort to curb it. However, as demonstrated previously, the statistic of corruption in Nigeria is on the increment, the 2019 and 2020 *corruption perception index* in Nigeria is ranked 146 and 149 respectively, out of 180 countries surveyed with a score of 20% of 100 (see Transparency International, 2020). Our investigation further shows that the corruption perception index continues to grow in Nigeria because of institutional failures (such as the police, judiciary, the legislative, and the failure of anti-corruption agencies like the EFCC, etc.). It is also the case that both *the agents* and *the principals* who are supposed to fight corruption are being accused of corruption, institutions that are supposed to fight it are now biased, public criticism and protest are becoming a crime, the judiciary is now bribed, the police are protecting them and becoming part of those benefiting from it. By implication, politicians and institutions of government do not live up to the required political standard in carrying out their constitutional responsibilities. There are also reports that corruption in Nigeria has also extended to the electoral process known as

'electoral corruption' (Idada & Uhunmwuangho, 2012). It explains how politicians and electoral officers/voters are involved in vote-buying, and engaging in all forms of electoral manipulations/violence that lead to the killing and maiming of people. It is now the case that *the agents* and *the principals* are now hunting the rest of the citizens they are supposed to protect during electoral processes. By implication, an average Nigerian sees political institutions as benefactors of corruption, and bribery for public services has also become a norm for the rest of the public while the poor are intimidated and could not afford public services. Generally, politicians who embezzle public funds work free without proper litigation in the open court, thereby denying the masses their constitutional rights. The poor are silent because of a lack of financial capacity to seek justice in court, even the few court cases reported lack the will to sanction defaulters. This shows a clear case of institutional failure that needs collective action in order to fight against corruption in Nigeria. Within the above purview, it is evident that corruption in Nigeria happens to be a collective action problem, the principal-agent model as is used in Nigeria in fighting corruption is one of the reasons why the anti-corruption effort in Nigeria does not produce a meaningful result.

The thought-provoking question is, how can we respond to this issue using Rawls's thought experiment? From the above illustration, it can be argued that features *of* Rawls's thought experiment or features of the system of which the thought experiment is part would offer an insight into how political corruption can be addressed in Nigeria. The idea is that so far there is a political problem in Nigeria that is rooted in corruption, particularly crimes like embezzlement of public funds, political racketeering caused by the leadership, and institutional failures, which leads to a lack of trust. Therefore, their action falls short of the collective action plan in the quest for justice in Nigeria. Politicians who are benefiting from corruption need to understand how they could feel if they are in the position of the poor masses (and others who morally disapprove of it) that have been continuously exploited by their corrupt actions. However, with the present cases of corruption in Nigeria, it is the case that the few political elites in governance are using their portfolio to tailor the greater part of the primary social goods for their interest and to the benefit of their loved ones against the interest of the masses. Such corrupt, selfish, and nepotic practices in any manner is a typical example of unfair play that could become an impediment to the fulfilment of one's rational plans, and it is against Rawls's thinking of collective action plan (or basic agreement) at *the original position*. Meanwhile, the real-world impact of this thought experiment is for politicians to become aware of how governments can be just or fair, and individuals can be seen as part of the state where

their collective voices and action needs to complement that of the state in the quest for justice for all. Indeed, Rawls's principles suggest the possibility of building a strong institution for everyone's interest, for instance, it protects equal opportunity which must be granted to everyone, and every corrupt official must be given a fair hearing in the court of law. Perhaps, it should be the case that a judgement that is equal to any crime committed against the collective interest of all can be passed without fear or favour.

Below I will briefly demonstrate the action plan of Rawls's thought experiment within the logic of collective action theory. The aim is to unravel the suitability of the *original position* in tackling the problem of collective action as is the case in the fight against corruption in Nigeria. The argument I am raising here is that Rawls's thought experiment can be used to solve the problem of collective action as posted in the anti-corruption reforms in African countries in general and Nigeria in particular. It can be argued that Rawl's case provides the audience with the needed mental insight to act collectively in responding to crime or injustice which is against everyone's collective interest in the state, such insight can also help everyone to imbibe the culture of collective responsibility to refrain from corrupt behaviours which is an act of injustice. For instance, going by the various claims in Rawls's thought experiment, parties in the *original position* might not see corruption as part of their considered principle in the new society. It can be deduced that they will collectively try as much as possible to make laws to sanction defaulters or create strong institutions where their collective interests can be protected. This of course will be responded to by considering policies devoid of all forms of impartiality in the state. Similarly, it is important to note that a collective agreement based on individual cooperation with others is something that they will do to protect their interest and that of the unborn. Thus, *the veil of ignorance* and the fear of not losing out will become a key motivator for collective agreements and it is the responsibility of all to fight for such rights. This also implies that an individual will be obligated to keep to his/her part of the agreement while believing that others will do the same. However, as suggested by Rawls the parties are rational individuals, meaning they can also understand the risk involved if any of them end up becoming *principals or agents* including taking any role that might likely betray their collective trust. As rational people, they would opt for principles where the individuals (commoners and disadvantaged) can be part of the polity and can contribute meaningfully to the state. Perhaps their right in bringing about revolutionary changes when the need arises can never be denied. By implication, there will be no one that will exploit others' interests, such that no one will be seen as a corrupt *agent* that only acts in his or her self-interest (as is the case in the principal-

agent framework) that will work against their collective goal. As such, no one will have absolute power (as the actor) to tailor the social goods for personal interest. In any case, if they found themselves as an ordinary citizen, they must ensure that the expected decision and the behaviour of the principal must represent their collective interest when they are out of the veil. For instance, they must ensure that anyone in the future who is found to be corrupt will face the law. The implication of this is that in the real world, political leaders in Nigeria with such insight are prone to imbibe a political culture that is devoid of individualism and partiality, especially during the policymaking process against corruption. Although literally, politicians in Nigeria cannot get behind the "veil of ignorance" as suggested in the thought experiment to make fair decisions or create just policies against corruption in the country. But, in a practical sense, they necessarily must attempt to set aside their interests and political position and show the readiness to welcome collective contribution. Of course, no one shall be above the law-immunity clause (law) should be discontinued. They will begin to make laws that represent the people and offer a collective solution to issues, and finally, anyone who has fallen short of their agreement will be sanctioned as permitted by Nigerian laws.

Critics of the collective action framework might say that in a corrupt state, it might be difficult for an individual who is not corrupt to believe that others will keep to the same behaviour, especially in situations when being honest becomes a risk. These are some of the collective action problems that are concerning and can be remedied by countries whose rule of law and press freedom are guaranteed. It is the case that these two principles as proposed by Rawls can guide the few who are not corrupt to advance their right of reporting those who are corrupt and seek justice through litigation against those whose interest falls short of the agreed collective behaviour. This problem will be further explained under the subheading below.

8.6.1 The Effect of Rule of law and Press freedom on Corruption

In the section above, I have explained the suitability of Rawls's *original position* as one of those political thought experiments that can be applied within the framework of collective action theory in tackling corruption. I have also argued that via its institutional possibilities, it can become a tool for responding to the problem of anti-corruption reforms globally and in Nigeria in particular. The question is how then can *the original position* be used to respond to the collective action problem, especially on issues relating to poor implementation, reporting/punishment of corrupt behaviour, impartiality, and greed in Nigeria? And how can this bring us real-world institutional changes to cope with the existing crisis of anti-corruption

reforms in Nigeria? Below, I suggest two important features from Rawls's argument that I believe are capable of creating huge impacts when it comes to the implementation, reporting, and sanctioning strategy as part of the anti-corruption reforms in Nigeria. These two important features include as follows: (i) press freedom (freedom of speech) and (i) the rule of law. Below, I will explicate these features and further provide empirical information to justify my position on how Rawls's thought experiment could serve as policy guidance in the fight against corruption in Nigeria.

Briefly, in Rawls's thought experiment, the culture of impartiality and equal liberty for all is crucial and is part of his argument for the future arrangement for parties at the original position (OP), which is also applicable in today's democratic settings. As part of the liberty for all especially in a fair society, it is important to note that Rawls acknowledges the existence of an "impartial administration of public rules which then is the rule of law that guides and gives form to the system" (See Rawls, 1971, 206). Besides, Rawl believes that the principle of equal liberty must also guarantee freedom of speech as one of the basic requirements for all that must be respected by everyone (See Rawls, 1971, 313). The idea is that the rule of law and freedom of speech is central to any arrangement procedure at the OP, and part of the expected liberty for all. This thesis believes such claims provide sufficient support for the collective action thesis and it is also important for understanding how press freedom and equality before the law in a democratic society can provide policy guidance when addressing political corruption or the issues relating to anticorruption reforms in Nigeria.

However, the culture of impartiality (as regards equality before the law) and freedom of speech (as regards press freedom) suggested by Rawls can be demonstrated in democratic settings when politicians make laws that guarantee the rule of law and free reporting by the press and the general public without intimidation and unwarranted arrest. In Nigeria, there are numerous issues where the rule of law and press freedom have been compromised and human rights seriously violated, this is one of the main reasons why corruption remains high. It is the lack of freedom in reporting that corrupt behaviour especially those committed by politicians in the country is hidden from the people. However, to make things worst, the present political dispensation in Nigeria especially the Buhari administration has promoted laws that are anti-people voices. Some of these laws are enjoyed by the elite against the interest of the masses, for example, the Immunity Bill, the Rural Grazing Area (RUGA) bill, the Social Media Bill of 2020, and the bill that bans Twitter are key examples. This thesis believes that the immunity bill for instance is not in conformity with 'the rule of law' because it lacks equality before the

law. This immunity only protects the political elite from an open trial in a court of law when they are accused of corruption. Also, the RUGA bill which gives a certain group of people the right to seize the native land(s) of other tribes is another clear case of injustice (See Adesoji, 2019). Critics believe that such bills "lack economic sense" and are nepotic, partial, and only provide special privileges to the president and his kinsmen. Of course, the RUGA bill is to promote the 'Fulinization and Islamization' of Nigeria, and it only offers protection to the elite against the will of the masses (See Adesoji, 2019). Similarly, different Social Media Bills have been promoted by the same administration to protect politicians (the ruling party) from public criticism on issues to do with corruption. Currently, Twitter which is a social media platform for easy communication was once banned and there is a general trait that other media platforms shall follow suit. As a result of this, the press and the general public lack freedom of speech that is needed for objective reporting against *the 'principals'* and its '*agents'* who are real perpetrators of corruption. According to Reporters Without Borders (2021), Nigeria is ranked 120 in the 2021 World Press Freedom Index and there are also claims that most Nigerian reporters are always forced to flee after reporting political corruption/crimes (see Reporters Without Borders, 2021). Of course, the press in Nigeria represents the people's voice and is supposed to serve as a "transparent agency, as a determinant of government performance and a detector of corruption" (see Hollyer et al., 2014). In China, the "online anti-corruption" approach has been adopted as one of the strategies in the fight against corruption. Social media in China now provide a channel for public participation in anti-corruption activities (see Yong and Songfeng, 2015). But this seems to be lacking in Nigeria because the Nigerian government is censoring the media, and once suspended Twitter which of course does not provide a secure environment for free reporting. It can be argued that Nigeria has become a haven for corrupt politicians. Therefore, to have revolutionary changes against corruption, freedom of the press must be granted, and public participation must be adopted using social media as one of the mediums for free reporting and monitoring of corrupt behaviours in line with the collective action framework.

Meanwhile, corruption, as we know it is a typical example of the kind of issue that Rawls's idea of just arrangement could correct, perhaps if freedom of the press is guaranteed then the collective voices of people can be heard and corruption will be reduced in Nigeria. By implication, I believe that lack of press freedom is a situation that denies the people their fundamental right of seeking justice and providing public opinion using the media in the fight against corruption. Besides, it is the case that such denial cannot create a good environment for accountability, transparency, and political development. However, Rawls's case predicts that

when Nigerian leaders perform their constitutional duties with the behaviour of an impartial spectator, they are bound to be altruistic, meaning there will be no room for denying others their freedom of speech or any other biased practices in the system as a result of the altruistic mindset. Also, it limits them from considering how policies would benefit them and their immediate families against the wishes of all including the poor masses who voted them to power. I believe that a law that has an altruistic motive can be impartial and can be universally applied to everyone- this can become a game-changer in the fight against corruption in Nigeria. The idea is that Rawls's thought experiment can provide cognitive awareness that supports the importance of the rule of law and freedom of speech, in particular, press freedom in governance which I believe is crucial in the anti-corruption reforms particularly in Nigeria and Africa in general. This approach will guarantee an improved level of transparency and expansion of information disclosure to ensure accountability in Nigeria.

8.7 Responses to Possible Criticisms

There are various ways this chapter can be criticised. Some critics might argue that most thought experiments are done from an individual level, as such how can they be used to achieve a collective solution? For instance, Rawls thought experiment on the original position can only be conducted from an individual point of view, therefore the idea of its usefulness from the collective point of view can be questioned. In responding to this concern, it is important to note that thought experiments in political philosophy are arguments. As an argument, they can be used to provide evidence that offers a solution for collective interest. For instance, in Rawls's case, it can be deduced that the thought experiment gives us an argument that some laws are unjust as such to achieve a just society there is a need for a fair redistribution of social goods. This chapter suggests that Rawls's thought experiment can be used to offer a collective solution to the issues of corruption in Nigeria. One way to achieve this aim is that in Nigeria, the idea of corruption is heavily linked to the immunity clause which is considered as an unjust law because it favours only the politicians against the interest of the majority of the citizens. Rawls's argument suggests that such self-interest law is unjust, hence the need for a fair redistribution of social goods. It suggests that the immunity clause is unjust and should be replaced by a law that protects collective interest through which everybody's interests are guaranteed given his two principles of justice. This of course suggests that even though a political thought experiment can be done from the individual point of view, it can create a significant impact given the idea of collective interest emanating from Rawls's thought experiment.

Critics of my view might also think that freedom of the press and the rule of law as suggested in Rawls' thought experiment as part of the collective action response and supported by this thesis cannot guarantee that corruption in Nigeria can be addressed. The thought-provoking question this claim might bring is what are the empirical pieces of evidence that these features 'rule of law' and 'press freedom' can be useful in eradicating political corruption? Or do they play any impact on the mitigation of corruption in a country? There are existing empirical pieces of evidence that explain the crucial role of the rule of law and press freedom on corruption which I think can be useful in justifying my claims. For instance, an empirical investigation conducted by Hamada et al (2019) shows where these features (also known in the work as 'variables') were put to test in democratic settings to study their impacts on corruption. The study involves findings from different study samples explored across 111 countries, in particular, those with complete information on all the required variables, press freedom, corruption, democracy, and rule of law (RL), for a period of 12 years from 2004 to 2015. Based on data availability, the study represents all six inhabited continents, democratic and authoritarian contexts, as well as developed and developing countries.

For press freedom, data from the freedom of the press index published annually by Freedom House (via http://www.freedomhouse.org) was used, while the Corruption Perception Index for each country was from Transparency International. In the study, all three major variables, democracy, press freedom, and the rule of law, were all interacted with corruption to generate their impacts. To examine how differences in press freedom influence corruption, the study divided the data into three models: "the first includes only 'free countries', the second includes the 'partly free countries' and the third includes the 'not free countries'," as identified by Freedom House (see Hamada et al, 2019, 312).

The objectives of the study are based on five major hypotheses, which include as follows:

(i) *H1*. Countries with higher levels of press freedom have lower levels of corruption than countries with partial or no press freedom.

(ii) *H2*. The negative correlation between democracy and corruption will be stronger in countries with higher levels of press freedom than in countries with partial or no press freedom.

(iii) *H3*. The negative correlation between RL and corruption will be stronger in countries with higher levels of press freedom than in countries with partial or no press freedom.

(iv) *H4*. Press freedom, democracy, and RL are complements rather than substitutes in their fight against corruption.

(v) *H5*. The interaction effect of press freedom with RL is stronger than the interaction effect of press freedom with democracy in reducing corruption (see Hamada et al, 2019, 312).

In this study, several regression models were used to test the hypotheses, including one independent variable, two independent and all the variables. The effects of Democracy and the Rule of Law on corruption per the status of Press Freedom, cover from 2004–2015. According to the study, the rule of law is the strongest variable in explaining variation in corruption (85%), followed by press freedom (41.1%) and then democracy (15.2%). The estimates from the single variable model are statistically significant, the model indicates that the increase by one standard deviation in the rule of law, democracy, and press freedom triggers a decrease of .92, .39, and .64, respectively, in the corruption level. (see Hamada et al, 2019, 311). The researchers also believe that there is a significant mean difference between countries with a free press and those with partial or no press freedom. As seen from the means plot in Figure 1 (below). The implication according to the study is that press freedom is a key component of the general level of corruption in a country.

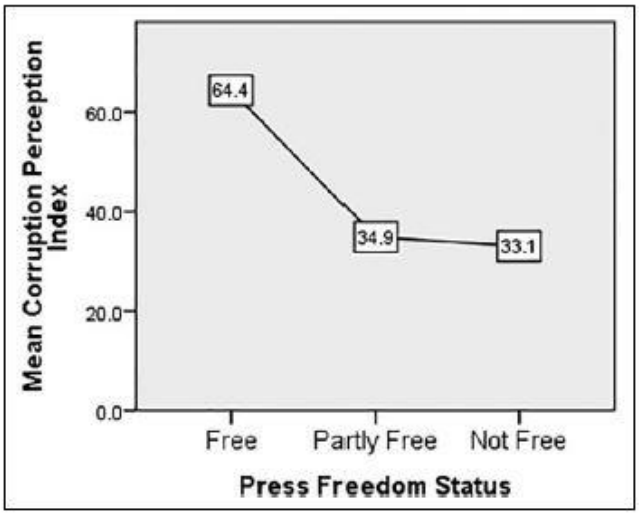


Figure 1: Press freedom status. (Retrieved from Hamada et al, 2019)

Figure 1 above explains how the changes in the state of press freedom have happened in tandem with changes in corruption. By implication "the increase in the level of press freedom is accompanied by a notable decline in corruption and vice versa" (see Hamada et al, 2019,

312). Furthermore, using the 'interaction model', the study also examines the interactive effect of the variables with corruption in a country. The aim is to establish the effect of democracy and rule of law on corruption as per the status of press freedom. Findings from this interaction were used to justify a significant relationship between higher levels of press freedom and lower levels of corruption. Evidence from this study also shows that democracy prevents corruption only in countries that enjoy press freedom, and that rule of law reduces corruption whether the press is free or unfree. The study further concludes that based on the statistical analysis, "it is evident that a free press is an important institution that increases the probability of the detection of corruption and magnifies the cost corrupt officials and institutions are likely to incur, and therefore acts as a deterrent to corruption" (Hamada et al, 2019, 316). Thus, press freedom, rule of law, and democracy are complements rather than substitutes in the fight against corruption.

Generally, the Hamada et al, study suggests to us the inevitable role of rule of law, and freedom of speech through press freedom in the fight against corruption. There are important points to deduce from this study about the anti-corruption reforms in Nigeria vis-a-vis the institutional possibility of political thought experiments. There is a strong need for the implementation of policies that advance press freedom and rule of law in Nigeria to satisfy the collective action framework. Such is needed to guarantee an all-inclusive government where everyone's voice can be heard, where issues can be addressed and collective action can be taken in the fight against corruption. However, I argue in this thesis that Rawls's thought experiment could help political leaders become fully aware of these variables, such that, press freedom, rule of law can become practicable in Nigeria's democracy. Through such awareness, politicians can begin to engage in political conversation on policies, and laws that will lead to the direct implementation of these variables in order to reduce corruption in Nigeria. It can be argued that since these variables have been proven as factors that help reduce or check corruption in a country, then political thought experiments can be institutionally possible because they provide cognitive awareness on how these variables can be legitimately implemented in a free and fair society as suggested by Rawls's thought experiment.

8.8 Summary and Conclusion

In summary, this chapter adopts the institutional possibility of Rawls's thought experiment to provide an alternative approach in response to the problem of political corruption in Nigeria. It also claims that the problem of political corruption can be addressed with a clear

understanding and implementation of the rule of law, and freedom of speech via press freedom as part of the collective action plan in anti-corruption reforms as suggested by Rawls's thought experiment. The chapter agrees with Persson et al (2010) proposal that the problem of anticorruption reforms needs *collective action* against the predominant *principal-agent framework* which has failed to produce a meaningful result in the present anti-corruption campaign in Africa including Nigeria. I have argued that Rawls's thought experiment provides a framework for understanding how the collective action problem can be resolved especially when it comes to issues like impartiality, greed, and other corrupt behaviour from politicians who are sometimes seen as the principals and agents. While Rawls's case provides the idea that it would be good if people cooperate, the Hamada et al, study suggests that in the real world, the rule of law and freedom of speech can be used to protect collective cooperation for all individuals in a state. Second, the institutional possibility of Rawls's thought experiment suggests that in the real world, political corruption in Nigeria can be eradicated if we can enforce and implement policies that enhance collective action plans and all-inclusive roles against the particularistic ones as projected by the principal-agent framework. By implication to have an equality of opportunity for all, the full actualization and implementation of the rule of law and press freedom as part of the fundamental right for all is the way forward. By implication, 'the immunity clause' that protects politicians in Nigeria from a fair trial in court must also be jettisoned because it discourages the fight against corruption in the country. I have also suggested that everyone should have the freedom of speech including press freedom in reporting corrupt individuals for proper legal investigations. Finally, the chapter suggests that policy-making procedures in Nigeria must corroborate the idea of the rule of law and press freedom, such that politicians who are the real perpetrators of corruption can be checked by objective rules and standards that are equal to the same rules and standards that apply to all as suggested by Rawls's thought experiment. This of course is the possible way of eradicating corruption in Nigeria, because, through it, transparency and accountability will be part of policy-making procedures as demonstrated throughout this chapter.

CHAPTER NINE

9. SUMMARY AND CONCLUSION

9.1 Objectives Revisited

I have aimed to defend the use of thought experiments in moral and political philosophy that has been largely dismissed by critics. I have observed that to date, intuitions elicited by political thought experiments have not received the required recognition they deserve, perhaps they have been misconstrued mostly by empirically minded scholars. For instance, the use of unrealistic hypotheticals in political philosophy has brought the attention of different scholars claiming that intuitions elicited by this kind of hypotheticals in the discipline are generally problematic. Thaler (2017) in particular, believes that the present state of hypotheticals used during thought experiments must involve realistic scenarios to evoke a trustworthy intuition that can be relevant in combating real-world political dilemmas. He thinks that imaginary scenarios via thought experimentation must explain political situations as they occur in the real world as we know it for them to be considered trustworthy.

Similarly, Klampfer (2017) also believes that thought experiments in moral and political philosophy provide a divisive response meaning their intuitions are prone to be problematic, especially when used as prima facie evidence for normative truth. He thinks that for any thought experiments to be considered trustworthy or relevant in the discipline, hypotheticals must be able to generate a non-divisive conclusion, such that their intuition can be considered stable for political action. Other empirically minded scholars also believe that intuitions from the discipline are prone to be affected by personal biases, our cultural values, gender, language, and the ethnic backgrounds of the reader, etc (see Tobia, 2018; Costa, 2014; Friedman, 2012; Stich, 2004; Costa, 2014). For these critics, nothing good can come out of philosophical intuitions whether via thought experiments or by any other means because they are prone to be manipulated by these factors, thus, the use of it for normative or political/moral reasoning should be rejected.

I have argued that such claims that result in the rejection of the use of intuitions in political philosophy, must not be taken seriously. Hence, these critics fail to demonstrate the proper understanding of political thought experiments, in particular, the role of argument during thought experimentation. I have argued that some thought experiments in political philosophy can be considered trustworthy in solving normative issues which also include those political issues as they occur in the real world as we know it. In responding to the general criticisms of

thought experiments in political philosophy, in chapter two, I endorsed Nenad Miščević's account of political thought experiments, to trace the various uses of hypotheticals in political philosophy. Here, I tried to offer a substantive account of the role of thought experiments that were used by noble political thinkers. Using Miščević's proposal, I have explained the various desiderata of successful thought experiments including the desiderata of relevance, detailed coverage, realisticity, and stability. The desiderata are used to explain the various components of successful political thought experiments, and their normative functions, including their motivational and institutional possibilities. Although Miščević's account/ desiderata fail to respond to the concerns that were raised by critics, which shows the gap in the present literature on political thought experiments (PTEs) as clearly summarised in (Chapter Three).

However, the explanation of the existing relationship between political thought experiments and arguments supports our position that some thought experiments in political philosophy can be considered trustworthy because of the valid nature of their arguments. I have argued that those whose premises necessitate their conclusion are capable of eliciting a stable intuition. The thesis also suggests that readers of hypotheticals in political theory need to carefully follow the premises of each political thought experiment to attain a non-divisive conclusion, from where stable intuitions can be evoked. Also, with the problem involving the use of unrealistic hypotheticals in political thought experiments (PTEs) as presented by Thaler, in Chapter Four, I offer an alternative approach to PTEs by providing a satisfactory explanation of how *the reflective equilibrium method* might be necessary for understanding legitimate use of unrealistic thought experiments with the real-world. Firstly, I have shown that unrealistic hypotheticals in political philosophy are important because it helps the thought experimenter gain the right intuitions against all empirical barriers. Perhaps since some thought experiments are also arguments as discussed in chapters four and five, the implication is that logically, premises and conclusions need not be realistic for any argument to be considered valid. Therefore, such use of unrealistic hypotheticals in political philosophy can be justified because political thought experiments (PTEs) are arguments and can be reconstructed as one without any epistemic loss. I further claim that *the reflective equilibrium method* can be used in political philosophy to connect the unrealistic imaginary claims of PTEs with their intuitive equivalence in the real world. I suggest that readers of hypothetical scenarios need not worry about the real-world situations of hypotheticals during thought experimentation, because they offer no epistemic significance to the argument of the thought experiment. I argue that readers of PTEs

only need to unravel the logic of comparison between the imaginary conclusion and certain aspects of reality for which they can both be compared in the real world as we know it.

In responding to the various criticisms of political thought experiments, I have made room for another approach to political thought experiments called *'the arguments view of PTEs'* (in Chapter Five) to explain the relationship between political thought experiments and arguments. This approach is an extension of *'the argument view'* drawn from John Norton's (2002) account of scientific thought experiments. Norton believes that scientific thought experiments are dressed-up arguments. He also thinks that we can as well reconstruct imaginary claims to premises and a conclusion similar to a standard argument in logic without any epistemic loss. I endorse the *argument view* especially his 'reconstruction thesis' which is used in this thesis to explain that political thought experiments can as well be considered as arguments, and can also be reconstructed as standard arguments, in line with the 'reconstruction thesis' of John Norton. In addition to conveying the existing relationship between arguments and political thought experiments, John Rawl's thought experiments on *the Original Position* have been carefully examined. In virtue of the findings and the reconstruction of the same thought experiment, I have concluded that Rawl's thought experiment presents a clear picture of a political thought experiment with a satisfactory structure of a valid argument without losing the intuitive functions. I argued that readers of hypotheticals, therefore, stand to benefit more from the existing features of political thought experiments, because intuitions that are elicited by it can be considered stable and trustworthy if their arguments are valid and if the imaginary scenario can be compared to real-world cases by virtue of their logical equivalence. Similarly, in Chapter Six, I use the 'argument view' of PTEs to respond to the issues that were raised by both Klampfer (concerning the divisive nature of PTEs) and other criticisms against the efficacy of intuitions as elicited via PTEs. The chapter submits that readers can all arrive at a stable conclusion via thought experimentation only when they can establish the argumentative structure of hypotheticals and reconstruct them as arguments. In this thesis, the reconstructed thought experiments (the violinist case and the pond and envelope case) all suggest that readers must understand the exact premises of complex imaginary claims which must be followed accordingly in order to produce a non-divisive response. Perhaps, new premises that were not initially mentioned by the thought experimental claims must be left out to avoid arriving at a divisive or extrapolated conclusion.

Also, I have studied the motivational and institutional possibilities of political thought experiments to address the problem of how to make political thought experiments relevant to the present political realities in Nigeria as captured in chapters seven, and eight. The issue concerning how thought experiments can be useful in responding to the problem of political corruption in present-day Nigeria is another important focus of this thesis that I have also tried to address in these last two chapters.

In Chapter seven, I explained how political thought experiments could reasonably hope to motivate their readers. This finding also explains the motivational possibility of thought experiments. The idea is that political thought experiments could hope to shape the moral and political behaviour of their audience via their motivational power. To defend these claims, I explained that there is a close connection between political thought experiments and narratives. This approach drew data from McVey's (2020) empirical studies on the relationship/impacts of narratives and philosophical arguments in behavioural changes to provide a good strategy for presenting political thought experiments as both arguments as well as narratives. I argued that like stories, PTEs are capable of facilitating narrative transport, meaning hypotheticals with these features can result in political motivations during the thought process. I also argued that some political thought experiments (like *the Original Position* by John Rawls) can produce narrative transport, which means it can be hoped to motivate its audience a bit, which could result in shaping their normative judgements and moral behaviours in politics. I also contend that if we redesign political thought experiments to meet the right features of narrative transport, there is strong hope that they can motivate the reader, such that they could guide our moral judgments.

In chapter eight, I present the general concern of political corruption using Nigeria as a case study, to understand how they affect society which includes, infrastructural development, health, and economic and social welfare of the Nigerian people. I then attempted to account for the impact of political corruption, which explain the effect of corruption on both individual and institutional behaviours. From the various empirical facts involving political corruption in Nigeria, I observed that many people in political authority in Nigeria use political power to direct social goods for personal needs. I have also noted that, in the present situation of Nigeria, corruption has become a norm leading to a high level of poverty and an increase in mortality rates due to poor infrastructures and lack of good medical care. I have conceded that corruption in Nigeria is increasing due to a lack of suitable policies to address it. Currently, addressing the

problem of corruption in Nigeria has failed because policies that are implemented in responding to it are surrounded by the principal-agent theory. Thus, because of the failures of this approach, corruption in Nigeria has been on the increase. In addition, chapter nine also discusses the importance of some features (such as collective action, rule of law and freedom of speech) that can be deduced from Rawls's thought experiment in the fight against corruption in Nigeria. First and foremost, I draw our attention to the institutional possibility of political thought experiments, to discuss how these features from Rawls's thought experiment can provide an alternative framework for responding to the issues of political corruption in Nigeria. I have pointed out that one of the key functions of Rawls's thought experiment as deduced from chapter seven is that it provides useful features as a framework for institutional development. With the aid of features such as collective action, rule of law, and freedom of speech, policymakers in Nigeria and world politics would gain useful information through this thought experiment to address their political issues. This idea has been developed further to explain the collective action framework as the best approach in policy reforms/implementation to address the problem of corruption in Nigeria, as against the principal-agent approach that seems to yield no good result in the fight against corruption in Nigeria. These findings suggest that a close study of PTEs by politicians could help them understand how to respond to political corruption in Nigeria through policy changes and reforms, firstly by implementing the collective action approach and the enactment of objective laws as suggested by Rawls's thought experiment.

9.2 Conclusion

As we finally draw to the finish, there are various implications of my thesis. My defence of political thought experiments (PTEs) is not mistaken, it assumes that the use of hypotheticals in political philosophy should not be rejected, especially in responding to real-world issues in politics, because of their enormous intuitive functions. Besides, thought experiments are used legitimately in political philosophy in different directions: (i) to enhance our knowledge about politics and (ii) to make us acquire new knowledge (iii) it provides awareness that enhances our political way of life, etc. Because of the role of arguments and the reflective equilibrium method, I have argued that thought experiments in political philosophy should not be limited to the use of realistic hypotheticals as critics have claimed. Going through this thesis, I hope that I have contributed some novel insights to our understanding of the usefulness of unrealistic hypotheticals in political thought experiments. More specifically, I would like to believe that this thesis enhances our understanding of the enormous role of intuitions elicited via political

thought experiments, in particular how to draw their real-world equivalence in responding to moral and political claims in the real world using the method of reflective equilibrium. Perhaps, the role of this method does not deny the application of *the argument view*, both approaches can work simultaneously in unravelling the efficacy of PTEs. The argument view as used in this thesis is of huge epistemic impact. It explains the relationship between PTEs and arguments, the implication is that the use of hypotheticals in political theory is guided by logical rules and operates under that same law to guide readers in accessing thought experiments with little or zero controversies Part of this thesis suggests that most political thought experiments also have a clear relationship with narratives, an idea that connects our understanding of political thought experiments to stories. I believe that the narrative approach to thought experimenting does not in any way invalidate the argument view of thought experiments because a thought experiment can both be argument and narrative. This framework is used in this thesis to explain how PTEs that contain both narrative and valid arguments would work side by side especially, in providing legitimate reasons on how readers can gain moral and political motivations without being biased via thought experimentation. However, as I bring this thesis to an end, it is important to note that PTEs can be of practical use. As argued throughout this thesis, the use of unrealistic thought experiments can create an action-guiding role if and only if their moral intuitions would be used to establish a parallel with the real world via the equilibrium test. To defend this claim, I have explored the method of *reflective equilibrium,* which explains the relationship between unrealistic hypotheticals and the real world. This method has been used to show how unrealistic thought experiments can be justified, particularly how such justification can be used to enable an understanding of the real world. These findings allow us to argue that unrealistic thought experiments would help to explain facts as they occur in the real world for the purpose of providing action guidance. In addition, most political thought experiments would provide useful information in responding to political issues, as demonstrated in this thesis, using the issue of political corruption in Nigeria as a case study. The thesis submits that important factors drawn from Rawls's thought experiment can act as policy guidance and are capable of providing guidelines concerning what policy can best be used in responding to the corruption issues in Nigeria.

Bibliography

Abbas J. (2019) '69% of Nigerians living below poverty line harms the economy, says Oxfam'. Retrieved from https://www.dailytrust.com.ng/69-of-nigerians-living-below-poverty-line-harms-economy-says-oxfam.html. Retrieved on 16/05/2020

Abramson CI. (2013) "Problems of teaching the behaviourist perspective in the cognitive revolution". *Behav Sci (Basel).* 3(1):55-71.

Adesoji, B, (2019) "Why the RUGA settlement initiative never made economic sense" cited in https://nairametrics.com/2019/07/04/why-the-ruga-settlement-initiative-never-made-economic-sense/. Retrieved on 25/04/2021

Arcangeli, M, (2010) 'Imagination in Thought Experimentation: Sketching a Cognitive Approach to Thought Experiments. In: Magnani, L., Carnielli, W., Pizzi, C. (eds.) *Model-Based Reasoning in Science and Technology*. SCI, vol. 314, 571–587.

Arthur, R. (2018) "Thought Experiments in Newton and Leibniz" in *The Routledge Companion to Thought Experiments* James Brown et al (Eds): New York, Routledge Publishers. 111-127

Bartlett, F. C. (1928), "Types of Imagination," *Journal of Philosophical Studies*., 78-85.

BBC News (2012), *Former Nigeria governor James Ibori jailed for 13 years*. Retrieved from https://www.bbc.co.uk/news/world-africa-17739388. Retrieved on 09/03/2020

BBC News, (2018) *'Abacha loot' to be given to poor Nigerians'*. Retrieved from https://www.bbc.co.uk/news/world-africa-44654931. Retrieved on 09/03/2020.

Beck, S & de Wijze S.(2015) "Interrogating the Ticking Bomb Scenario: Reassessing the Thought Experiment". *International Journal of Applied Philosophy* 29, 1:53-70.

Binzak Azeez (2018), 'Immunity Clause under the Nigerian Constitution: Matters Arising' Retrieved from https://www.pmnewsnigeria.com/2018/12/01/immunity-clause-under-the-nigerian-constitution-matters-arising/. Retrieved on 11/03/2020.

Bishop, M.A. (1999) "Why thought experiments are not arguments," *Philosophy of Science* 66: 534-541

Bourget, D., & Chalmers, D. (2013). What do philosophers believe? Cited in https://philpapers.org/archive/BOUWDP, Accessed Nov. 2019.

Brendel, E. (2018) “Argument view: Are thought experiments mere picturesque arguments?” in *The Routledge Companion to Thought Experiments* James Brown et al (Eds): New York, Routledge Publishers. 281-292

Brown, J. (1991) *The Laboratory of the Mind: Thought Experiments in the Natural Sciences.* London: Routledge.

Brown, J. (1992) "Why Empiricism Won't Work." pp. 271-279 in Hull *et al.* (1992).

Brown, J. (1993) "Author's Response" (to Norton, 1993) *Metascience, 3* (new series), 38-40.

Brown, J. (manuscript) "Peeking into Plato's Heaven," Prepared for Philosophy of Science Association Biennial Meeting, 2002, Milwaukee, Wisconsin.

Brownlee, K. & Stemplowska, Z. (2017). Thought experiments. in A. Blau (Ed.), *Methods in Analytical Political Theory*. Cambridge: Cambridge University Press. 21-45

Brun, G.. (2018) “Thought experiments in ethics” in *The Routledge Companion to Thought Experiments* James Brown et al (Eds): New York, Routledge Publishers. 195-210

Buckwalter, W. and Stich, S. (2014) “Gender and philosophical intuition,” in *Experimental Philosophy*, vol. 2, edited by J. Knobe and S. Nichols, Oxford: Oxford University Press.

Bulirsch R and J Stoer (1966) “Numerical treatment of ordinary differential equations by extrapolation methods” *Numer Math, 8:1-13*

Bulirsch R and J Stoer (1966) “Numerical treatment of ordinary differential equations by extrapolation methods” *Numer Math, 8:1-13*

Butler, Judith. (2007) Torture and the Ethics of Photography. *Environment and Planning D: Society and Space* 25:951–966.

Buzzuni, Macro (2010) “Empirical thought experiments: A Transcendental-Operational View” *Epistemologia* XXXIII: 5-26

Chalmers D.J. (2002). "Does Conceivability Entail Possibility?" In *Conceivability and Possibility*, ed. T. Gendler and J. Hawthorne. Oxford: Oxford University Press, 145-200

Chilton, Paul (2004) *Analysing Political Discourse: Theory and Practice*. London: Routledge Publishers.

Clare Murphy (2008) "Is it wrong to select a deaf embryo?" cited in http://news.bbc.co.uk/1/hi/health/7287508.stm Accessed 14/01/2019

Cole, D. (2009). "Section 2.3 The Chinese Nation". The Chinese Room Argument. *Stanford Encyclopedia of Philosophy*.

Cooper, R. (2005) Thought Experiment. *Metaphilosophy* 36(3): 328-347

Costa, A. et al (2014) "Your morals depend on language," *PLoS ONE* 9: e94842.

Dennett, Daniel C. (1980), 'The Milk of Human Intentionality', *Behavioral and Brain Sciences*, 3 (4), 428-430.

Dennett, Daniel C. (2013), *Intuition Pumps and Other Tools for Thinking* London: Allen Lane.

Dirac, P.A.M. (1947). *The Principles of Quantum Mechanics* (2nd edition). Clarendon Press.

Einstein, A (1949) Albert Einstein, Philosopher-Scientist, Paul A Schilpp (Ed), *Library of Living Philosophers*, Evanston

Einstein, A. (1931) *Einstein on Cosmic Religion with other opinions and Aphorisms*. Mineola Network: Dover Publications

Ekmekci, P. E., & Arda, B. (2015). Enhancing John Rawls's Theory of Justice to cover health and social determinants of health. *Acta Bioethica*, 21(2), 227-236

Ellen et al, (1997) *Self-Interest.* Cambridge: University of Cambridge Press

Fabamise Sesan (2017) "Constitutional immunity clause and the fight against corruption in Nigeria" in *Journal of Sustainable Development Law and Policy*. Cited in https://www.ajol.info/index.php/jsdlp/article/view/163328. Retrieved on 26/11/2021

Felana, Femi, (2019) "Falana accuses US, UK of nurturing corruption in Nigeria" https://www.sunnewsonline.com/falana-accuses-us-uk-nurturing-corruption-nigeria/

Frappie, J. et al (2013). “Introduction” In J. R. Brown, M. Frappier, & L. Meynell (Eds.), *Thought experiments in science, philosophy, and the arts*, New York, NY: Routledge, 177– 190

Galileo, G. 1914. *Discourses and Mathematical Demonstrations Relating to Two New Sciences(1638)*, Edited by Crew, H. and de Salvio, A. London: MacMillian.

Gan, Stepen (2018) “History and Key Concepts of Behavioral Psychology” Cited in https://www.verywellmind.com/behavioral-psychology-4157183. Retrieved on 5/12/2019

Gendler, T.S. (2000). *Thought Experiment: On the Powers and Limits of Imaginary Cases*. New York: Garland

Gilbert, et al (2000) “Thought experiments in science education: Potential and current realization”, *International Journal of Science Education* 22: 265-283

Goodman, N. (1965) *Fact, Fiction and Forecast*, Bobbs-Merrill, Indianapolis

Grundman, T. (2018) “Platonism and the a priori in thought experiments” in *The Routledge Companion to Thought Experiments.* James Brown et al (Eds): New York, Routledge Publishers. 293- 308

Gun University (2021) “Gun sales and manufacturing statisitics 2021” cited in https://gununiversity.com/gun-sales-stats/#2020. Retrieved on 13/11/2021

Gunjita, Saxena (2018, September 17) “*Galileo's experiment on falling objects*” [Video file] https://www.youtube.com/watch?v=gewDxLEryno

Häggqvist, S. (1996) *Thought Experiments in Philosophy*. *Stockholm Studies in Philosophy 18*. Stockholm: Almqvist & Wiksel International

Hamada, B. I., et al (2019). “Press Freedom and Corruption: An Examination of the Relationship.” *Global Media and Communication* 15 (3): 303–21

Hawkins Stephen (1988) *A brief history of time* New York: Buntam

Hess, Karl (2005) *Einstein was Right!.* Singapore: Pan Stanford Publishing

Hume, D. 1968: Treatise of Human Nature. Oxford: Clarendon Press. Hume, D. 1963: *An Enquiry Concerning Human Understanding*. La Salle: Open Court Press

Idada, W., & Uhunmwuangho, S. O. (2012). Problems of democratic governance in Nigeria: The way forward. *Journal of Sociology and Social Anthropology, 3*, 49-54.

Ijewereme, O. B. (2015). Anatomy of corruption in the Nigerian public sector: Theoretical perspectives and some empirical explanations. *SAGE Open*, 5(2). doi:10.1177/2158244015581188

Joel Feinberg, (1985). *The moral limits of the criminal Law: Offense to Others*. New York: Oxford University Press

Katerina Ierodiakonou (2018) "The Triple Life of thought Experiment in Ancient Philosophy" in *The Routledge Companion to Thought Experiments* James Brown et al. (Eds): New York, Routledge Publishers. 31-43

Klampfer, F. (2017) "The False Promise of Thought-Experimentation in Moral and Political Philosophy" in *Thought Experiments Between Nature and Society for Nenad Miscevic* Borstner M et al (Eds): Cambridge Scholars Publishing 328-348

Klampfer, F. (2018) "Moral Thought-Experiments, Intuitions, and Heuristics". *Croatian Journal of Philosophy* 52:133-160

Kukathas C. and Pettit P. (1990) *Rawls: A theory of Justice and its critics*. Oxford: Polity Press

Lederman, Norman G. (et al), (2013) Nature of Science and Scientific Inquiry as Contexts for the Learning of Science and Achievement of Scientific Literacy in *International Journal of Education in Mathematics, Science and Technology vol* 1, Number 3, 138-147

Liu, E. H.-L. (2019) ''Fundamental methods of numerical extrapolation with applications,'' *MIT Open Courseware, Massachusetts Inst. Technol.*, Cambridge, MA, USA, Tech. Rep., 2006. Accessed: Oct. 28, 2019. [Online]

Lyons David (1975) "Nature and Soundness of the Contract and Coherence Arguments" in *Reading Rawls: Critical Studies of A Theory of Justice*. Norman Daniel (Ed): Oxford: Blackwell Mott Ltd.

Machery, E. (2011). Thought experiments and philosophical knowledge. *Metaphilosophy,* 42, 191–214

Macro Trends (2020) 'Nigeria Death Rate 1950-2020' https://www.macrotrends.net/countries/NGA/nigeria/death-rate

Malami (2020) 'We have no record of how $5bn Abacha loot was spent' Retrieved from https://www.premiumtimesng.com/news/top-news/381938-we-have-no-record-of-how-5bn-abacha-loot-was-spent-agf-malami.html. Retrieved on 16/05/2020

Matulovic, Miomir (2017) "Miscevic, Mental Models, and Thought Experiments in Political Philosophy" in *Thought Experiments between Nature and Society: A Festschrift for Nenad Miscevic* Borstner Bojan et al (Eds): Cambridge Scholar Publishing. 354-369

McAllister James (2018) "Historicism and Cross-Cultural Comparison" in *The Routledge Companion to Thought Experiments* James Brown et al (Eds): New York, Routledge Publishers. 425-438

McAllister, J.W. (2004) "Thought Experiments and the belief in phenomena," *Philosophy of Science* 71: 1164-1175

McMahan, Jeff. "Torture and Method in Moral Philosophy." In *Torture, Law, and War,* edited by Scott Anderson and Martha Nussbaum, 1-18. Chicago: University of Chicago Press, 2015. Retrieved: http://jeffersonmcmahan.com/wp-content/uploads/2012/11/Torture-and-Method-in-Moral-Philosophy.pdf

McMahan, Jeff. 2002. *The Ethics of Killing: Problems at the Margins of Life.* New York: Oxford University Press.

McMahan, Jeff. 2008. Torture in Principle and Practice. *Public Affairs Quarterly* 22:91–108.

McMahan, Jeff. 2009. *Killing in War*. Oxford: Clarendon Press

McMahan, Jeff. forthcoming. Torture and Method in Moral Philosophy. in *Torture and the Rule of Law,* edited by Scott Anderson and Martha Nussbaum. Chicago: Chicago University Press.

McVey Christopher (2020) The Moral Importance of Narrative: A Philosophical Analysis of Narrative Transport. A PhD thesis submitted to the University of California

Riverside. Cited in https://escholarship.org/uc/item/0674x8bv. Retrieved on 31/08/2020

Merriam-Webster FAQ. Retrieved 2019-10-24.

Miščević N. (2013b) "Political Thought experiments from Plato to Rawks" in *Thought Experiment in Philosophy, Science, and Arts* James Brown et al (Eds): New York, Routledge Publishers. 191-206

Miščević N. (2018) "Thought experiments in Political Philosophy" in *The Routledge Companion to Thought Experiment* James Brown et al (Eds): New York, Routledge Publishers. 153-171

Miščević, N. (2012). Plato's Republic as a Political Thought Experiment. *Croatian Journal of Philosophy* 12 (2): 153-165

Miščević, N. (2013a) "In search of the reason and the right- Rousseau's Social contract as a thought experiment" *Acta Analytica* 28: 509-526

Miščević, N. (2013c) "A Hierarchy of Armchairs: Gerald Gaus on Political Thought Experiments" *European Journal of Analytic Philosophy* 9: 52-63

Miščević, Nenad. 2013. Political Thought Experiments from Plato to Rawls. Pp. 191–206 in *Thought Experiments in Science, Philosophy,* and the Arts, edited by Melanie Frappier, Letitia Meynell, and Robert Brown. New York: Routledge.

Mondragon, D.I. (2017) "Evaluating the Cognitive Science of thought experiments" in *International Journal of the Historiography of science.* 3, 68-76.

Mongabay (2020) "Deforestation: facts, figures, and pictures" Retrieved from https://rainforests.mongabay.com/deforestation/. Retrieved on 12/11/2021

Moon, B.M., Hoffman, R.R., Novak, J.D. & Cañas, J.J. (2011). *Applied Concept mapping: Capturing, analyzing, and organizing knowledge*. New York: CRC Press.

Moue A. S. et al (2006), "Tracing the Development of Thought Experiments in the Philosophy of Natural Sciences" *Journal for General Philosophy of Science*, Vol. 37, No. 1, 61-75

Neisser, U. (1967). *Cognitive Psychology*. Englewood Cliffs: Prentice-Hall.

Neisser, U. (1976). *Cognition and reality: Principles and implications of cognitive psychology*. New York: Freeman

Neisser, U. (2003). Cognitive psychology. In, *The history of psychology: Fundamental questions* (pp. 447–466). New York, NY US: Oxford University Press

Nersessian, N. (2018) "Cognitive science, mental modelling, and thought experiments" in *The Routledge Companion to Thought Experiments* James Brown et al (Eds): New York, Routledge Publishers. 309-326

Newton, I. (1999) *The Principia: Mathematical Principles of Mathematical Philosophy*, edited and translated by I. B. Cohen and A. Whitman, Berkeley: University of California Press.

Norton, J. D. (1991) "Thought Experiments in Einstein's Work." in: *Thought Experiment in Philosophy and Science* Horowitz and Massey (Eds), Rowman and Littlefield, Savage, MD, pp129-148.

Norton, J. D. (1993) "Seeing the Laws of Nature," (Review of Brown, 1991) *Metascience, 3* (new series), 33-40.

Norton, J. D. (1996) "Are Thought Experiments Just What You Thought?" *Canadian Journal of Philosophy, 26*, pp. 333-66.

Norton, J. D. (2004). "Einstein's Investigations of Galilean covariant electrodynamics prior to 1905". *Archive for History of Exact Sciences* 59: 45–105

Norton, J. D. (2004a). Why thought experiments do not transcend empiricism. In C. Hitchcock (Ed.), *Contemporary Debates In Philosophy Of Science* (pp. 44–66). Oxford, UK: Blackwell Publishing.

Norton, J. D. (2004b). On thought experiments: Is there more to the argument? *Philosophy of Science*, 71(5), 1139–1151.

Norton, R. (2006, November 4). *How to train a cat to operate a light switch* [Video file]. Retrieved from http://www.youtube.com/watch?v=Vja83KLQXZs

Novak, J.D. (2010). Learning, Creating, and Using Knowledge: Concept Maps as Facilitative Tools in Schools and Corporations (2nd Ed.) NY: Routledge.

Obasanjo, O (2019) "Nepotism is corruption, Obasanjo tells presidency" Retrieved from https://guardian.ng/news/nepotism-is-corruption-obasanjo-tells-presidency/ Retrieved on 25/03/2020

Office of National Statistics(2021) Adult smoking habits in the UK 2019. Retrieved from https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthandlifeexpectancies/bulletins/adultsmokinghabitsingreatbritain/2019. Retrieved on 12/11/202.

Persson, Anna, Bo Rothstein and Jan Teorell (2010) The Failure of Anti-Corruption Policies: A Theoretical Mischaracterization of the Problem. QoG Working Paper 2010:19. Gothenburg: University of Gothenburg Quality of Government Institute.

Popper, Karl (1968). The Logic of Scientific Discovery, New York: Harper & Row Publishers.

Prezenski S., Brechmann A., Wolff S., Russwinkel N.(2017) "A cognitive modelling approach to strategy formation in dynamic decision making" *Frontiers in Psychology*, Art. no. 1335.

Pryor, Jim (2006) Thought-Experiments and Counter-Examples. Cited in http://www.jimpryor.net/teaching/vocab/analyses.html. Retrieved on 25/01/2022.

Rawad, S. (2018) "The function and limit of Galileo's falling bodies thought experiments: Absolute weight, specific weight, and the medium's resistance" in *Croatian Journal of Philosophy* Vol. 17, 50, 37-58

Rawls John, (1972) A *Theory of Justice* Cambridge, (Massachusetts: Harvard University Press, 1972).

Reichenbach, H. (1938) "On Probability and Induction," Philosophy of Science 5, no. 1 21-45.

Reiss, J. (2013). "Genealogical thought experiments in economics" In J. R. Brown, M. Frappier, & L. Meynell (Eds.), *Thought experiments in science, philosophy, and the arts*, New York, NY: Routledge, 177– 190

Rescher, N. (2005) *What if?* New Jersey: Transaction Publishers

Rescher, Nicholas. (1991) "Thought Experiments in Presocratic philosophy*"* in *Thought Experiments in Science and Philosophy,* Tamara Horowitz and Gerald Massey (eds). *31-41*

Robinson-Riegler et al (2008) *Cognitive Psychology: Applying the Science of the Mind.* Boston: Pearson Education, Inc.

Sachs, Albie (2013, June 24) "Four Tales of Terrorism (Torture, Law, and War Keynote Address)" (Video file). Retrieved at https://www.youtube.com/watch?v=t7QLnGp4CWc (Accessed: 26/03/2019)

Schabas, M. (2018) "Thought experiments in Economics" in *The Routledge Companion to Thought Experiments* James Brown et al (Eds): New York, Routledge Publishers. 171-182

Schickore, Jutta & Friedrich (2006). *Revisiting Discovery and Justification. Historical and Philosophical perspective on the context distinction,* New York: Springer,

Schlaepfer, G.& Weber M. (2018) "Thought experiment in Biology" in *The Routledge Companion to Thought Experiment*s James Brown et al (Eds): New York, Routledge Publishers. 243-256.

Schwitzgebel, E. (2020) Philosophy Contest: Write a Philosophical Argument That Convinces Research Participants to Donate to Charity. Cited in http://schwitzsplinters.blogspot.com/2019/10/philosophy-contest-write-philosophical.html. Retrieved on 27/03/2021

Seo M., Chang H. (2013) Context of Discovery and Context of Justification. In: Gunstone R. (eds) *Encyclopedia of Science Education*. Springer, Dordrecht

Shettima, Y. (2017) "Buhari Biased in Anti-corruption Fight – Arewa Youths". Retrieved from https://www.concise.ng/2017/10/26/arewa-youths-accuse-buhari-of-bias-in-fight-against-corruption/. Retrieved on 25/03/2020.

Shoemaker, David (2015) "Thought experiments in philosophy" Retrieved from https://blog.oup.com/2015/04/thought-experiments-philosophy/

Singer, Peter (1997) The Drowning Child and the Expanding Circle. In: *New Internationalist*, April 1997.

Singer, Peter. (1972) "Famine, Affluence, and Morality" *Philosophy and Public Affairs* 1: 229-243

Singer, Peter. (2003) 'The why and how of effective altruism' Cited from https://www.ted.com/talks/peter_singer_the_why_and_how_of_effective_altruism#t-1022356. Retrieved on 7/04/2021

Solso, R. (2001) *Cognitive Psychology*. Boston: Pearson Education, Inc.

Sontag Susan (2004) 'Regarding The Torture Of Others' Retrieved from https://www.nytimes.com/2004/05/23/magazine/regarding-the-torture-of-others.html Retrieved on 16/05/2020

Sorensen, R. A. (1992) *Thought Experiments*. Oxford: Oxford University Press

Stanford Encyclopedia of Philosophy (2022) **'**Doctrine of Double Effect**' cited in** https://plato.stanford.edu/entries/double-effect/ Retrieved on 30/3/2022.

Starmans, C. and Friedman, O. (2012) "The folk conception of knowledge" *Cognition* 124: 272–283.

Steinhoff, Uwe. 2010. Defusing the Ticking Social Bomb Argument: The Right to Self-Defensive Torture. Global Dialogue 12. Available at http://hub.hku.hk/handle/10722/90486, accessed 14 January 2019.

Steinhoff, Uwe. 2013. *On the Ethics of Torture*. Albany: the State University of New York Press

Sternberg R. (2009) *Cognitive Psychology. Beltmont:* Berky Cross

Stitch, S. (1988) "Reflective Equilibrium, Analytic Epistemology and the Problem of Cognitive Diversity" *Synthese* 74(3): 391-413

Stuart, Michael T. (2014). Imagination: A sine qua non of science. *Croatian Journal of Philosophy* 17, (49) 9-32.

Stuart, Micheal (2018) “How thought experiments increase understanding” in *The Routledge Companion to Thought Experiments* James Brown et al (Eds): New York, Routledge Publishers. 526-544.

Stuart, Micheal. (2012) Brown’s Laboratory of the Mind. Spontaneous Generation: *A Journal for the History and Philosophy of Science*, Vol. 6. 1. 237-241

Thagard P. (2014) “Thought experiments considered harmful” in *Perspectives on Science* 22, 288–305

Thaler M. (2016) Unhinged Frames: Assessing Thought Experiments in Normative Political Theory. *British Journal of Political Science*. Advance online publication. Retrieved from https://doi.org/10.1017/S0007123416000259.

Thomson, Judith Jarvis. (1971), “A Defence of Abortion”, *Philosophy and Public Affairs* 1: 47-66

Thomson, Judith Jarvis. (1985). “The Trolley Problem”, *Yale Law Journal* 94: 1395-1415

Tobia, K. et al (2018) “Intuitions and its critics” in *The Routledge Companion to Thought Experiments.* James Brown et al (Eds): New York, Routledge Publishers. 369-384

Tocqueville, Alexis de (2018) “The Fairness Principle: How the Veil of Ignorance Helps Test Fairness” in https://fs.blog/2017/10/veil-ignorance/

Transparency International (2019) ‘Citizens’ Views And Experiences Of Corruption. Retrieved from https://www.transparency.org/files/content/pages/2019_GCB_Africa.pdf. Retrieved on 18/03/2020

Transparency International (2019), *Corruption Perception Index* Retrieved from https://www.transparency.org/files/content/pages/2019_CPI_Report_EN.pdf. Retrieved on 16/05/2020.

United Nation & National Bureau of Statistics (2017) Corruption in Nigeria Bribery: public experience and response. Retrieved from https://drive.google.com/file/d/0B6jj-ulM0cLrOXFpMDh1Q1l1bm8/view. Retrieved on 05/03/2020.

United Nation (2022) Preventive Anti-Corruption Body Or Bodies. Retrieved from https://www.unodc.org/documents/corruption/WG-Prevention/Art_6_Preventive_anti-corruption_bodies/Nigeria.pdf. Retrieved on 07/03/2022.

University of Cambridge (2012) 'Morality Prevents Crime', cited in http://www.cam.ac.uk/research/news/morality-prevents-crime. Retrieved on 20/05/2021

Uzan, J-P and Leclercq, B (2008). The Natural Laws of the Universe: Understanding Fundamental Constants.*Chichester:* Springer *Press*

Van Woudenberg, Rene. (2006) "Introduction" *Metaphilosophy* 37, 151-161

Weisberg, Micheal. (2007) "Three kinds of idealization," *Journal of Philosophy* 104: 639–659. https://www.sas.upenn.edu/~weisberg//documents/threekindsfinal.pdf

Wilkes Kathleen (2003) *Real People: Personal Identity without thought experiments*, Oxford: Clarendon Press.

Williamson T. (2007). *The Philosophy of Philosophy*. Malden, Mass.: Blackwell

Wilson, L. (1954). A study of opinions related to the nature of science and its purpose in society. *Science Education*, 38(2), 236-242.

Wimmel, Hermann (1992). *Quantum Physics & Observed Reality: A Critical Interpretation of Quantum Mechanics*. World Scientific.