DIRECTORATE OF DISTANCE EDUCATION

UNIVERSITY OF NORTH BENGAL

MASTER OF ARTS- PHILOSOPHY

SEMESTER -IV

EXISTENTIALISM

ELECTIVE 405

BLOCK-2

FOREWORD

The Self-Learning Material (SLM) is written with the aim of providing simple and organized study content to all the learners. The SLMs are prepared on the framework of being mutually cohesive, internally consistent and structured as per the university's syllabi. It is a humble attempt to give glimpses of the various approaches and dimensions to the topic of study and to kindle the learner's interest to the subject

We have tried to put together information from various sources into this book that has been written in an engaging style with interesting and relevant examples. It introduces you to the insights of subject concepts and theories and presents them in a way that is easy to understand and comprehend.

We always believe in continuous improvement and would periodically update the content in the very interest of the learners. It may be added that despite enormous efforts and coordination, there is every possibility for some omission or inadequacy in few areas or topics, which would definitely be rectified in future.

We hope you enjoy learning from this book and the experience truly enrich your learning and help you to advance in your career and future endeavours.

EXISTENTIALISM

BLOCK-1

BLOCK-2

BLOCK 2: EXISTENTIALISM

Introduction to the Block 2

Unit 8 deals with Decision and choice. Decision making is the process of choosing actions that are directed towards the resolution

Unit 9 deals with Existence. Existence raises deep and important problems in metaphysics, philosophy of language, and philosophical logic.

Unit 10 deals with The necessity of existence. It is commonly accepted that there are two sorts of existent entities: those that exist but could have failed to exist, and those that could not have failed to exist.

Unit 11 deals with The Definition of Death. Examples of standards for human death are the traditional cardiopulmonary standard and the whole-brain standard. Insofar as clinical tests are primarily a medical concern, the present entry will not address them.

Unit 12 deals with the concept of Death is life's ending. To clarify death further, we will need to say a bit about the nature of life, and ask whether life can be suspended or restored, and how it relates to our continued existence.

Unit 13 deals with Temporality. In philosophy, temporality is traditionally the linear progression of past, present, and future. However, some modern-century philosophers have interpreted temporality in ways other than this linear manner.

Unit 14 deals with Authentic and non-authentic. The term 'authentic' is used either in the strong sense of being "of undisputed origin or authorship", or in a weaker sense of being "faithful to an original" or a "reliable, accurate representation".

UNIT 8: DECISION AND CHOICE

STRUCTURE

8.0   OBJECTIVES

After going through this unit, you should be able to:

- To discuss the importance and process of decision making,
- To discuss the models of decision making,
- To explain the relativity of creativity and decision making,
- To discuss common errors in decision making.

8.1   INTRODUCTION

Decision making is the process of choosing actions that are directed towards the resolution. It can be defined as "the selection from among alternatives of a course of action: it is at the core of planning". The decision making process can be carried out either by individuals acting alone or by groups. There are several models and theories which are developed to explain decision making and how effectively you can make

a decision. Decision making is a process of selection from a set of alternative courses of action which is thought to fulfill the objectives of the decision problem more satisfactorily than others. Decision making is an essential part of every function of management. In the words of Peter F. Drucker, "Whatever a manager does, he does through decision making." When we talk of teachers it can be seen that a teacher is continuously involved in decision making whether it is regarding school activities or related student centered activities etc. Knowingly or unknowingly a teacher is always at decision making. Decision making involves thinking and deciding before doing and so is inherent in every activity. That is the reason decision making is often called the "essence" of managing. No one can survive without effective decision making. Some of the decisions may be of a routine type and repetitive in nature and some may be strategic in nature which may require a lot of systematic and scientific analysis. In the educational sector, a teacher is always a decision makedTeachers are expected to make decisions that affect the growth and development of the students in their care.

Decision theory is concerned with the reasoning underlying an agent's choices, whether this is a mundane choice between taking the bus or getting a taxi, or a more far-reaching choice about whether to pursue a demanding political career. (Note that "agent" here stands for an entity, usually an individual person, that is capable of deliberation and action.) Standard thinking is that what an agent does on any given occasion is completely determined by her beliefs and desires/values, but this is not uncontroversial, as will be noted below. In any case, decision theory is as much a theory of beliefs, desires and other relevant attitudes as it is a theory of choice; what matters is how these various attitudes (call them "preference attitudes") cohere together.

The focus of this entry is normative decision theory. That is, the main question of interest is what criteria an agent's preference attitudes should satisfy in any generic circumstances. This amounts to a minimal account of rationality, one that sets aside more substantial questions about appropriate values and preferences, and reasonable beliefs, given the situation at hand.

8.2 SIGNIFICANCE OF DECISION MAKING

Decision making is important for organizational effectiveness because of its central role in the overall process of directing and controlling the behaviour of organizational members. Decisions are made that cover the setting of goals, strategic planning, organizational design, personnel actions, and individual and group actions. Besides its organizational effect, however, decision making also has an individual effect. The quality of a decision has a bearing on his or her professional success and sense of satisfaction. So studying decision making is important from both an organizational and an individual perspective. Another major reason for studying decisions is to enable us to make better quality decisions than we do presently. This point must be emphasized strongly because the quality of our decisions is often much poorer than we realize. Selective perception tends to bias the information we use in making decisions and our attitudes and values influence how we interpret that information. Drives for consistency lead to oversimplified interpretations. Our willingness to attribute positive outcomes to ourselves. (e.g., taking credit for good decisions) and to attribute negative outcomes to forces outside our control makes us remember the results of decisions in a personally favourable light. All these forces degrade our decisions. and at the same time, limit our understanding of the decision making problem, Besides being unaware of our human limitations in the decision making process, we are often unaware of the methods that can be used to increase our decision effectiveness. Very little training that emphasizes the actual decision making process is available either inside or outside organizations. In most cases experience is our guide and while experience can be a good teacher, it can be misleading as well. In many cases we may learn the wrong way to do something or we may obtain information that is actually irrelevant for the quality of the decision. In order to increase our effectiveness in decision making, we must first understand the decision making process. Decision making and planning are deeply interlinked. The determination of objectives, policies, programmes, strategies, etc. involves decision making. The most outstanding quality of a teacher to be

successful is hisher ability to make sound decisions, A teacher may be in a situation where he/she has to make up hisher mind quickly on certain matters, It is not correct to say that he/she has to make ~pur of the moment decisions all the tima, While taking many decisions, he/she gets enough time Pot ~areffil fact fiadin~, merlysis sf alternative8 and eheiee sf the best alternative. Decision making is a human process. When a teacher decides, he/she chooses a course which he/she thinks is the best.

8.3 DECISION MAKING PROCESS

The basic characteristics of decision making are as follows: It is the process of choosing a course of action from among the alternative courses of action. It is a human process involving to a great extent the application of intellectual abilities. It is the end process preceded by deliberation and reasoning. It is always related to the environment. A decision may be taken in a particular set of circumstances and another in a different set of circumstances. It involves a time dimension and a time lag. It always has a curpose. Keeping this in view, there may just be a decision to not to decide. It involves all actions like defining the problem and probing and analyzing the various alternatives which take place before a final choice is made. The decision making process includes the following components: The decision maker. The decision problem. The environment in which the decision is to be made. The objectives of the decision maker. The alternative courses of, action. The outcome expected from various alternatives. The final choice of the alternative. The stages of decision making are indicated in Figure 8.1

Flg.8.l: Stages of Decision making

Flg.8.2: Decision Making Elements

The final leg of decision making goals and objective. The seeand istags is problem reeagnitian, Here the deeision:m&er has to be alert to know what

is happening and also to recognize the discrepancies which exist. During the third stage, the decision-maker must evaluate the discrepancy whether it is an important one or not. Next, it has to be found out how the problem occurred i.e., 'information-search phase'. This stage is crucial but least handled well. In the next stage 'course of action' must be explored i.e., number of alternatives to be explored. This is the 'alternative-generation' phase. Next comes the evaluation of alternatives that is the 'choice phase'. Here the pros and cons of each alternative have to be thought about before taking a decision which is known as choice of action. The last phase of the process involves the implementation and evaluation of the decision.

8.4   TYPES OF DECISIONS

Decisions may be classified into five major types. These are: Organisational and personal decisions Routine and strategic decisions Policy and operating decisions Programmed and non-programmed decisions Individual and group decisions Let us discuss each type in brief.

(i)   Organizational and personal decisions: Personal decisions are those decisions that cannot be delegated to others. These decisions are meant only to achieve personal goals. Organisational decisions are those decisions that are taken to achieve organizational goals. For example you want to solve food habits related problems of your students. Advising them to take i nutritious food becomes a personal decision. As a teacher you adopt different kinds of teaching methods so that your students are able to I understand science and mathematics better. These are for organizational goals because good performance enhances the credibility of the school.

Activity 1 Illustrate with reference to your school on what occasions you had to follow organizational or individual decisions.

(ii)   Routine and strategic decisions: Routine decisions are those which are repetitive in nature. For example, certain established rules, procedures and policies are to be followed. You might

have experienced that when a teacher goes on leave another teacher who is free at that time has to engage the class. This is a routine decision. 'Strategic' decisions are those decisions which have to be deliberated upon in depth. For example, highlighting the characteristics of the school, before giving an advertisement for admissions, can bring more revenue to the school.

Activity 2 Cite an example with reference ta your school regding a strategic decision.

(iii) Policy and operating decisions: Policy decisions are those decisions which are taken at the higher level. For example, fixing pay scales for teachers. Operating decisions are those decisions which mean procedure of execution of the policy made. For example, how to disburse the arrears accumulated to a teacher (e.g. calculations).

Activity 3 What do you mean by operating' decision?

(iv) Programmed and non-programmed decisions: Non-programmed decisions are those decisions which are unstructured. For example, if a child is often absent, the class teacher can analyse the reasons for hislher absenteeism from the information provided by the child and then advise as to how to recoup with the situation. Whereas programmed decisions are of routine type and repetitive in nature. For example, when should children take their breakfast, hlnch etc.

Activity 4 differentiate between programmed and non-programmed decisions citing appropriate examples.

(v) Individual and group decisions: A decision taken by an individual in the organisation is known as 'individual' decision, where autocratic style of functioning prevails. For example, if

only the principal takes a decision without the participation of teachers, it is an individual decision. 'Group' decisions are collective decisions which are taken by a committee with a proper representation. For example, decisions taken collectively by parents, teachers and principal for the welfare of students.

Activity 5 Describe a situation where individual and group decisions have to be taken?

Some other types of decisions: Decisions can also be classified on the basis of dimensionality i.e., complexity of the problem and certainty of outcome of following the decision. These are described below:

Mechanistic decisions: Mechanistic decision is routine and repetitive in nature where the outcomes are known. For example, if a child misbehaves in the class, the teacher raises voice to control it.

Analytical decisions: In this type of decision one has to analyse the situation and take a decision. For example, if students are 30t performing well in science, the reasons have to be explored. It cx be because of the teacher or the method of teaching science, lab-facilities provided, etc.

Adaptive decisions: In this kind of decisions outcomes are not hewn and offen unpredictable, It variety from Bibation to Bibation, For example, a decisions taken by without prior experience of the outcome.

Activity 6 Cite examples regarding the above types of decisions with reference to your school or a school you are familiar with.

Check Your Progress 1
Note: a) Use the space provided for your answer
b) Check your answers with those provided at the end of the unit

1.   Discuss the Significance of Decision making.

_____

_____

_____

2. Write about the Decision making Process.

_____

_____

_____

3. What are the Types of Decisions?

_____

_____

_____

8.5   MODELS OF DECISION MAKING

The following are important decision making models which enable us to know more about decision making: Contingency model Economic man model Administrative man model Social man model Let us discuss each model in brief.

(i) Contingency model: Beach and Mitchell (1978) felt that the decision maker uses one of three general types of decision strategies: aided analytic, unaided analytic, and no analytic. The aided analytic strategy employs some sort of formal model or formula, or an aid such as a checklist. An unaided analytic strategy is one in which the decision maker is very systematic in his or her approach to the problem and perhaps follows some sort of model, but does it all in his or her head. Thinking of all the pros and cons for each alternative or trying to imagine the consequences of each action would fall in this category. Finally there is the category of no analytic strategy. Here the decision maker chooses by habit or uses some simple rule of thumb ("nothing ventured, nothing gained" or "better safe than sorry") to make the choice. Decision Making Which strategy is to be selected depends on the personal characteristic of the decision maker and the demands of the task. The underlying
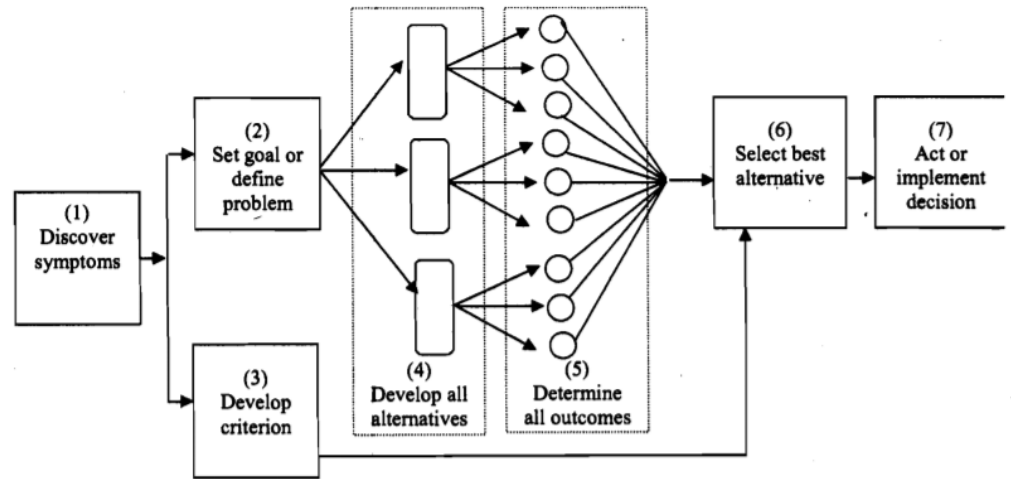
assumption of this model is that a person will choose a strategy that requires the least amount of time and effort to reach a satisfactory decision. The more analytic a strategy, the more time and effort are required to use it. Since aided analytic techniques take the most effort and analysis, the use of such techniques requires that

1) the individual should have the personal characteristics necessary to employ them (e.g., knowledge, ability, and motivation) and

2) such techniques are demanded by the characteristics of the decision problem. The characteristics of the problem are divided into two groups: the decision problem itself and the decision environment. The model suggests that as the decision problem becomes less familiar and more ambiguous, complex, and unstable, the decision maker will use more time and analysis (more analytic strategies) ta reduce the uncertainty caused by these factors. However, this process continues only up to a point. When the uncertainty due to these factors becomes too great, the decision maker is likely to return to a simpler rule. The reason is that when there is an extremely high degree of uncertainty in the decision problem, the potential gains of a more accurate analytic decision are small and are often far outweighed by the cost (e.g., time and effort) required to arrive at that decision. The decision environment is composed of four factors. The model suggests that more analytic strategies will be .selected when decisions are not reversible and very important, and when the decision maker is personally accountable. Also, analytic.procedures are more likely to be used where there are no time or money constraints.

(ii) Economic man model: In this model, it is believed that man is completely rational in taking decisions. It is accepted that man takes decisions based on the best alternatives available. An ecologic model of decision making is given in Figure 8.3.

Flg.8.3: An Econologlc Model of Declsion maldng Source: Behling and Schriesheim, 1976

(iii) Administrative man model: This model assumes that though people would like to have best solution, they settle for less because the decisions may require more information which they may not possess. Thus, there is a kind of bounded (or limited) rationality in decisions. The following three steps are involved in the process of this model. Sequential attention to alternative solutions: In this step, all the alternatives are identified and evaluated one at a time. If one of the alternatives fails then the next alternative is considered Use of heuristics: A heuristic is a rule which guides the search for alternatives into areas that have a high probability for yielding satisfactory solutions. In this step if the previous solution was working then a similar set of alternatives are used in that situation. Satisfying: Here the alternatives which are workable are found to be satisfying. I A bounded rationality model of decision making is explained in Figure 8.4

Fig.8.4: A Bounded Rationality Model of Decision Making Source: Behling and Schriesheim, 1976
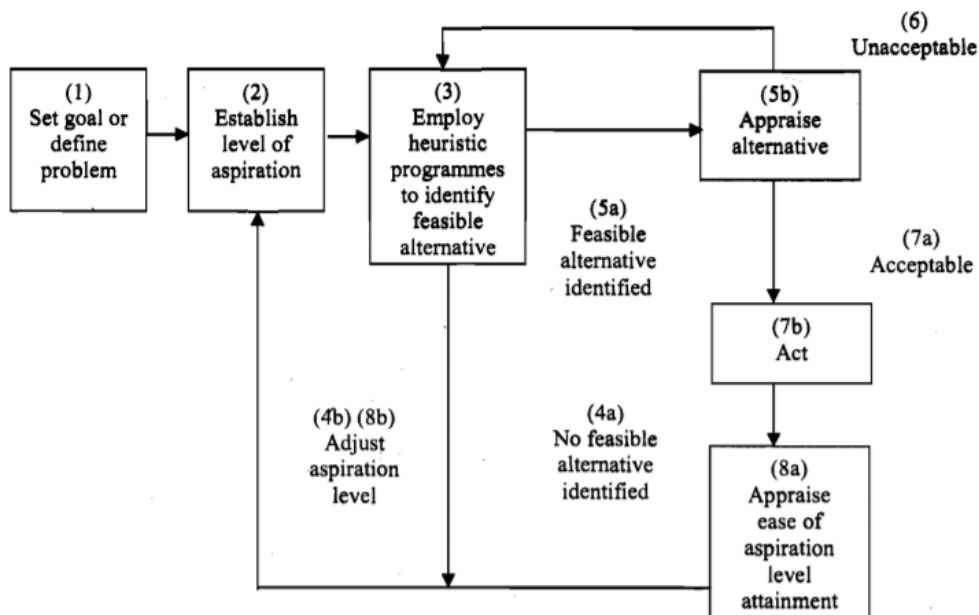
(i) Social man model: This model was developed by the classical psychologists. This model feels that man being a social animal is subjected to social pressures and influences. Here the decisions are taken under the following conditions:

- Certainty: Because of certainty, accurate decisions can be taken.
- Uncertainty and risk: Several decisions are taken under conditions of risk.

Identification of Alternatives

In order to generate alternatives three main processes are generally used. These are brainstorming, synectics and nominal grouping.

(i)   Brain storming: This is developed by Alex F. Osborn. It is the best technique in stimulating creative thinking. The objective of this method is to produce as many ideas as possible. In this method 'criticism' is prohibited. 'Freewheeling' is welcome. Generating a number of alternatives is the motto. Combination and improvement are sought. This method does have limitations. They are time consuming and costly. Care should be taken to select group members who are familiar with the

problem to be considered (e.g. Parent - Teacher Association meetings).

(ii)   Synectics: Here members are selected from different backgrounds and training. The leader poses the problem in such a way that the members deviate from traditional ways of thinking. Various methods employed include role playing, use of analogies, paradoxes, metaphors and other thought provoking exercises. This is a widely used method and though it has limitations like brain storming, it is very useful for complex and technical problems.

Activity 7 Illustrate the above mentioned decision making technique and situation where it can be used.

(iii)   Nominal grouping

It means group in name only. This model is useful when it requires a high degree of innovation and idea generation. Here the search process is proactive rather than reactive. It is also time consuming and costly.

8.6   CREATIVITY AND DECISION MAKING

Creativity involves a novel combination of ideas which must have theoretical or social value or make an emotional impact on other people. Creative decisions and the quality of such decisions is influenced by many factors. It would depend upon the quality of the information input and any prejudices introduced because of our perceptual processes and cognitive constraints. In addition to the outside factors, the characteristics of the decision maker greatly affect the quality of the decision. The primary characteristics are the attitude of the decision maker towards risk that he/she may be facing and the types of social and cultural influences on him/her. Some of the factors and. personal characteristics that have an impact on the decision maker are:

(i)   Information inputs: It is very important to have adequate and accurate information about the situation for decision making; otherwise the quality of the decision will suffer. It must be recognized, however, that an individual has certain mental consmints which limit the amount of information that he/she can adequately handle. Less information is as dangerous as too much information, even though some risk takers and highly authoritative individuals do make decisions on the basis of comparatively less information than more conservative decision makers.

(ii)   Prejudice: Prejudice and bias are introduced by our perceptual processes and may cause us to make ineffective decisions. First of all, the perception is highly selective, which means that we only accept what we want to accept and hence only such type of information filters down to our senses and secondly, perception is highly subjective meaning that the information gets distorted to coincide with our pre-established beliefs, attitudes and values. For example, a pre-conceived idea that a given person or an organization is honest or deceptive, good or poor source of information, late or prompt on delivery can have a considerable effect on the objective ability of the decision maker and the quality of the decision.

(iii) Cognitive constraints: A human brain, which is the source of thinking, creativity and thus decision making, is limited in capacity in a number of ways. For example, except in unique circumstances, our memory is short term with a capacity of only a few ideas, words and symbols. Secondly, we cannot perform more than a very limited number of calculations in our heads which are not enough to compare all the possible alternatives and make a choice. Finally, psychologically, we are always uncomfortable with making decisions. We are never really sure if our choice of the alternative was correct and optimal, until the impact of the implication of the decision has been felt. This makes us feel very insecure. These

constraints limit us to use 'Heuristics', which means limiting the search for facts and data and using the limited information for decision making. This leads to 'satisfactory' decisions rather than optimal decisions.

(iv)   Attitudes about risk and uncertainty: These attitudes are developed in a person, partly due to certain personal characteristics and partly due to organizational characteristics. If the organizational policy is such that it penalizes losses ignore than it rewards gains, then the decision maker would tend to avoid such alternatives that have some chances of failure even though the probability of substantial potential gains is very high. The risk taking attitude is influenced by the following variables: Intelligence of the decision maker. Higher intelligence results in highly conservative attitudes and highly conservative decision makers are low risk takers. The less intelligent decision makers are generally more willing to take calculated risks if the potential rewards are large and there is some chance of success.

Expectations of the decision maker.

People with high expectations are generally highly optimistic in nature and are willing to make decisions even with less information. The decision makers with low expectations of success will require more and more information to decide upon a course of action.

Decision Making Time constraints.

As the complexity of the personal habits of the decision maker and the complexity of the decision variables increases, so does the time required to make a rational decision. Even though, there are certain individuals who work best under time pressures and Interpersonal Process and Conflict Resolution may out-perform others under severe time constraints, most people, by and large, require time to gather all the available information for evaluation purposes. However, most people under time

pressures rely on 'Heuristic' approach, considering few characteristics of alternatives and focusing on reasons to reject some alternatives. This approach may also be in use when the cost of gathering information and evaluating all such information is high.

    (v)   Personal habits: Personal habits of the decision maker, even though formed through social environmental impact and personal perceptual processes, must be studied in order to predict his decision making style. Some people stick to their decisions even when these decisions are not optimal and try to shift the blame for failure on outside factors rather than their own mistakes. For example, Hitler found himself bound by his own decisions. Once he decided to attack Russia, there was no coming back even when it was realised that the decision was not the right one. Some people cannot admit that they are wrong and they continue with their decisions as before even ignoring such evidence which indicates that a change is necessary. These personal habits have a great impact on organizational operations and effectiveness.

    (vi)   Social and cultural influences: The social and group norms exert considerable influence on the style of the decision maker. Ebert and Mitchell define a social norm to be "an evaluating scale designating an acceptable latitude and objectionable latitude for behaviour, activity, events, beliefs or any object of concern to members of a social unit. In other words, social norm is the standard and accepted way of making judgements". Similarly, cultural upbringing and various cultural dimensions have a profound impact on the decision making style of an individual. For example, in the Japanese organizational system, a decision maker arrives at decisions in consensus with others. This style is culturally oriented and makes implementation of the decision much easier, since everybody participates in the decision making process. In America, on the contrary, the decision making style is highly

individualistic with the of decision models and decision techniques.

Check Your Progress 2

Note: a) Use the space provided for your answer

b) Check your answers with those provided at the end of the unit

4.   Discuss the Models of Decision Making.

_____

_____

_____

5.   Discuss the Creativity and Decision making.

_____

_____

_____

8.7 SOME COMMON ERRORS IN DECISION MAKING

Since the importance of the right decision cannot be overestimated, because the quality of the decision can make the difference between success and failure, it is imperative that all factors affecting the decision be properly looked at and fully investigated. In addition to technical and operational factors which can be quantified and analyzed, other factors such as personal values, personality traits, psychological assessment, perceptions about the environment, intuitional and judgmental capabilities and emotional interference must also be understood and credited. Some researchers have pinpointed certain areas where managerial thinking needs to be re-assessed and where some common mistakes are usually made. These mistakes that affect the decision making process as well as the efficiency of the decision should be avoided as far as possible. Some other errors are:

(i)   Indecisiveness: Decision making is a very heavy responsibility. The fear of its outcome can make some

people timid about making a decision. This timidity may result in taking a long time for making a decision and this may result in the loss of a good opportunity. This trait is a personality trait and must be looked into seriously.

(ii)   Postponing the decision until the last moment: This is quite a common practice and results in decision making under pressure of time which generally eliminates the possibility of a thorough analysis of the problem since such analysis is time consuming. It also makes it practically impossible to establish and compare all possible alternatives. For example, many students who postpone studying until their final exams usually do not I fare well in the exams.

(iii)   Failure to isolate the root cause of the problem: It is a very common practice to cure the symptoms, rather than the causes. For example, a headache may be a symptom of some deep rooted emotional problem so that just a medicine for the headache would not cure the problem. It is necessary to separate the symptoms from the causes. Success of a decision is dependent upon the correct definition of the problem.

(iv)   Failure to assess the reliability of informational sources: Very often, we take it for granted that the other person's opinion is very reliable and trustworthy and we do not check for the accuracy of such information for ourselves. Many times, the opinion of the other person is taken so that if the decision fails to bring the desired results, the blame for the failure can be shifted to the person who had provided the information. However, this is a poor reflection on the manager's ability and integrity and the manager must be held responsible for the outcome of the decision. Accordingly, it is his moral duty to analytically judge the accuracy and reliability of the information that is provided to him.

(v)   The method for analyzing the information may not be a sound one: Since most decisions and specially the non-

programmed ones have to be based upon a lot of information, and many factors and variables, the procedures to identify, isolate and select the useful information must be sound and dependable. Usually, it is not operationally feasible to objectively analyse more than five or six pieces of information at any given time. Hence, g model must be built which incorporates and handles many variables in order to aid the decision maker. Also, it is desirable to define the objectives, criteria and constraints as early in the decision making process as possible. This would assist in making the process more formal so that no conditions or alternatives would be overlooked.

Stages in the creative process

Stage	Type	Behaviours
Preparation	Conscious	**Saturation:** Investigating the problem in all directions to become fully familiar with it, its setting, causes, and effects. **Deliberation:** Mulling over these ideas, analyzing and challenging them, viewing them from different optics.
Latent period	Unconscious	**Incubation:** Relaxing, switching off, and turning the problem over to the unconscious mind. **Illumination:** Emerging with possible answers – dramatic, perhaps off beat, but fresh and new.
Presentation	Conscious	**Verification:** Clarifying and flushing out the idea, testing it against the criterion of appropriateness. **Accommodation:** Trying out the solution on other people and other problems.

Activity 8 Do you use any other model for making decisions other than the models presented above? Illustrate.

8.8 CHOICE THEORY

Decisions vs. Choices

With decision, it is more of a process orientation, meaning we are going through analysis and steps to eliminate (or, cut off) options.

With choice, it is more of a mindset approach, meaning we have a perception of what the right or wrong choice may be.

Does this all matter? I believe it does.

We can easily setup processes to enable the best decisions possible. The decisions can range in scope from being low impact to high impact, and we can build in checks and balances along the way in reaching a decision. It can be a thoughtful, thorough approach.

With choices, we face opportunities – large and small – to select or choose an option. Although we may put thought into the larger choices we make, the smaller ones may be more instinctively made.

We make choices based on our values, beliefs, and perceptions of where a selected one may take us. One of life's responsibilities is centered here in that we need to spend time building our choice senses and systems.

Choices are more difficult. At times, we cannot collect all the data, analyze the options, and reach a sound conclusion. Time escapes us to "cut off" certain options because life choices fit a different model. Choices involve our life more in which path we select and the direction – intended or unintended – it then takes us.

A Distinction: Decisions vs. Choices

Here is the kicker. We may make many decisions during a day, week, or month, but how many life or leadership choices do we really make?

We can go through our life making decisions on where to live, work, and play, but do we make the choice of how to best live or lead?

We should take the time to make more proactive choices in setting a life and leadership direction.

We can spend our lifetime making all sorts of decisions, yet we spend little, if any, time making distinctive life choices.

I believe we may need to focus more on making choices than making decisions. We need to make real and necessary choices on how to lead our life in the most purposeful way possible.

Three key points:

Life choices set a direction, so spend the time to develop a mindset on how to build a meaningful future.
Choose how to lead. How we lead our teams and our work demands a thoughtful approach, meaning we need to define the presence we want through our actions and interactions.
Decisions still matter, so use a process to prune out the bad options and select the best ones possible.
Make life and leadership choices. Live inspired choices.

What is the difference between decision and choice? When you think about it, are they the same thing? In practice, although we often use the words interchangeably, it depends on how you define the words.

Choice

Life may be likened to a path. We walk along the pathway of our lives, doing what we do. And sometimes we come to a fork in the path, where we must choose which way to go. Sometimes these choices are minor, for example whether to have a cup of coffee or tea. Minor choices do not really affect our lives much as we continue on the major route. Other choices are major and life-changing, such as what career we will follow.

Choice, then, is selection from alternatives.
This means we must see the alternatives from which we can choose. Sometimes these are obvious but often they are not and the path we walk can have a significant random element. Being alert and able to see the

choices we have is a critical ability for living deliberately.

Choosing is the process of selection. Classically, we weigh up each option, considering pros and cons. We then select the most advantageous option. In practice we are limited by time and the linear nature of conscious thought, so we leave a lot to our unconscious minds, which use intuition, rules of thumb, habit and so on. We seldom have complete information and may have to guess. We may also copy others or be swayed by their arguments.

Decision

Decision is a more general term that does not imply the existence of alternatives. It is driven more by needs, goals and problems than by simply encountering a set of choices.

Decision is a process that can vary depending on the situation.

- Decision is the same as choice when it is 'deciding between' pre-existing or provided alternatives.
- Decision can be a part of choice when choosing is not simple, for example where deeper consideration is needed in deciding whether to marry a person or not.
- Decision can be generative when it creates alternatives from which to choose, such as when someone investigates different towns and houses as they decide on where to live.
- Decision can be predictive, such as when we 'decide if' it will rain. While this still includes the choice between rain and other weather states, the process of forecasting is more than a simple choice.
- Decision can also be evaluative, for example where we 'decide how' friendly a person is. Choice here is less distinct, although we may subsequently choose to ignore the person or to spend more time with them.
- Decision can be more about direction than selection, such as when a person decides to travel more, thus setting up a new vista of holiday choices.
- Decision can be slow, for example when a person mulls over their life so far, wondering who the want to be. It can also be fast, such

as in the quick decisions of racing driver.

Check Your Progress 3

Note: a) Use the space provided for your answer

b) Check your answers with those provided at the end of the unit

6.   Discuss the Some Common Errors in Decision making.

_____

_____

_____

7.   Write about the Choice Theory.

_____

_____

_____

8.9   LET US SUM UP

In this unit we have discussed the importance of decision making, the process involved and different models of decision making and how decision making is helpful in everyday life. We discussed various characteristics of decision making. Four types of decisions have been discussed in this unit. We have also discussed four major models: contingency model, economic man model, administrative man model and social man model. Various factors and personal characteristics that have an impact on decision making have also been discussed.

8.10 KEY WORDS

Decision: the act of or need for making up one's mind.

Choice: the right, power, or opportunity to choose.

8.11 QUESTIONS FOR REVIEW

1. Explain the concept of rationality in decision making? What are its limitations?

2. Discuss the economic man model of decision making. How does it differ from administrative man model of decision making?

3. Discuss various types of decision making and illustrate with reference to the school you are working in.

4. Discuss the process and relationship of creativity and decision making.

8.12 SUGGESTED READINGS AND REFERENCES

- Chhabra, T.N., (2002): Principles and Practice sf Management. Dhanpat I i Rai &. Co., Delhi.

- IGNOU, (1987): Management Functions and Rehaviour ,MS-1. IGNOU I Study Material, New Delhi.

- JITS, Chandan, (1998): Organisational Rehaviour. Vikas Publishing House I b Pvt. Ltd.,Delhi.

- Mitchell, T.R. and Larson, J.R., (1987): People in Organisations: An I Introduction to Organizational Behaviour. McGraw Hill International - Edition.

- Raju, R.S. and Parthasarathy, A., (2000): Management Text and Cases. Prentice-Hall of India Private Limited, India.

8.13 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

1. See Section 8.2
2. See Section 8.3
3. See Section 8.4

Check Your Progress 2

1. See Section 8.4
2. See Section 8.5

Check Your Progress 3

1. See Section 8.6
2. See Section 8.7

UNIT 9: EXISTENCE

9.0   OBJECTIVES

After this unit, we can able to know:

- To know about the Frege and Russell: Existence is not a Property of Individuals
- To discuss about the Meinongianism
- To know the Anti-Meinongian First-Order View

9.1   INTRODUCTION

Existence raises deep and important problems in metaphysics, philosophy of language, and philosophical logic. Many of the issues can be organized around the following two questions: Is existence a property of individuals? and Assuming that existence is a property of individuals, are there individuals that lack it?

What does it mean to ask if existence is a property? A full answer to this question requires a general theory of properties, which is well beyond the scope of this article. I briefly sketch the landscape to set up our discussion of existence. (See the entries on properties and substance for deeper discussion.) Properties contrast with individuals. This distinction can be

explicated using the instantiation relation. My cat instantiates the property of being hungry, as that is a way he is, and, being an individual, is not himself instantiated by anything. While properties also instantiate—the property of being red, for example, has the property of being a color—only properties are instantiated; individuals only instantiate. So, our first question is whether existence is instantiated and, if so, whether it is instantiated by individuals like Obama, my chair, and the fig tree in my backyard. Do individuals, in addition to ordinary properties like being human, being comfortable to sit in, and needing more water, instantiate a property expressed by the English verb 'exists'?

There is a debate in the literature on properties between the abundant conception of properties, according to which there is a property corresponding to every natural language predicate and, more generally, every class of individuals, and the sparse conception of properties, according to which a predicate expresses a property only if the objects that predicate is true of resemble one another in an intrinsic way. If the abundant conception is true, then our first question may seem trivial: Existence is a property of individuals because sentences like 'Bill Gates exists' are grammatical and there is a class of all individuals and hence a corresponding property of existing. It seems, then, that our first question has bite only if the sparse conception is true. But appearances deceive. As we will see in section 1, there is a controversy as to whether the logical form of a sentence like 'Bill Gates exists' is really subject-predicate in structure and so whether the English verb 'exists' really is predicated of individuals. The question whether existence is a property of individuals is perhaps more straightforward on the sparse conception of properties. But the question can still be raised even on the abundant conception as the question whether existence is a property of individuals involved in our talk of what exists and what does not, which is then a question about the logical form of the sentences used in our existential discourse.

We can trace the issue of whether existence is a property to a disagreement between the ancient Greek philosopher Aristotle and some of his medieval followers over the relationship between an individual's essence and its existence. The debate requires some background. We begin with the distinction between accident and essence and that

distinction's relation to contingency and necessity. Some of a thing's properties are contingent, in the sense that the thing might not have had them. I am writing right now, but I might have been out for a run instead. So working on a paper right now is one of my contingent properties. Contingent properties contrast with necessary properties. I am necessarily human, in the sense that it is impossible that I am a nonhuman. All contingent properties are accidents and all essences are necessary but, according to the Aristotelian, some necessary properties are accidents. A thing's essential properties are inseparable from the bearer, not only in the sense that the property is necessarily had by that object but in the deeper sense that any adequate account of what that object is involves that property; they are part of any adequate definition of the thing or answer to the question 'What is it?'. I am essentially a human and perhaps essentially the person I in fact am, if there are individual essences in addition to general essences. I am necessarily identical to something and necessarily such that 2+2=4, but these properties are among my accidents, as they are not part of any adequate account of what I am and what distinguishes me from others. While this distinction between a thing's necessary properties and its essential properties is controversial, there is no doubt that it was Aristotle's view, shared by most of his medieval followers, and so informed the first historical occurrence of the debate about existence under discussion. Aristotle seems to have seen nothing more to existence than essence; there is not a space between an articulation of what a thing is and that thing's existing. Saint Thomas Aquinas, on the other hand, famously distinguished a thing's essence from its existence. Aquinas argued something as follows, in chapter 4 of his On Being and Essence. One can have an understanding of what a man or a phoenix is without knowing whether it exists. So, existence is something in addition to essence. In short, Aquinas argued that existence is a separate property as existence is not part of most objects's natures and so those objects can be conceived or thought of separately from their existing.

There is a long and distinguished line of philosophers, including David Hume, Immanuel Kant, Gottlob Frege, and Bertrand Russell, who followed Aristotle in denying that existence is a property of individuals,

even as they rejected other aspects of Aristotle's views. Hume argued (in A Treatise of Human Nature 1.2.6) that there is no impression of existence distinct from the impression of an object, which is ultimately on Hume's view a bundle of qualities. As all of our contentful ideas derive from impressions, Hume concluded that existence is not a separate property of an object. Kant's criticism of the ontological arguments for the existence of God rested on a rejection of the claim that existence is a property of an object. Proponents of the ontological argument argue that the concept of God as an entity with all perfections or a being of which no greater can be conceived entails God's existence, as existence is a perfection and a being that exists is greater than a being that does not exist. Kant objected (in his Critique of Pure Reason, A596/B624-A602/B630) that existence is not a property. "Thus when I think a thing, through whichever and however many predicates I like (even in its thoroughgoing determination), not the least bit gets added to the thing when I posit in addition that this thing is. For otherwise what would exist would not be the same as what I had thought in my concept, but more than that, and I could not say that the very object of my concept exists" (A600/B628). Finally, both Frege and Russell maintained that existence is not a property of individuals but instead a second-order property—a property of concepts, for Frege, and of propositional functions, for Russell. Crudely, to say that dinosaurs do not exist is to say that the property of being a dinosaur is not instantiated; to say that Jean-Baptiste Botul does not exist is to say that some property—say, the property of being a unique post-war critic of Kant and father of Butolism—is not instantiated. In neither case is it to say of some individual that it does not exist, which neither Frege nor Russell thought made sense.

The view that existence is not a property of individuals became the common view in the early 20th Century. While Aristotle, Hume, and Kant's related reasons for the thesis have persuaded some, the dominant force behind this agreement is the thought, developed most forcefully in [Russell 1905b], that denying that existence is a first-order property is the only way to avoid the consequence that there are things that do not exist and thus that there is a distinction between being and existing, our second framing question from above. The thesis that there are things that do not

exist was held by the Austrian philosopher Alexius Meinong. Existence is a genuine property of individuals, Meinong maintained, but not universally had. While there are many ways to motivate Meinongianism, a primary motivation is the puzzle of negative singular existentials—sentences that seem to truly assert the nonexistence of an individual, such as the sentence 'Jean-Baptiste Botul does not exist'. In order to be true, it seems, the subject position must designate some entity of which nonexistence is truly predicated, in which case there are things—the designation of these singular terms—that do not exist. Frege and Russell, by contrast, take the same sentences to demonstrate that those expressions are not genuine singular terms at all and that negative existentials all have a general form, asserting the noninstantiation of a property. In the following section I discuss Frege and Russell's account of true negative existentials. In the second section I discuss Meinongianism, comparing the account of true negative existentials on offer by the Meinongian with the Russellian account discussed in the first section. The article ends with a discussion of an anti-Meinongian account according to which existence is a universal property of individuals and a discussion of the related issues of existence in the context of quantified tense and modal logics.

9.2   FREGE AND RUSSELL: EXISTENCE IS NOT A PROPERTY OF INDIVIDUALS

There are two sets of reasons for denying that existence is a property of individuals. The first is Hume and Kant's puzzlement over what existence would add to an object. What is the difference between a red apple and a red existing apple? To be red (or even to be an apple) it must already exist, as only existing things instantiate properties. (This principle—that existence is conceptually prior to predication—is rejected by Meinongians.) Saying it is red and an apple and furthermore exists is to say one thing too many. The thought seems to be that instantiating any property whatsoever presupposes existence and so existence is not a further property over and above a thing's genuine properties. The thought is not merely that everything that instantiates any property exists, as the same is true of being self-identical, being either human or not human—assuming the law of excluded middle—and being such that 2+2=4, all of

which seem to be unproblematic properties of individuals even if that status is denied of existence. Instead the thought is that instantiating any property whatsoever conceptually presupposes the existence of a subject in a way that makes it incoherent to then think of existence as a further property of that thing. The thing's existence is prior to any predication to it and so it is incoherent to think of existence as a property had by the thing. This thought is behind Aristotle's thesis that existence is not a further feature of a thing beyond its essence.

The second consideration favoring the thesis that existence is not a property of individuals concerns the puzzle of negative singular existentials. Suppose that existence is a property of the designation of the subject term in a singular existential sentence. Then 'Ronald McDonnald does not exist' predicates nonexistence of the designation of the subject term, in which case reality includes an entity—the designation of the singular term and subject of predication—that has the property of not existing. That, Russell complained, runs contrary to a robust sense of reality, according to which everything exists. So, we should reject the claim that existence is a property of the designation of subject terms in existential sentences.

To appreciate Russell's alternative account, consider first general nonexistence claims. To say that foxes exist is to say that there are some things that are foxes; that is, the property of being a fox is instantiated. This is reflected in the standard regimentation of the sentences 'Foxes exist' and 'There are foxes' in first-order quantificational logic as $\exists x Fx$, where Fx is the translation for the predicate 'is a fox'. General kind terms do not, then, designate individuals, which we then (redundantly) say exist when using the predicate 'exists' or (paradoxically) say are not when using the predicate 'does not exist'. Instead, kind terms designate properties and simple seeming subject-predicate sentences like 'Foxes are carnivores' are claimed to possess a more complicated logical form, $\forall x (Fx \rightarrow Cx)$, where Cx translates the predicate 'is a carnivore'. (I ignore the difficult question whether generics are really quantifiers at all, made more troubling by the fact that some generics seem to admit of exceptions—'Birds fly' is true, even though penguins are birds and don't fly; 'Cats have four legs' is true, even though there is a three-legged cat

wandering my neighborhood.) Given this analysis, general nonexistence claims are unproblematic. The sentence 'Dragons do not exist' says, on this analysis, that the property of being a dragon is not instantiated. Take the most inclusive class of what there is; nothing in that class has the property of being a dragon. That is what ¬∃xDx says, letting Dx translate the predicate 'is a dragon'. This is significant because it does not require identifying some entity to then predicate of that thing the property of nonexistence.

The Frege-Russell view that existence is a second-order property is based on the idea that seemingly singular existential and negative existential sentence like 'Bill Gates exists' and 'Ronald McDonnald does not exist' are, in their deeper logical form, general existential and negative existential claims. I focus on Russell's version of the view.

Russell claimed that ordinary proper names like 'Bill Gates' are disguised definite descriptions, something like 'the richest man in the world'. And definite descriptions, given Russell's view of definite descriptions, are not genuine referring terms but are instead quantificational expressions. The sentence 'The richest man in the world lives in Washington' has, as its logical form, a quantificational structure and not a subject-predicate structure, equivalent to something like the following: There is a unique richest person who lives in Washington. Individuals do not enter directly into the proposition expressed by the sentence and are not part of the sentence's truth conditions.

These features of Russell's account of definite descriptions are significant for the treatment of seemingly singular existential and negative existentials as they remove the need for entities to serve as the designation of the singular terms for the meaningfulness and truth of negative existentials. Seemingly singular existentials like 'Bill Gates exists' are assimilated to general existentials like 'Foxes exist'. Assuming the proper name 'Bill Gates' is analyzed as the definite description 'the richest person alive', the sentence 'Bill Gates exists' has a logical form that can be more accurately expressed as There is someone that is uniquely richer than anyone else alive. This is neither redundant nor uninformative, assuming that we can grasp in thought properties while coherently and rationally wondering whether or not they are instantiated. Russell's

account similarly dissolves the problems generated by seemingly singular negative existentials like 'Ronald McDonald does not exist'. The truth of this sentence does not require a designation for the term of which nonexistence is then predicated. 'Ronald McDonald' is short for a definite description, say, 'the happy hamburger clown'. The sentence 'Ronald McDonald does not exist' expresses a proposition of the form It is not the case that there is a unique happy hamburger clown. This proposition is true even if absolutely everything there is exists. The proposition concerns the property of being a happy hamburger clown and says of that property that it is not uniquely instantiated. As the property—the true subject of predication—exists, however, we are not forced to countenance the reality of entities that do not exist in order to recognize this sentence as saying something true.

Russell's strategy depends on two claims. The first is that the negation in a negative existential takes wide scope, applying to the whole sub-sentence and not just the predicate. So, 'Ronald McDonald does not exist' does not involve ascribing the predicate 'is nonexistent' to the subject 'Ronald McDonald'. Instead, it is more faithfully represented as 'It is not the case that [Ronald McDonald exists]'. The second is that 'Ronald McDonald' is not a genuine referring expression and the predicate 'exists' really means something like is instantiated. Notice that the first in solitude is not sufficient to overcome the problems generated by seemingly singular negative existentials. Even if the deep form of 'Ronald McDonald does not exist' is 'It is not the case that [Ronald McDonald exists]', assuming that 'Ronald McDonald' is a genuine singular term, the problem remains of finding in reality some entity to serve as the designation of 'Ronald McDonald'. That entity is then part of reality and so, assuming that Meinongianism is false, is existent. In that case, the sub-proposition Ronald McDonald exists is true and so its negation false. The problem of true singular negative existentials does not rest on the supposition that they involve ascribing the property of nonexistence. So, it is the second of the above claims that carries the weight of Russell's solution to the problem of singular negative existentials.

The second component of the Russellian solution—the claim that ordinary proper names like 'Bill Gates' are disguised definite descriptions—faces a

number of objections. One is the semantic argument. (See [Kripke 1972].) Suppose that the descriptive equivalent of the name 'Bill Gates' is 'the richest living person in the world'. An adequate semantic understanding of the sentence 'Bill Gates is richer than everyone else alive' would then be sufficient for recognition of its truth. (More precisely, the sentence, 'If anyone is richer than everyone else alive, then Bill Gates is richer than everyone else alive' has this feature.) But that seems implausible. Surely we must collect empirical data to determine its truth. One who wonders whether someone, say, Warren Buffett, is wealthier than Bill Gates does not display irrationality or semantic ignorance, comparable to one who wonders whether a fortnight is longer than 14 nights.

Perhaps these considerations should motivate the descriptivist to abandon "great deeds" descriptions in favor of metalinguistic descriptions like 'the person named 'Bill Gates'' or causal descriptions like 'the person that stands at the origin of this chain of uses of the name 'Bill Gates''. It is plausible that semantic competence suffices to know that the sentence 'Bill Gates is named 'Bill Gates'' is true and, while perhaps importing extra-semantic facts about language use, it is plausible that any reflective speaker of English knows that any token of 'Bill Gates is the person that stands at the origin of this chain of uses of the name 'Bill Gates'' is true. So these descriptions seem to survive the semantic argument presented in the previous paragraph. But they face another objection, also facing simpler versions of descriptivism: Namely, the modal objection ([Kripke 1972]). While it is absolutely impossible that Bill Gates is not Bill Gates, it is, it seems metaphysically possible that Bill Gates is not the richest person alive, instead being a middle American, and it seems metaphysically possible that he is not named 'Bill Gates' and does not stand at the causal origin of any particular chain of uses of the name 'Bill Gates'. This suggests that ordinary proper names and their alleged descriptive equivalents considered above are not, in fact, semantically equivalent, as they embed differently under modals like 'it is necessary that'. [See the entry on names for further discussion of these problems.]

In response to the modal argument, the descriptivist might avail herself of individual essence descriptions like 'the person identical to Bill Gates', 'the person that Bill-Gatizes', or rigidifications of the above descriptions,

'the person actually named 'Bill Gates'' and 'the person that actually stands at the origin of this chain of uses of the name 'Bill Gates'', all of which designate the same person in every possible world in which they designate anything. It is plausible that semantic competence suffices for recognition of the truth of the sentence 'Bill Gates is the person identical to Bill Gates' and that that sentence expresses a necessary truth. So these versions of descriptivism seem to escape the problems discussed in the previous paragraphs. The first two candidates, however, do not hold much promise for solving the problem of apparently true singular negative existentials. We know what the property of being identical to Bill Gates is, but only because we know the result of plugging up one of the relata in the two-place relation is identical to with the individual Bill Gates. Insofar as we think that reality does not include any entity identical to Ronald McDonald, however, we are then left to wonder what the property of being identical to Ronald McDonald is. Because the contents of these properties are derivative from the individuals that serve as the referents of their names, they are poor candidate descriptive equivalences for a robust version of descriptivism and unlikely to shed light on the truth of seemingly singular negative existentials like 'Ronald McDonald does not exist'. Similar considerations apply to the predicating view.

The last candidates, rigidified metalinguistic and causal descriptions, are the most promising. But some have claimed to discern important differences in the functioning of a name and its alleged semantically equivalent rigidified description, of any flavor. First, some have claimed that the name 'Bill Gates' designates Bill Gates with respect to every possible world, including worlds at which Bill Gates does not exist; otherwise the sentence 'Bill Gates does not exist' would not be true with respect to those worlds. But the rigidified description 'the person actually named 'Bill Gates'', for example, does not designate anything with respect to such a world, as nothing in the domain of that world satisfies the condition being named 'Bill Gates' at the actual world. That is because it is only Bill Gates that satisfies that condition and he is not a member of the domain of the possible world in question. So, differences in how a name and a rigidified description embed under modal operators can still be discerned. (See [Salmon 1981] for further discussion.) This

objection assumes that the domain of quantification varies from world to world and that individuals that serve as the designation of ordinary names are genuine contingent existents, which some may deny. The objection also assumes that the range of the description is the domain of the world with respect to which the description is being evaluated, the actuality operator rigidifying only the condition of the description, which may also be denied. The second objection to rigidified descriptivism concerns the differences some have claimed between how names and rigidified descriptions embed under propositional attitude verbs. Intuitively, Jones would have still believed that Bill Gates is wealthy even if things had been ever so slightly different than they actually are—say, I bought a poppy bagel instead of a sesame bagel this morning. But if the content of Jones's belief concerns the actual world, as rigidified descriptivism dictates, then, to retain his actual belief in that counterfactual situation, he would have to believe something about another possible world—the actual world. But it is implausible that Jones would have a belief about another possible world. So, the content of Jones's belief does not concern the actual world and so one can believe what is expressed by 'Bill Gates is wealthy' without believing what is expressed by 'The person that actually stands at the origin of this chain of uses of the name 'Bill Gates' is wealthy'. (See [Soames 1998].)

In this section I examined the thesis that seemingly singular existential and negative existential sentences are really general existentials, which are then treated as ascribing the property of being instantiated or not instantiated to some property. The need to grant being to entities that do not exist in order to account for the truth of sentences like 'Ronald McDonald does not exist' is then avoided, which is no small victory. The success of that proposal, however, was seen to rest on the claim that ordinary proper names have descriptive equivalences, which many philosophers of language reject.

9.3 MEINONGIANISM

Perhaps, then, we should reject descriptivism and accept that ordinary proper names are devices of direct reference, that there are true genuinely singular negative existentials, and so that there are nonexistent objects.

'Ronald McDonald' seems like a referring term, open to existential generalization, in the sense that a sentence like 'Ronald McDonald does not exist' entails 'There is something that does not exist', and 'exists' seems like a predicate that applies or fails to apply to the designation of subject-place terms. The Meinongian accepts these appearances and concludes that reality includes referents for empty names and those referents do not exist. The Meinongian trades logical and semantic simplicity for metaphysical abundance.

Meinongianism is the thesis that there are objects that do not exist, nonexistent entities being included in the most unrestricted domain of quantification and discourse. One immediate challenge to the Meinongian is to offer individuating conditions for nonexistents. The most straightforward comprehension principle is the naive principle that, for any condition on objects, there is a unique object satisfying exactly that condition. For our purposes, we can conceive of a condition as determining a set of properties; crudely, the properties expressed by the predicates composing the condition. In that case, condition C is the same condition as C′ when they determine the same set of properties. It follows that corresponding to any set of properties, there is exactly one object with exactly those properties. The naive comprehension principle faces several problems. In what remains of this section, I survey these problems and distinguish different versions of Meinongianism in terms of the devices employed to develop a restricted comprehension principle for objects that avoids them.

The first is the problem of incomplete objects. Conditions need not be total; that is, we do not require that the set of properties a condition determines is such that, for every property, either it or its complement is a member of that set. So, by the naive comprehension principle, the condition of being a singer defines an object with exactly that property— being a singer—and no other properties. A set with other properties as well is a distinct set of properties and so corresponds to a different condition and hence a different object. Some find incomplete objects problematic in themselves, as they are counterexamples to bivalence: Our singer, for example, is neither wearing a dress nor not wearing a dress.

But they also lead to more general threats of paradox. Our singer is an object with exactly one property: That of being a singer. This is its sole defining characteristic. So having a exactly one property is also a property of our singer and that property is distinct from the property of being a singer, which our singer also has. So, the singer has two properties Contradiction. One simple solution is to restrict the comprehension principle to total conditions. The resulting proposal, however, leads to a questionable application of Meinongian metaphysics to problems of fictional truth, as many want to claim that there is simply no fact of the matter as to whether or not Sherlock Holmes has a mole on his left shoulder, as that is left underdetermined by the Holmes stories and there are no deeper grounds for either predication. The promise of employing nonexistent objects in explaining apparent truths about fiction is one of the theory's main virtues. Relatedly, this solution undermines a primary motivation for Meinongianism—namely, the idea that there is a subject of predication corresponding to any object of thought, as we certainly do not think only of complete objects.

The second is the problem of contradiction. A naive comprehension principle generates objects that violate the principle of noncontradiction. Consider the condition of being taller than everything. By the naive comprehension principle, this condition determines an object and so there is an object that has exactly the property of being taller than everything. But then it is taller than itself, which is a contradiction given the irreflexivity of the taller than relation. The irreflexivity of the taller than relation is nonlogical. But not so with the identity relation, as = is typically taken to be a logical predicate. It is a logical truth that everything is self-identical; i.e., the sentence $\forall x\ x=x$ is true under every interpretation. But consider the property of being self-distinct. By the naive comprehension principle this condition determines an object and that object is self-distinct. But then that object does not satisfy the condition $x=x$. So our logically true sentence has a counterinstance. Contradiction.

A third problem, one of Russell's objections to Meinongianism (see [Russell 1905a, 1907]), turns on the fact that existence is, on Meinongianism, a property and hence figures into the base of the naive comprehension principle. So, consider the condition of being winged, being a horse, and existing. By the naive comprehension principle, there is an object with exactly these features. But then this object exists, as existing is one of its characterizing features. Intuitively, however, there is no existent winged horse; existing seems to require a bit more substance. Indeed, for every intuitively nonexistent object that motivates Meinongianism—Zeus, Pegasus, Santa Clause, and Ronald McDonald— there is, by the naive abstraction principle, an object just like it but with additional the property of existing. But then there is an existing Zeus, an existing Pegasus, etc.. This is overpopulation not of being but of existence as well.

The naive comprehension principle, then, must be rejected and a restricted principle connecting sets of properties with objects found. The principle should generate enough objects to serve the Meinongian purpose of ensuring a corresponding object for every thought while avoiding the problems discussed above. We can distinguish two strategies, both suggested by Meinong's student Ernst Mally [Mally 1912]. The first distinguishes two kinds of properties, what, following Terence Parsons [Parsons 1980], we shall call nuclear and extra-nuclear properties. While the distinction remains ultimately unclear, the key idea is that nuclear properties are part of a thing's nature, broadly construed, and extra-nuclear properties are external to a thing's nature; more precisely, nuclear properties, but not extra-nuclear properties, are part of the characterization of what the object is. The comprehension principle is then restricted to conditions involving only nuclear predicates. Problematic properties, like existing, etc., are deemed extra-nuclear and beyond the scope of the comprehension principle, not determining the objects that there are. Nuclear, not extra-nuclear, properties individuate objects. The second Meinongian camp distinguishes two modes of predication: What Mally called determining and satisfying, Hector-Neri Castañeda [Castañeda 1974] called internal and external predication, William Rapaport

Notes

[Rapaport 1978] called constituency and exemplification, Kit Fine [Fine 1982] called implicit and explicit, and Edward Zalta [Zalta 1983, 1988] called encoding and exemplifying. There is, on this view, a single class of properties that the comprehension principle ranges over, but the principle determines the properties encoded not exemplified (to follow Zalta's terminology). For every condition, there is a unique object that encodes just those properties. An object may or may not exemplify the properties it encodes. Sherlock Holmes encodes the properties of being a detective and living at 221B Baker Street, etc., but he does not exemplify those properties. He exemplifies (but does not encode) the properties of being a fictional character and being the hero of Arthur Conan Doyle's Holmes stories.

How do these distinctions solve the problems raised above for the naive comprehension principle? I begin with Parsons's view. Parsons focuses on the problems of contradiction and of the existent winged horse. Following Russell's discussion of Meinong, in [Russell 1905a, 1907], Parsons considers the threat of contradiction generated by impossible objects like the round square. Meinong claimed that there is a round square, but that, complained Russell, leads to violations of the principle of noncontradiction, as that entity is then both round and not round, in light of the fact that it is square, which entails that it is not round. Parsons's response (see [Parsons 1980], 38-42) seems to be to deny that being square entails not being round, in which case it is simply false that the round square is not round. He thinks that that implication holds only for "real" objects. He claims that there are counterexamples to the claim that all square objects are not round; after all, the round square is a square object that is round! This solution, however, does not seem to solve the more general threat of contradiction, as discussed above. Indeed, Parsons himself recognizes the limited success of his response (see [Parsons, 1980, 42n8]). He allows that being non-squared is a nuclear property. But then his comprehension principle entails that there is an object corresponding to the condition of being a non-squared square, where that object instantiates the incompatible properties of being a square and being a non-square.

Let's turn to Parsons's response to the existence problem. The naive comprehension principle faced the problem of generating an existent winged horse. Because existence is an extra-nuclear property, however, Parsons's version of the comprehension principle, which correlates sets of only nuclear properties to objects, avoids this problem. The condition of being an existent winged horse is not composed solely of nuclear properties and so Parsons's principle does not correlate it to an object. Parsons's distinction between nuclear and extra-nuclear properties similarly promises to solve the problem of incomplete objects. Recall our singer from above. That object does not have exactly one property; instead, it has exactly one nuclear property. As having exactly one nuclear property is itself an extra-nuclear property, much as being a complete object is on Parsons's view, the threat of contradiction is avoided.

The distinction between nuclear and extra-nuclear properties remains unclear. Parsons introduced the distinction with lists of nuclear predicates ('is blue', 'is tall', 'kicked Socrates', 'is a mountain') and extra-nuclear predicates ('exists', 'is thought about by Meinong', 'is complete'). He then tells us that the extra-nuclear are those that do not stand for properties of individuals ([Parsons 1980, 24]). And, of course, it is nuclear and not extra-nuclear properties by which objects are individuated. Parsons's individuation principle for objects is the following: "(1) No two objects (real or unreal) have exactly the same nuclear properties; and (2) For any set of nuclear properties, some object has all the properties in that set and no other nuclear properties" ([Parsons 1980, 19]). But it is not clear what status individual identity properties—properties like being identical to A, where A is an individual substance like, say, Parsons himself—have with respect to this distinction. He sometimes claims that they are extra-nuclear properties ([Parsons 1980, 28]). In that case, however, Parsons is committed to the problematic thesis of the identity of indiscernibles and so the impossibility of two primitively distinct but qualitatively identical objects. (For further discussion, see the entry on the identity of indiscernibles.) Most contemporary philosophers agree that objects are not individuated qualitatively, their identity and diversity

being primitive. Max Black's two qualitatively indiscernible spheres are primitively distinct, in virtue of which one has the property of being that very thing and the other lacking that property (see [Black 1953]). Furthermore, it is hard to see why identity properties are not properties of individuals. Suppose, then, that we count individual identity properties like being identical to A as nuclear properties, those properties entering the range of Parsons's restricted comprehension principle. Then the nuclear negations of those properties are also nuclear. But then we can take the set of all objects, construct the individual identity property for each, construct the nuclear negation of each of those properties, and then construct a condition from those properties that, by Parsons's comprehension principle, corresponds to an object. Then there is an object that is distinct from every object that there is, which is a contradiction. It is unclear, then, that the distinction between nuclear and extra-nuclear properties and the restriction of the comprehension principle to nuclear properties solves the problems facing the naive comprehension principle. (For further discussion of Parsons's view, see [Fine 1982, 1984] and [Zalta 1992].)

Earlier I distinguished two versions of sophisticated Meinongianism. The first, based on the distinction between nuclear and extra-nuclear properties, was found lacking. I turn now to the second, based on the distinction between encoding and exemplifying a property, focusing on Zalta's version. Whereas Parsons distinguishes different kinds of properties, restricting the comprehension principle to only nuclear properties in the hope of thereby avoiding the problems plaguing the naive comprehension principle, Zalta distinguishes two different modes of having a property for the same effect. Exemplifying a property is the familiar way in which an individual has a property; it is what I called instantiation above. Obama exemplifies humanity, my chair exemplifies being comfortable, and the fig tree in my backyard exemplifies needing water. What the comprehension principle does is say not what properties object exemplify, in this sense, but rather what properties they encode. So, for any condition C on properties, there is an object that encodes exactly

those properties, which leaves open whether or not those objects also exemplify those properties.

Let's apply this distinction to the problems facing Meinongianism presented earlier in this section. By the comprehension principle, the condition of being a singer determines an object with exactly that property. The object does not exemplify the property of being a singer but rather encodes it. Indeed, exemplifying the property of being a singer requires exemplifying other properties like having a spatial location, having a voice box, etc., all properties our singer does not have, neither encoding nor exemplifying these further properties. That object does exemplify some properties: Like the property of encoding exactly one property. There is no contradiction here, as the singer encodes exactly one property—the property of being a singer—and exemplifies multiple properties, including the property of encoding exactly one property, being self-identical, etc.. More generally, Zalta's comprehension principle correlates sets of properties with objects that encode, not (necessarily) exemplify, those properties. Insofar as the set of properties characterizing an object are not complete, the resulting object will be incomplete with respect to the properties it encodes. But it need not be incomplete with respect to the properties it exemplifies. While our singer encodes neither the property of wearing blue shoes nor the property of not wearing blue shoes, we can say that it exemplifies the property of not wearing blue shoes. Restricting the comprehension principle to the properties encoded also promises to avoid the other threats of contradiction presented above. Recall the logically impossible condition of being self-distinct. The principle of noncontradiction concerns the properties objects exemplify, not the properties objects encode (assuming our Meinongian is going to account for impossible objects). Because an object can encode inconsistent properties without exemplifying them, impossible objects do not violate the principle of noncontradiction. Finally, Russell's worry that a Meinongian comprehension principle generates existent winged horses can be answered. There is an object correlated to the condition of being an existent winged horse, but that object encodes and does not exemplify the property of existing. Being existent can characterize an object without that

object exemplifying existence. So we do not need to worry about an overpopulation of existent beings, as existent beings exemplify existence, which the existent winged horse does not.

A view based on the distinction between encoding and exemplifying avoids the standard objections to Meinongianism while promising to deliver that view's many benefits. The semantics and logic is straightforward and simple and surface forms of the natural language sentences of interest in this article match their deep logical forms. But the ontological costs are evident. There is semantic and logical simplicity at a metaphysical price.

9.4   AN ANTI-MEINONGIAN FIRST-ORDER VIEW

In the previous two sections I discussed views that deny that existence is a property of individuals and views that deny that existence is a universal property. In this section I consider views according to which existence is a universal property of individuals, in the hope of reaping the benefits of both the earlier views. I then explore the interaction between quantifiers, tense operators, modal operators, and a universal, first-order existence predicate in an attempt to expose some difficulties such a view faces.

For both the Meinongian and the proponent of the proposal under consideration, proper names are directly referential and simple sentences in which they occur express singular propositions. However, unlike the Meinongian, a proponent of this view insists that absolutely everything exists. In that case, a sentence like 'Ronald McDonald does not exist' either expresses a fully articulate singular proposition and so is false, as in that case there is a referent of the subject-place singular term which exists, or does not express a truth evaluative proposition at all, as the singular term lacks a semantic content. In neither case is the sentence true. Avoiding that consequence was a primary motivation behind all of the alternative accounts discussed in the previous two sections. Our first challenge, then, is to explain how sentences like 'Ronald McDonald does not exist' are both meaningful and sometimes apparently true without

abandoning the theses that absolutely everything exists and all names are devices of direct reference.

One important suggestion, adopted by Saul Kripke [Kripke 1973], Peter Inwagen [van Inwagen 1977, 1983, 2003], Nathan Salmon [Salmon 1998], David Braun [Braun 1993, 2005], and Amie Thomasson [Thomasson 1999, 2003, 2009], among others, is that seemingly empty names like 'Ronald McDonald' refer to existent, albeit abstract, fictional characters. Fictional characters have both being and existence. As we don't run into them on the street, see them on the bus, or feel them in our beds, given their lack of spatiotemporal location, it is plausible that what a speaker means when she utters the sentence 'Ronald McDonald does not exist' is not the false proposition that that sentence expresses but instead the true proposition that (the fictional character) Ronald McDonald is not a real person or is not concrete. Indeed, this is suggested by the natural amendment, ''Ronald McDonald does not exist; he's a creation of advertisement!'' On this view, then, there are no genuinely true singular negative existentials. All meaningful singular existentials are true and their negations false. We mistakenly take some singular negative existentials to be true because we conflate or do not sharply distinguish existing from being concrete. (Edward Zalta suggests the possibility of reinterpreting his Meinongian object theory along similar lines, trading his primitive nonuniversal, sometimes contingent existence predicate E!x in for a primitive nonuniversal, sometimes contingent concreteness predicate C!x. I discuss this version of object theory view below.) The benefit of this account is the simple semantics of proper names and the sparse metaphysics. The cost is revisionism regarding what we mean when we use apparently true singular negative existentials.

I end this section by briefly discussing issues that arise with the interaction between quantifiers, tense and modal operators, and a universal, first-order existence predicate, as this interaction is the source of another important cost of this account of existence. There are two sets of intuitions that seem to pull in opposite directions. The first concerns the transience and contingency of existence. Things come in and go out of

existence through time. While Plato and Descartes used to exist, they no longer do; when Plato existed, Descartes did not yet exist; and right now the first child to be born in 2150 does not exist, but that individual will. So, it seems, different things exist at different times. Likewise, of the things that in fact exist, some of them might not have and different things—things that in fact do not exist—might have existed instead or in addition. So, it seems, different things exist at different worlds. These intuitions are quite robust. The second set of intuitions concern the ontological status of nonactual and nonpresent objects. Many philosophers are drawn to the thesis of actualism: The thesis that absolutely everything is actual and how an object is simpliciter is how it actually is, unactualized possibilities for an object being in some sense hypothetical ways of being for that object. While less popular, many philosophers accept the temporal analog of actualism, the thesis of presentism, according to which absolutely everything is present and how an object is simpliciter is how it presently is, how an object was and will be are in some sense hypothetical ways of being for that object. These two sets of intuitions combined with the view of existence under consideration in this section lead to difficulties.

Let's begin with the modal problem. There could have been an object distinct from all actually existing objects. For example, I could have had a brother and, given origin essentialism, if I had a brother, he would have been distinct from every actually existing objects, as no actually existing thing could have been my brother. Our intuitions concerning how things might have been lead us to accept this claim as true. But, by the thesis of actualism, absolutely everything is actual and, by our view of existence, exists and so actually exists. So, it seems that actualism and our view of existence are incompatible with the intuitive possibility of there being an object distinct from all existing objects and the intuition that I might have had a brother.

We can regiment the contingency of existence intuition as followings, letting A be the actuality operator, where Aφ is true with respect to a world w under an interpretation I just in case φ is true with respect to the

distinguished world of I: $\Diamond\exists x\neg A\exists y(y=x)$. Call this sentence Alien. The worry is that the truth of Alien carries ontological commitment to merely possible individuals and so the falsity of the thesis of actualism. One way to substantiate this worry is to invoke the Barcan Formula, or one of the mixing axioms proposed for modal operators and quantifiers in Ruth Barcan Marcus's groundbreaking work in quantified modal logic [Marcus 1946], according to which all instances of the sentence $\Diamond\exists x\varphi(x)$ $\rightarrow\exists x\Diamond\varphi(x)$ are logical true. (In the foregoing formulas, $\varphi(x)$ stands for any formula in which the variable x may or may not be free.) We can then transpose the modal and quantifier in Alien to derive $\exists x\Diamond\neg A\exists y(y=x)$. The truth of this second sentence evidently requires that there is something that is not actual, contrary to the dictates of actualism. The Barcan Formula, and in part for this very reason, is controversial and rejected by those that subscribe to a varying domains possible worlds semantics for modal discourse. So this line of argument is not likely to convince everyone that our modal intuitions lead to problems.

There is, however, a second line of argument that does not rest on the validity of the Barcan Formula, relying instead on combining standard truth definitions for quantified and modal sentences in the most straightforward way. Alien is true under an interpretation I just in case there is a world w accessible from the distinguished world of I with an object in its domain that is not in the domain of the distinguished world of I. This is because its truth requires that $\exists x\neg A\exists y(y=x)$ is true with respect to w and so, it is tempting conclude, that there is a witness that satisfies the condition $\neg A\exists y(y=x)$ at w. But that is the rub. If actualism is true, then there is no such witness, although there could have been. If we are realistic about possible worlds semantics, the model theory for modal talk itself does not contain primitive modality, instead containing worlds as points of evaluation and the notion of truth at a world, in which case 'there is' in the above truth recursion does not occur inside the scope of a possibility operator. So, if Alien is true, there is some object o and accessible possible world w such that o satisfies $\neg A\exists y(y=x)$ at w, which seems to run contrary to the thesis of actualism, as that witness does not

actually exist. (For further discussion of this problem and some of the solutions considered below, see the entry on actualism.)

One solution is to abandon actualism and accept that there are merely possible objects. According to this possibilist position, merely possible objects are among the most unrestricted domain of quantification, being constituents of fundamental reality. It is simply false, then, that absolutely everything is actual. While this position deserves serious consideration, I propose to set it aside and explore only actualist solutions. A second solution rests on the Meinongian distinction between being, in the sense of being a member of the most inclusive domain of quantification and discourse, and existing and the claim that there are objects that do not exist. Armed with a Meinongian metaphysics, we can reject Alien as capturing the intuition that existence is contingent, opting instead for the following sentence as doing that, where E!x is the Meinongian logically primitive existence predicate: $\exists x(\neg E!x \land \Diamond E!x)$. The Meinongian can then deny Alien and appeal to the truth of this sentence to explain our intuitions concerning the contingency of existence. Everything is actual, on this view, although some of those actual things do not exist but could have. This solution to our problems, however, is unavailable to a proponent of the view that existence is a universal property of individuals.

Bernard Linsky and Edward Zalta [Linsky and Zalta 1994] present a novel solution to this problem that promises to be consistent with the tenets of the view of existence under consideration in this section. (A similar account is defended by Timothy Williamson [Williamson 1998, 1999, 2000, 2002].) I begin with the intuitive possibility that I have a brother. An entity that encodes the property of being my brother actually exists, but as a nonconcrete object. That object is only contingently nonconcrete; it could have been concrete, and had it been concrete, it would have exemplified the property of being my brother (along with the other properties that it encodes and the necessary consequences thereof). On this view, my possible brother (as well as any alleged merely possible object) actually exists, just as a nonconcrete individual that could have been concrete. The view is both actualistic, as absolutely everything is

actual, and anti-Meinongian, as absolutely everything exists. Note, however, that Alien, the sentence two paragraphs above purported to capture our intuitions concerning the contingency of existence, is false on this view, as every individual is a necessary existent. Instead, our intuition that what there is is contingent is to be explained in terms of the contingency of what is concrete and nonconcrete. So, where C!x is a logically primitive predicate of concreteness, it is the truth of the following sentence, not Alien, that explains those intuitions: $\exists x(\neg C!x \land \Diamond C!x)$. While the explanation of the contingency of existence bears a structural similarity to the Meinongian explanation discussed in the previous paragraph, the metaphysics is importantly different and so the two views should not be collapsed.

Linsky and Zalta's view requires that concreteness is an accidental property. The self-same individual that is nonconcrete (my possible brother, for example) could have been concrete and the self-same individual that is concrete (me, for example) could have been nonconcrete. This is problematic and becomes more problematic, I believe, when one considers its temporal analog. To see, we turn from the modal problem of contingent existents to its temporal analogy, the problem of temporary existents. While one's accounts of the two problems do not need to swing together (Linsky and Zalta, for example, do not offer the temporal analog of their account of contingent existents), as there are differences between alethic modality and temporality, it is useful to consider them as a pair for our purposes. Intuitively things come in and go out of existence; what exists at one time does not exist at another. The temporal analog of Linsky and Zalta's view of contingent existents entails that everything always exists. What there is and what exists at one time is the same as what is and what exists at any other time; the domain of quantification is fixed across all times. What varies from time to time is which of those individuals are concrete. Socrates still exists now, although as a nonconcrete individual, who was concrete in 450 BCE, and similarly for the first child to be born in 2150. This view requires that a thing can survive the change from being nonconcrete to concrete, intuitive generation, and the change from being concrete to nonconcrete, intuitive

destruction. On this view, then, seeming generation and destruction or substantial change are really forms of qualitative change; a change in the quality of concreteness. This runs contrary to the common conception that concreteness is necessary and eternal to any object that instantiates it, the divide between concrete and nonconcrete individuals marking a divide between categories of being that an individual cannot migrate across. Just as one and the same thing cannot go from being an individual to being a property, so too one and the same thing cannot go from being nonconcrete to being concrete. Suppose then that this view of the transience of existence is rejected. What account can be given to ground the asymmetry between the permanence of concreteness and its contingency? The above explanation, that concreteness marks a category of being, explains the permanence of concreteness, but entails its necessity as well. While not conclusive, this suggests that we look for an alternative solution to our problem.

I now turn to the prospects of a theory according to which existence is a universal, genuinely contingent and transient property of individuals. Such a view requires that the domain of quantification varies from world to world and time to time, as everything that is exists but different individuals exist at different possible worlds and at different times. In that case, Alien is true. What, though, of the earlier argument that the truth of Alien is inconsistent with the thesis of actualism? One response rejects the last step, insisting that there are general claims that could have been true, and so are true at some accessible possible world, without there being specific instances of those claims that are true at those accessible worlds. In nonmodal environments, the quantified sentence $\exists x\varphi(x)$ is true just in case there is some witness o that satisfies the condition $\varphi(x)$. The argument that the truth of Aliens requires that there are merely possible individuals quite naturally imports this truth definition to the truth of a quantified sentence at a merely possible world. And that is the step I am proposing rejecting. The model theory for a model language—with its space of possible worlds and individuals populating the domains of possible worlds—contains, like everything else if actualism is true, only actually existing entities. Still, sentences like Alien are true and so there

are merely possible worlds at which $\exists x \neg A \exists y(y=x)$ is true. There is no individual of which $\neg A \exists y(y=x)$ is true in virtue of which the quantified sentence is true, although there would have been had that merely possible world been actual. This is perhaps more clear when we turn from sentences being true with respect to merely possible worlds to propositions being true at merely possible worlds. There is a world w at which the existential proposition [there is something such that there is no actual thing identical to it] is true but there is no singular proposition [there is no actual thing identical to o] true at w as there is no such entity o. Had we been actual, however, then there would have been such an entity and it is this fact that grounds the truth of the existential proposition at w. (See [Adams 1981].) This suggestion, then, runs contrary to the standard semantics for quantificational sentences, which reduces the truth or falsity of existential and universal sentences to the truth or falsity of their instances, when extending that semantics to truth at a world. The suggestion also requires a distinction between truth at a world, in terms of which the model theory for modal operators is given, and truth in a world, which involves the notion of considering what would have been had a nonactual world been actual.

Check Your Progress 1
Note: a) Use the space provided for your answer
b) Check your answers with those provided at the end of the unit

1. What do you know about the Frege and Russell: Existence is not a Property of Individuals?

    _____

    _____

    _____

2. Discuss about the Meinongianism.

    _____

    _____

    _____

9.5   LET US SUM UP

I began by saying that existence raises a number of deep and important problems in metaphysics, philosophy of language, and philosophical logic. I have examined some of those problems and surveyed a number of different accounts of existence. None of the theories surveyed is wholly satisfying and without cost. The first view proposed by Frege and Russell treats existence as a second-order property and assimilates seemingly singular existentials to general existentials. The proposal requires descriptivism, the thesis that ordinary proper names have descriptive equivalences, which many find to be a problematic thesis. The second Meinongian view requires countenancing individuals that do not exist. We have seen the view face challenges in giving coherent and yet informative and compelling individuation principles for nonexistent individuals and all versions of the view suffer from the problem of metaphysical overpopulation. Finally I presented the naive view that existence is a universal property of individuals. That view faced the problem of having to reject the truth of highly intuitively true singular negative existential sentences like 'Ronald McDonald does not exist'. The view also faces difficulties in properly accounting for the interaction of quantifiers and modal and tense operators. Existence remains, then, itself a serious problem in philosophy of language, metaphysics, and logic and one connected to some of the deepest and most important problems in those areas.

Critical exposition is offered of Kripke's actualist interpretation of the meaning of fiction, against the background of his actualist modal metaphysics. Kripke is committed to the proposition that Sherlock Holmes not only does not happen to exist in the actual world, but for that reason cannot possibly exist in any nonactual merely possible world. Difficulties in Kripke's analysis are highlighted, and contrasted with the Meinongian alternative account of the intended objects of fictional discourse as nonexistent objects satisfying the same generic intensional Leibnizian identity criteria, despite their relevant predicational incompletenesses, as any existent entities. Kripke emphasizes the role of pretending in creating and experiencing works of fiction, which is correct

as far as it goes. However, it does not take into account the fact that pretending is itself an intentional relation, and that there is no satisfactory solution in Kripke's lectures as to how pretending that Anna Karenina has actually been named and actually has the other properties associated with her in one novel is different from pretending that Sherlock Holmes has actually been named and actually has different properties associated with him in another novel can be distinguished, without bringing the distinct intended fictional nonexistent Meinongian objects Sherlock Holmes and Anna Karenina into the referential semantic domain by which the meaning of fictional discourse is explained.

9.6 KEY WORDS

Anti-Meinongian: Critical exposition is offered of Kripke's actualist interpretation of the meaning of fiction, against the background of his actualist modal metaphysics. Kripke is committed to the proposition that Sherlock Holmes not only does not happen to exist in the actual world, but for that reason cannot possibly exist in any nonactual merely possible world.

9.7 QUESTIONS FOR REVIEW

1. What do you know the Anti-Meinongian First-Order View?

9.8 SUGGESTED READINGS AND REFERENCES

*   Adams, R., 1981, "Actualism and Thisness," Synthese, 49: 3–41.
*   Aquinas, T., On Being and Essence, trans. Armand Maurer, 1968, Toronto: Pontifical Institute Medieval Studies.
*   Black, M., 1952, "The Identity of Indiscernibles," Mind, 61: 153–164.
*   Braun, D.,1993, "Empty Names," Noûs, 27: 449–469.
*   –––, 2005, "Empty Names, Fictional Names, Mythical Names," Noûs, 39: 596–631.

- Castañeda, H-N., 1975, "Individuation and Non-Identity: A New Look," American Philosophical Quarterly, 12: 131–140.
- –––, 1986, "Objects, Existence, and Reference: A Prolegomenon to Guise Theory," in R. Haller (ed.), Non-Existence and Predication, Amsterdam: Rodopi, 3–60.
- –––, 1989. Thinking, Language, and Experience, Minneapolis: University of Minnesota Press.
- Dancy, R., 1986, "Aristotle and Existence," in Knuuttilla & Hintikka 1986, pp. 49–80.
- Findlay, J., 1963, Meinong's Theory of Objects and Values, Oxford: Oxford University Press.
- Fine, K., 1982, "The Problem of Non-Existents. I. Internalism," Topoi, 1: 97–140.
- –––, 1984, "Critical Review of Parsons' Nonexistent Objects," Philosophical Studies, 45: 94–142.
- Frege, G., 1884, Die Grundlagen der Arithmetik. English translation: J.L. Austin, The Foundations of Arithmetic, Oxford: Blackwell, second revised edition 1974.

9.9 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

1. See Section 9.2
2. See Section 9.3

UNIT 10: THE NECESSITY OF EXISTENCE

STRUCTURE

10.0 OBJECTIVES

After this unit, we can able to know:

- To know about the Stating the Question
- Why Might Someone Believe God Grounds the Existence of Necessarily Existing Abstract Objects?

- God's Grounding Abstract Objects I: Views on Which Necessarily Existing

- God's Grounding Abstract Objects II: Views on Which There Aren't Necessarily Existing Abstract Objects that All Are Grounded in God

10.1 INTRODUCTION

It is commonly accepted that there are two sorts of existent entities: those that exist but could have failed to exist, and those that could not have failed to exist. Entities of the first sort are contingent beings; entities of the second sort are necessary beings. We will be concerned with the latter sort of entity in this article.

There are various entities which, if they exist, would be candidates for necessary beings: God, propositions, relations, properties, states of affairs, possible worlds, and numbers, among others. Note that the first entity in this list is a concrete entity, while the rest are abstract entities. Many interesting philosophical questions arise when one inquires about necessary beings: What makes it the case that they exist necessarily? Is there a grounding for their necessary existence? Do some of them depend on others? If so, how might one understand the dependence relation?

10.2 STATING THE QUESTION

The main question we will address in this article is: Does God ground the existence of necessarily existing abstract objects? It is perhaps a more general question than a question one might at first ask: Did God create necessarily existing abstracta? But it is the main question that philosophers who have written about the relation between God and abstract objects have sought to answer.

Over the last two decades, philosophers have done a great deal of work on the notion of grounding (see, e.g., Fine 2001, Rosen 2010, Audi 2012, Schaffer 2009 Koslicki 2012 and the Stanford Encyclopedia entry on grounding). It is thought by many currently working on issues in the metaphysics of grounding that grounding is a primitive, sui generis

relation. In particular, it is not to be understood as a supervenience or causal relation. How are we then to understand what it is? Philosophers point to particular cases where it is instanced: Dispositional properties are grounded in categorical properties, the mental is grounded in the physical, the semantic is grounded in the non-semantic, features like smiles or surfaces are grounded in facts about bodies, and so on. To this point, one might think that grounding talk can be captured by our ordinary notion of supervenience.[3] But Fine (2001) claims that Socrates' singleton set is grounded in Socrates; yet, necessarily one exists just if the other does. Thus our ordinary modal notion of supervenience won't capture this case of grounding. If we assume (as many in the grounding literature do) that the other cases of grounding are of the same sort as the Socrates-singleton case is, then our ordinary notion of supervenience won't capture them either.

Our discussion of the question of God's grounding the existence of necessarily existing abstracta bears on the general conversation about the nature of grounding. First, we can note that our divine grounding case stands alongside the Socrates-singleton case in showing that ordinary supervenience won't capture the grounding relationship properly. For instance, suppose we say that God grounds the existence of the number 2. We then can note that, necessarily God exists just if 2 does (that is, each exist in every possible world). According to ordinary notions of supervenience, the number 2 supervenes on God, and conversely. But we are to think that God grounds the existence of 2, and not vice versa. Second, we have here in the case of divine grounding of abstracta a case where the grounding relationship is typically spelled out in other, familiar terms (and thus isn't sui generis). As we will see, a number of different philosophers who think that God grounds the existence of necessarily existing abstract objects think that God does so in a causal manner. Others think that the grounding takes place in that necessarily existing abstracta are identical with divine mental states.

One might look at those who claim that God causes necessarily existing abstract objects or that they are identical with divine mental states as not

asserting that God grounds the existence of necessarily existing abstracta. But as we will see, each of these sorts of theorists really is saying that God grounds the existence of necessarily existing abstract objects. Thus, it might be better to cast our lot with those who are skeptical that there is a sui generis grounding relationship that metaphysicians investigate. Or, if there is such a relationship in some cases of grounding, it isn't present in all cases of grounding (it isn't "univocal"—see Hofweber 2009 and Daly 2012 for discussion). After all, it is perfectly sensible to recast "Do necessarily existing abstract objects depend on God?" as "Are necessarily existing abstract objects grounded in God?"

However we think of the dependence relationship between God and necessarily existing abstract objects, we will want to insist that on it God is somehow more fundamental than necessarily existing abstract objects. Fundamentality (Stanford Encyclopedia entry) is an asymmetric relationship. Thus we will construe those who think that God grounds the existence of necessarily existing abstracta as claiming that God is more fundamental than necessarily existing abstracta, and not conversely.

10.3 WHY MIGHT SOMEONE BELIEVE GOD GROUNDS THE EXISTENCE OF NECESSARILY EXISTING ABSTRACT OBJECTS?

There are at least two sorts of reasons why someone might be inclined to think that God grounds the existence of necessarily existing abstract objects. The first sort of reason involves central religious texts in monotheistic faiths like Judaism, Christianity, and Islam. Roughly, this sort of reason consists in these texts' assertions or suggestions that God has created everything. If God created everything, it must be that God has created necessarily existing abstract objects, as well. Thus, God grounds the existence of these abstract objects. For instance, there are statements in the Hebrew Bible such as Psalm 89:11: "The heavens are yours, the earth also is yours; the world and all that is in it—you have founded them". Also in the Hebrew Bible is Nehemiah 9:6:

And Ezra said: "You are the Lord, you alone; you have made heaven, the heaven of heavens, with all their host, the earth and all that is on it, the seas and all that is in them. To all of them you give life, and the host of heaven worships you".

In the New Testament, there are passages like John 1:1–1:4:

In the beginning was the Word, and the Word was with God, and the Word was God. He was in the beginning with God. All things came into being through him, and without him not one thing came into being. (The Word [logos in Greek] to which John refers is Jesus of Nazareth)

Paul states in Colossians 1:15–16,

He [Jesus] is the image of the invisible God, the firstborn of all creation; for in him all things in heaven and on earth were created, things visible and invisible, whether thrones or dominions or rulers or powers—all things have been created through him and for him.

One of the most important documents for Christian faith outside the Hebrew Bible and New Testament, the Nicene Creed of 325, says "We believe in one God, the Father almighty, maker of all things visible and invisible". The Niceno-Constantinopolitan Creed of 381, a modification of the older Nicene Creed of 325 that is used by the western Church begins similarly, "We believe in one God, the Father almighty, maker of heaven and earth, of all things that are visible and invisible".

According to the Qur'an, "God is the Creator of all things; He has charge of everything; the keys of the heavens and earth are His" (39:62–63). The Qur'an also says, "This is God, your Lord, there is no God but Him, the Creator of all things, so worship Him; He is in charge of everything " (6:102).

These reasons from authoritative religious texts may not be taken to be conclusive, however. One may take these sorts of texts seriously as an

adherent to faiths they define and still hold that God doesn't have creative control over necessarily existing abstract objects. For instance, Peter van Inwagen (2009) argues that the universal quantifier in claims like that of the Nicene Creed "maker of all things visible and invisible" is implicitly restricted to include only those things that are capable of being created. Necessarily existing abstract objects cannot enter into causal relations and thus can't be created. But it is worth noting that philosophers who think that if there are necessarily existing abstract objects, God must have some sort of control over them (e.g., Craig 2016) point to texts like those cited above for justification for this view.

There is a second sort of reasoning that may lead someone might think that God grounds the existence of necessarily existing abstract objects. That is by way of perfect being theology (see Morris 1987a, 1987b and Nagasawa 2017 for discussion). Perfect being theology is a way of theorizing a priori about God that goes back at least to Anselm of Canterbury. One begins with the claim that God is the greatest possible being, and from there one can derive attributes that God must have. This method is one way of arriving at God's being omnipotent, omniscient, and perfectly good. Anselm himself famously thought that via perfect being theology he could conclude that God existed. For our purposes here, we are to imagine two conceptually possible beings: One being has grounds or explains the existence of necessarily existing abstract objects, and the other doesn't. We are to see that the being that grounds these abstracta is greater than one who doesn't, and thus we are to conclude that God (the greatest possible being) has control over necessarily existing abstract objects. Sometimes the intuition that the former being is greater than the latter is put in terms of God's aseity, or independence from all other entities. A being with maximal aseity is greater than one without it (other things being equal); and if necessarily existing abstract objects don't depend on God, God lacks maximal aseity.

There likely would be little objection to reasoning to divine grounding of necessarily existing abstracta in the above way, if it were thought that God could have control over these sorts of abstracta. However, someone

might concur with van Inwagen that abstracta can't enter into causal relations, and say that the only way that abstracta might be grounded in God is via causation. Or someone might think that the idea of a necessarily existing object depending on anything is incoherent. If one took either of these positions, one would deny that the being who grounds necessarily existing abstract objects was greater than the one who didn't. (Just as she would deny that a being who can make a square circle is greater than one who can't—there can't be a being who can make a square circle.)

We have noted two sorts of reasons why a theist might think that God grounds the existence of necessarily existing abstract objects. We turn to a discussion of some different answers to our central question: Does God ground the existence of necessarily existing abstract objects? Each of the next two sections will begin with a list of views and will follow with considerations for and against each.

10.4 GOD'S GROUNDING ABSTRACT OBJECTS I: VIEWS ON WHICH NECESSARILY EXISTING ABSTRACTA ARE ALL GROUNDED IN GOD

The views discussed in this section are as follows:

Theistic Voluntarism:

Necessarily existing abstracta are caused to exist by God's will (or some other normally-contingent divine faculty). Example: Descartes.

Theistic Emanationism:

Necessarily existing abstracta are caused to exist by some non-contingent divine faculty (e.g., the right sort of divine cognition). Example: Leibniz, Morris-Menzel (1986).

Theistic Mentalism (without Divine Simplicity):

Necessarily existing abstracta are identical with divine mental states, and God isn't simple. Example: Welty (2014).

Theistic Mentalism (with Divine Simplicity):
Necessarily existing abstract are identical with divine mental states, and God is simple. Examples: Augustine, Aquinas.

10.4.1 Theistic Voluntarism

According to the theistic voluntarist, necessarily existing abstract objects depend on the divine will, or some other contingent feature of God. This is famously the view of Descartes. In a letter to Mersenne (27 May 1630), Descartes says:

You ask me by what kind of causality God established the eternal truths. I reply: by the same kind of causality as he created all things, that is to say, as their efficient and total cause. For it is certain that he is the author of the essence of created things no less than of their existence; and this essence is nothing other than the eternal truths. I do not conceive them as emanating from God like rays from the sun; but I know that God is the author of everything and that these truths are something and consequently that he is their author. (Descartes 1991: 25)

Descartes makes the same sorts of claims in his public writings, as well (e.g., in the reply to the Sixth Set of Objections (also from Mersenne)). This view seems to take seriously that God truly is maximally powerful; he even has volitional control over things like numbers, properties, and states of affairs. Indeed, even more than with views like theistic emanationism is God in control of abstracta on this view. According to the theistic voluntarist, God could have made different—or no—abstracta like propositions, properties, and states of affairs. God is in control of abstracta like God is in control of any other object: Their existence is subject to God's will.

Of course, the seriousness with which the theistic voluntarist takes divine aseity and sovereignty is also the source of problems for the view. If God could have failed to make the number 2, in what sense is 2 a necessary being? One might try to weaken the voluntarist view, by claiming that 2 is only weakly necessary: God had to create it, but it is possibly…possible that it not exist. All worlds that are accessible to the actual world have 2 existing in them. But some of these worlds have a somewhat-different divine will relative to the existence of 2 (maybe God somewhat-reluctantly wills the existence of 2 in them). And possible relative to those sorts of worlds (or relative to worlds possible to those worlds, etc.) are worlds in which God doesn't will that 2 exist. The key here is that the claim: Necessarily, 2 exists comes out true on this picture; 2 exists in every possible world relative to the world of evaluation (the actual world). But there are possibly…possible worlds in which God doesn't will the existence of 2. The voluntarist can say that abstracta depend on the will of God, and yet really do exist necessarily (just don't say in every possible world, full stop (see Plantinga 1980: 95 ff. for further discussion)).

Of course, this suggestion risks two sorts of problems. The first is that it doesn't take divine sovereignty seriously enough. Imagine a being who could—in a world possible relative to the actual world—make it the case that 2 doesn't exist. That being might be thought to be more powerful than a being that only possibly…possibly could do this. And Descartes (in places, at least) seems to have this intuition; and thus plumps for a God who could make it the case that 2 didn't exist.

The second concern is that it abandons S5-type modal logic, in which anything that is necessary is necessarily necessary. This is thought by many to be the appropriate system of modal logic to describe the way actual modality is. So there are concerns from both sides for this reply to the objection to voluntarism. On the one hand, it might be thought not to take divine power seriously enough. On the other, it might not make abstracta "necessary enough".

One reason why Descartes is famous for holding to theistic voluntarism view is because so few others in the history of theological thought hold to it. And perhaps the main reason why no one else holds to it is because many judge that the theistic voluntarist isn't able to account for the absolute necessity of necessarily existing abstracta. These are objects that if they exist should exist in every possible world, full stop; and once one allows for that, it makes it very difficult to see how it could be up to God's will that these exist. Rather, if they are up to God; one winds up with a view like theistic emanationism. We turn to it.

10.4.2 Theistic Emanationism

According to the theistic emanationist, necessarily existing abstract objects are caused to exist by some non-contingent feature of divine activity. The standard feature the emanationist appeals to is divine cognitive activity of some sort. So, the theistic emanationist will say something like that the number 2 exists because of God's cognitive activity. She will go on to say (and this is how the view is distinct from a theistic voluntarist view) it is not possible (it's true in no possible world, full stop) for God's cognitive activity in this respect to be other than it is. Thus, the theistic emanationist can hold that abstracta really exist in every possible world, full stop (allowing that God does, too).

One example of a theistic emanationist is Leibniz. In his Monadology he says:

It is also true that God is not only the source of existences but also that of essences insofar as they are real, that is, the source of that which is real in possibility. This is because God's understanding is the realm of eternal truths, or that of the ideas on which they depend; without him there would be nothing real in possibles, and not only would nothing exist, but also nothing would be possible.

For if there is a reality in essences or possibilities, or indeed in eternal truths, this reality must be grounded in something existent and actual, and

consequently, it must be grounded in the existence of the necessary being, in whom essence involves existence, that is, in whom possible being is sufficient for actual being. (Leibniz 1714 [1989: 218])

Here Leibniz seems to suggest that necessarily existing abstracta are grounded in divine cognitive activity. It's not clear exactly how to characterize the relation between the cognitive activity and the existence of the abstract objects, but saying that the former causes the latter to exist seems appropriate given his language.

Thomas Morris and Christopher Menzel (1986) also are theistic emanationists. They invoke explicitly causal language in setting out their view, which they call "theistic activism".

So our suggestion is that the platonistic framework of reality arises out of a creatively efficacious intellective activity of God. It is in this sense that God is the creator of the framework. It depends on him. (1986: 356)

They continue later:

Let us refer to this view, the view that a divine intellectual activity is responsible for the framework of reality, as "theistic activism". A theistic activist will hold God creatively responsible for the entire modal economy, for what is possible as well as what is necessary and what is impossible. The whole Platonic realm is thus seen as deriving from God. (1986: 356)

And on the next page:

God's creation of the framework of reality…is an activity which is conscious, intentional, and neither constrained nor compelled by anything independent of God and his causally efficacious power. (1986: 357)

Theistic emanationism allows the theist to take seriously the claims of religious documents that God creates everything (indeed, the name of

Morris/Menzel's paper is "Absolute Creation"), and it avoids the problems that beset theistic voluntarism. It has virtues. But it has problems of its own. First, some philosophers claim that God already has to have critical properties in order to be able to cause abstracta to exist. The theistic emanationist claims that God causes properties such as being omniscient, being omnipotent, existing necessarily, being able to cause abstracta to exist, and having cognitive activity to exist. She also claims that God causes his own haecceity, being God, to exist. However, to claim this is to get the dependence relationship backwards, one might charge. Surely, God's being able to cause abstract objects to exist must be posterior to his having properties like the ones mentioned above. And if God has these properties, they must exist. But, the proponent of this theory is committed to the existence of properties being posterior to God's causing them to exist. Thus, the objection concludes, theistic emanationism is false (see Leftow 1990, Davison 1991, Davidson 1999, Bergmann and Brower 2006 for discussion of this sort of objection).

This sort of argument has seemed to many to be decisive. However, there is a response that the theistic emanationist can give at this point. It might be claimed that although God's ability to cause abstracta to exist is logically dependent on his having certain properties, it is not causally dependent. The account would be problematically circular only if God's ability to cause abstracta to exist were causally dependent on his having certain properties, and his having these properties were in turn causally dependent on his having caused these properties to exist. There is a circle of logical dependence here (as there is between any two necessary truths), but there is no circle of causal dependence (see Morris and Menzel for this sort of reply).

The opponent of theistic emanationism might make the following retort. Certainly, the above response is right in that if there is a problem of circularity, it is one of causal circularity. Earlier, we saw that there for the theistic emanationist is a one-way causal relationship between God's cognitive activity and the existence of abstracta such as being the number two. We can say that the necessary existence of being the number two (or

any abstract object) causally depends on God's having the cognitive activity that he does. Or, perhaps we might say that the necessary existence of being the number two causally depends on God's being omniscient, omnipotent and existing necessarily. However, the entities on which being the number two causally depends are themselves properties. On what do they causally depend? It seems that on the emanationist account they wind up causally depending on themselves. But this is incoherent, one might charge.

Even if the emanationist successfully replies to this first problem for the view, there is a second, and perhaps more serious objection to the view. We will call this objection "the bootstrapping objection" (see Leftow 1990, Davidson 1999, Bergmann and Brower 2006, and Gould 2014b for discussion of this sort of objection). We can put the concern this way (following Davidson 1999). To cause something to exist is to cause its essence (or, in the terminology of Plantinga 1980, its nature) to be exemplified. Suppose God creates a certain table which has as a part of its essence being red. Then God causes the property being red to be exemplified by the table when he creates it. Consider the property being omnipotent. The property being exemplified by God is contained in its essence. So, God causes the property being exemplified by God to be exemplified by being omnipotent in causing being omnipotent to exist. Similar to the manner with which God causes being red to be exemplified by the table in exemplifying the table's essence, God causes being omnipotent to be exemplified by himself. But, surely God can't cause the property being omnipotent to be exemplified by himself: How can God make himself omnipotent? Furthermore, one might think that God's omnipotence should be causally prior to his causing properties to exist. However, on this occasion it is not. Then, if one does think that God's omnipotence should be causally prior to his causing properties to exist, this would be an instance of causal circularity. This sort of argument will work for other properties like being omniscient or having divine cognitive activity (although the causal circle may be more difficult to establish with the former, and the implausibility of self-exemplification may be more difficult to establish with the latter).

Furthermore, consider God's haecceity, the property being God. The property being necessarily exemplified is contained in the essence of this property. So, when God causes his haecceity to exist, he causes the property being necessarily exemplified to be exemplified by his haecceity. Just as God causes being red to be exemplified by the table when he causes it to exist, God causes being God to be exemplified necessarily. However, one might well think this incoherent. Indeed, it seems this is the divine causing his own existence: God is pulling himself up by his own bootstraps.

The theistic emanationist needs to address these sorts of concerns about bootstrapping, and it is not clear how that could be done.

10.4.3 Theistic Mentalism (without Divine Simplicity)

One sort of theistic mentalism is the view that necessarily existing abstract objects are divine mental states, and that God isn't simple.[10] On this view, God is distinct from his mental states, and abstracta are identical with these mental states. One proponent of this view is Welty (2014). He says

I maintain that [abstract objects] are constitutively dependent on God, for they are constituted by the divine ideas, which inhere in the divine mind and have no existence outside it…[Abstract objects] are necessarily existing, uncreated divine ideas that are distinct from God and dependent on God. (2014: 81)

Why might someone adopt theistic mentalism? One could make the following sort of case. Thoughts (e.g., sentences in the language of thought) are capable of representing the world as being a particular way. Propositions are capable of representing the world as being a particular way. Why do we need both of these sorts of intentional entities? We can simply identify propositions and thoughts, and we get a simpler ontology.

Of course, there is a problem here. If the thoughts we speak of here are human thoughts, there are continuum-many true propositions, and finitely many human thoughts. Also, there are propositions true in worlds where there are no human thoughts. So we can't identify propositions and human thoughts. But we don't have this problem with divine thoughts. God, we may grant, exists necessarily. And God has sufficiently many mental states to stand in for true propositions (see Plantinga 1980, 1982).

If we identify propositions with divine thoughts, we have enough of them in all possible situations. And one has one fewer kind of thing if one admits only thoughts (divine and otherwise) rather than thoughts and propositions. But there are reasons to think there are both thoughts and propositions and that the two shouldn't be identified, even if one identifies propositions and divine thoughts. The most straightforward reason is that thoughts are a different kind of entity from propositions. The former are concrete, and the latter abstract. Furthermore, it's worth noting the conceptual role propositions play. They are the sorts of things that can be affirmed, doubted, believed, and questioned. They can be true and false, necessary and possible. It is said by some that they are sets of possible worlds; and by others that they are composite entities, made up of properties and relations, and perhaps concrete individuals. It's not at all clear that thoughts, especially divine thoughts, satisfy any of these conceptual roles.

We also should ask about other necessarily existing abstracta. What sorts of mental entities are they? Do they relate to one another, as concrete mental tokens, in the right sort of way such that they mirror the ways that Platonic states of affairs, propositions, properties, relations, and numbers relate to each other?

What these considerations suggest is that theistic mentalism may actually be a sort of nominalism about abstract objects, in the way that Plantinga (2003: ch. 10) says Lewis' (1986) conception of possible worlds is a sort of nominalism about possible worlds. At best, we have concrete things

that play the roles of necessarily existing abstract objects. (And the theistic mentalist has a great deal more work in specifying concrete divine mental particulars such that we have all the requisite role-players among the various sorts of necessarily existing abstract objects. It is presumably not enough to say that propositions are divine thoughts and leave it at that.)

Let us return to the initial motivation for theistic mentalism: There are two sorts of intentional objects (propositions and thoughts), and it would be a simpler metaphysic to identify tokens of the two sorts. To assess this, we must ask if the tokens of the two sorts are enough like each other to be identified. That is, simplicity isn't the only consideration relevant here. After all, Spinoza's metaphysic (necessarily there is one object that is exactly as it is in the actual world) is maximally simple, yet has few proponents within western philosophy. Furthermore, if we are able to explain the intentionality of one of these sorts of entities by its relation to the other, it will seem less mysterious that we have two classes of intentional entities. That is precisely what many want to say about the intentionality of thoughts vis-à-vis that of propositions: Thoughts derive their intentionality by standing in the right sort of relation to propositions. So, the reason why my thought is a thought that grass is green is because it has the propositional content that grass is green. The proposition that grass is green has its intentionality intrinsically.

10.4.4 Theistic Mentalism (with Divine Simplicity)

Theistic mentalism with divine simplicity is the view that necessarily existing abstracta are identical with divine mental states, and that God is simple. Because God is simple, each abstract object is identical with God and thus each other. This is a view held most famously by Augustine and Aquinas. Because this is a mentalist view, the criticisms leveled in the section on theistic mentalism without divine simplicity will apply here. In addition, the person who accepts divine simplicity alongside her divine mentalism also will face criticisms of divine simplicity. Plantinga (1980)

is perhaps the locus classicus of contemporary criticism of divine simplicity. He argues that according to divine simplicity, God is identical with his attributes and has (all of) his attributes essentially. But, he argues, God isn't an attribute; and God has many different attributes. (For discussion more sympathetic to divine simplicity, see Mann 1982; Stump and Kretzmann 1985; Leftow 1990; Stump 1997; Wolterstorff 1991; and Bergmann and Brower 2006.) The sorts of difficulties that Plantinga has raised have seemed decisive to many. (It is beyond the scope of this essay to evaluate them, however.) This isn't to say that they can't be met. But the theistic mentalist who accepts divine simplicity has, prima facie, a great number of difficulties with her view.

10.5 GOD'S GROUNDING ABSTRACT OBJECTS II: VIEWS ON WHICH THERE AREN'T NECESSARILY EXISTING ABSTRACT OBJECTS THAT ALL ARE GROUNDED IN GOD

Theistic Platonism: There are necessarily existing abstract objects, and none of them are grounded in God. Example: van Inwagen (2009).

Theistic Nominalism: There are no necessarily existing abstract objects. Example: Craig (2016).

Mixed View 1: Mentalism-Platonism: Example: Gould and Davis (2014).

Mixed View 2: Anti-Bootstrapping Emanationism: Any abstracta that create "bootstrapping" problems aren't grounded in God. Theistic emanationism is true of the others.

10.5.1 Theistic Platonism

According to the theistic Platonist, there are at least some necessarily existing abstract objects (e.g., propositions, properties, relations, numbers, and states of affairs), and the existence of all of the necessarily existing abstracta is not grounded in God. Peter van Inwagen (2009) is a paradigm case of a theistic Platonist. As we saw earlier, van Inwagen argues that if necessarily existing abstract objects were grounded in God, they would be caused to exist by God. But necessarily existing abstracta can't enter into causal relations. So they aren't grounded in God. He says:

In the end, I can find no sense in the idea that God creates free abstract objects [things like propositions, relations, numbers, properties, etc.], no sense in the idea that the existence of free abstract objects in some way depends on the activities of God. (Recall that, although I believe that all abstract objects are free, that is not a position that I am concerned to defend in this chapter.) And that is because the existence of free abstract objects depends on nothing. Their existence has nothing to do with causation… Causation is simply irrelevant to the being (and the intrinsic properties) of abstract objects (2009: 18).

Van Inwagen takes his most serious challenge to be from religious texts that he, as a Christian, thinks are authoritative. He speaks particularly of the beginning of the Niceno-Constantinopolitan Creed, which begins, "I believe in one God, the Father almighty, maker of heaven and earth, of all that is seen and unseen". He thinks here that the quantifier in "all that is seen and unseen" is restricted to things that are capable of entering into causal relations and thus are capable of being created. As van Inwagen points out, there are other authoritative Christian texts in which a universal quantifier is read in a restricted manner (e.g., Matthew 19:26 "for God all things are possible"; see also Luke 1:37, Mark 10:27). No such passage should be taken as a prooftext for a Cartesian view of omnipotence. Rather, the quantifier is read in a restricted manner. Similarly, the quantifier is restricted in the case of the beginning of the Niceno-Constantinopolitan Creed.

It is worth noting that van Inwagen's argument here is slightly different from one that often occurs around these sorts of texts. Often, there is discussion as to whether the writers of the authoritative texts had in mind things like necessarily existing abstracta (e.g., Wolterstorff 1970: 293; Morris and Menzel 1986: 354). To the person who says that the writers of these texts didn't have, say, structured propositions in mind when they claimed God created everything; it is pointed out that neither did they have in mind (clearly-created) things like quarks and bosons (e.g., Davidson 1999: 278–279). Rather, van Inwagen argues that a text like the beginning of the Niceno-Constantinopolitan Creed has to be read in a restricted manner if it is to avoid asserting impossible propositions. And this seems the right way to go for the theistic Platonist. It is very difficult

to discern the scope of the universal quantifier in the usage of writers from nearly 2000 years ago. (This is apart from questions about the connection between intention and semantic content.)

Van Inwagen's main focus is on authoritative texts like the Niceno-Constantinopolitan Creed. But we noted earlier a second sort of reason for adopting a view on which any necessarily existing abstract objects depend on God. That second sort of reason is perfect-being theology. Again, the line of reasoning is that a being who is such that necessarily existing abstracta depend on it is greater than a being on whom they don't depend. And God is the greatest possible being. It is clear what van Inwagen would say at this point: It's not possible for necessarily existing abstract objects to depend on anything. Thus, being an x such that necessarily existing abstracta depend on x is not a great-making property. That this isn't a great-making property is the sort of thing the theistic Platonist needs to say to the perfect-being defender of divine grounding of necessarily existing abstracta. It must be that these things can't depend on God or anyone else. (One may or may not adopt Inwagen's particular argument that they can't.)

Thus, if the theistic Platonist thinks there are good arguments that necessarily existing abstract objects can't be grounded in God, she will have reason to do two things. First, she will have reason read the relevant universal quantifications in authoritative texts as restricted. Second, if she accepts the sort of reasoning in perfect being theology; she will have reason to insist that being an x such that necessarily existing abstracta depend on x is not a great-making property (any more than being able to create a square circle is).

10.5.2 Theistic Nominalism

The theistic nominalist doesn't think there are necessarily existing abstract objects. She may or may not think if there were necessarily existing abstract objects, they would be grounded in God. For instance, William Lane Craig (2016) who is a theistic nominalist; thinks that were there necessarily existing abstracta, they would have to be grounded in God. (We presumably should count Craig as someone who thinks there are true counterpossibles.) But one can imagine someone who thinks that

if things like numbers and properties did exist, theistic Platonism would be a plausible view to adopt. It is worth noting that very few—if any—of the realists about necessarily existing abstracta who are theists are themselves realists because they are theists. Rather, they are realists about necessarily existing abstracta for other sorts of reasons (e.g., indispensability arguments, arguments that we quantify over them with true sentences, or arguments that true sentences (e.g., that 2+3=5) require them as truthmakers (see Rodriguez-Pereyra (2005) for a defense of truthmakers). Van Inwagen himself believes in necessarily existing abstract objects because he thinks that we are committed to true existential quantifications over them (e.g., van Inwagen 2014: ch. 8). There is nothing in particular the theistic nominalist needs to say qua theist about the existence of necessarily existing abstracta that any other nominalist doesn't need to say. One advantage of theistic nominalism is that it allows one to avoid some of the sorts of difficult maneuvers that those who believe God grounds the existence of necessarily existing abstract objects wind up performing. Another advantage of theistic nominalism is that it allows one, if she wishes, to avoid debates about the semantics of universal quantifiers in ancient religious texts. Of course, theistic nominalism is open only to those who find plausible nominalistic replies to standard arguments for realism about necessarily existing abstracta. Craig himself thinks that he can give replies to these standard arguments for realism (2016: ch. 3, 6, 7).

We turn now to two "mixed views", views on which different types of abstracta stand in different grounding relations to God. Both of them try to bracket abstracta having to do with God (e.g., God's own attributes), and to say that God doesn't ground those. But God grounds all the other necessarily existing abstracta.

10.5.3 Mixed View 1: Mentalism-Platonism

On this view, propositions are identical with divine mental states; and properties and relations not exemplified by God are independent of God, a la theistic Platonism. This is the view of Gould and Davis (2014). They say, "Thus, abstract objects exist in two realms: the divine mind and Plato's heaven " (2014: 61). They decline to say in this particular essay

whether mentalism or Platonism is true of other sorts of abstract objects (e.g., numbers, states of affairs, possible worlds). So what we have in Gould and Davis is an initial sketch of a proposal. They are motivated by bootstrapping worries for theistic activism (itself an emanationist view). They think that they can evade bootstrapping objections by having some abstracta be identical with divine mental states, and having the others not grounded in God. Their own name for this view is "modified theistic activism".

It is perhaps strange that, having started with theistic activism (an emanationist view) and its bootstrapping worries, they wind up with a part-mentalist view. It would seem that they could have kept some abstracta causally grounded in God, and others independent of God (see Mixed View 2, below). Also, this first Mixed View will face objections of the sort faced by theistic Platonism; viz. that it doesn't take divine aseity seriously enough, and that it must read the quantifiers in the relevant religious texts in a restricted manner. Furthermore, it is peculiar that propositions wind up as divine mental states, but properties and relations wind up independent of God. One natural understanding of propositions is that they are structured entities, made up of properties and relations. Another is that they are sets of possible worlds. The former understanding seems unavailable to Gould and Davis, and the latter would seem to involve having sui generis primitive possible worlds identical with divine mental states. But at that point, why not just be a thoroughgoing theistic mentalist? After all, bootstrapping isn't a concern for the mentalist. Furthermore, bootstrapping worries arise with abstracta other than properties. For instance, consider the proposition God is omnipotent. In any possible world it exists, it is true. That is, it has being true as part of its essence. But then, if God causes it to exist, God causes it to be true. So we have the same sorts of bootstrapping concerns as we did with a property like being omnipotent.

10.5.4 Mixed View 2: Anti-Bootstrapping Emanationism

This is a view designed wholly to avoid bootstrapping worries that affect theistic emanationism. It really is a sort of modified theistic activism, and

it may actually be the sort of view Gould and Davis (2014) would like to hold. The idea is this: Ascertain the necessarily existing abstracta that cause bootstrapping problems (e.g., being God, being omnipotent, etc.) for the emanationist. Those exist independently of God. All other necessarily existing abstracta are causally grounded in God in the way the theistic emanationist thinks abstracta are grounded in God.

So far as I can tell, no one holds this first mixed view. Perhaps the reader is thinking that that is because it is obviously ad hoc: The sole motivation for the two classes of abstracta in the theory is avoidance of bootstrapping worries. This is too strong, I think. There is a reason why certain abstracta create bootstrapping problem. That reason is that they have to do with God in a way that other abstracta that don't cause bootstrapping problems don't. So why not say that Leibniz or Menzel and Morris are right about all the non-God related abstracta? To put it another way, why not be an emanationist about all the abstracta one can be an emanationist about— those that don't have to do with God?

Check Your Progress 1

Note: a) Use the space provided for your answer

b) Check your answers with those provided at the end of the unit

1.   What do you know about the Stating the Question?

_____

_____

_____

2.   Why Might Someone Believe God Grounds the Existence of Necessarily Existing Abstract Objects?

_____

_____

_____

10.6 LET US SUM UP

That said, there is at least a whiff of ad-hocness here. The motivation for this theory presumably would be that of perfect being theology. For the

proponent of this first mixed view must think that the quantifiers in relevant religious texts are actually restricted. They aren't as restricted as the theistic Platonist thinks they are. But she will agree with the theistic Platonist that it's false that all (read the quantifier wide open) entities are created by/depend on/grounded in God. It's worth pointing out that it's not clear how to delineate precisely those abstracta that lead to bootstrapping problems and those that don't. The best one can do seems to be to say that those that cause bootstrapping problems don't depend on God, and all others God causes to exist. But presumably for each necessarily existing abstract object, either it gives rise to bootstrapping problems, or it doesn't. So there should be two non-overlapping classes of abstract objects at hand here, even if we're not able to specify more descriptively which abstracta are in which class.

It would be better if the emanationist could find a cogent reply to bootstrapping concerns. But if she can't, she may plump for being an emanationist about all abstracta save those having to do with God.

10.7 KEY WORDS

Existing: Existence is the ability of an entity to interact with physical or mental reality. In philosophy, it refers to the ontological property of being.

Grounding: basic training or instruction in a subject.

10.8 QUESTIONS FOR REVIEW

1. Discuss the God's Grounding Abstract Objects I: Views on Which Necessarily Existing.
2. Discuss God's Grounding Abstract Objects II: Views on Which There Aren't Necessarily Existing Abstract Objects that All Are Grounded in God.

10.9 SUGGESTED READINGS AND REFERENCES

- Adams, Robert Merrihew, 1983, "Divine Necessity", The Journal of Philosophy, 80(11): 741–752. doi:10.2307/2026018
- Audi, Paul, 2012, "A Clarification and Defense of the Notion of Grounding", in Correia and Schnieder 2012: 101–121. doi:10.1017/CBO9781139149136.004
- Augustine, "On Ideas", from Eighty-Three Different Questions, David L. Mosher (trans.), Washington, DC: Catholic University of America Press, 1982.
- –––, City of God, Marcus Dods (trans.), Modern Library, 1993.
- Aquinas, Thomas, [ST], Summa Theologica, Fathers of English Dominican Province (trans.), Encyclopedia Britannica (Great Books of the Western World), 1922 [2013].
- –––, Disputed Questions on Truth, Robert J. Mulligan (trans.), Chicago: H. Regnery, 1954.
- Beaty, Michael (ed.), 1990, Christian Theism and the Problems of Philosophy, Notre Dame, IN: Notre Dame Press.
- Beebee, Helen and Julian Dodd (eds.), 2005, Truthmakers: The Contemporary Debate, Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199283569.001.0001
- Bergmann, Michael and Jeffrey E. Brower, 2006, "A Theistic Argument Against Platonism (and in Support of Truthmakers and Divine Simplicity)", Oxford Studies in Metaphysics, 2: 357–386.
- Chandler, Hugh S., 1976, "Plantinga and the Contingently Possible", Analysis, 36(2): 106–109. doi:10.1093/analys/36.2.106

10.10 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

1. See Section 10.2
2. See Section 10.3

UNIT 11: THE DEFINITION OF DEATH

STRUCTURE

11.0 OBJECTIVES

After this unit 11, students can able to understand:

- To know about the Current Mainstream View: The Whole-Brain Approach.

- To discuss A Progressive Alternative: The Higher-Brain Approach.

- To highlight a Proposed Return To Tradition: An Updated Cardiopulmonary Approach

- To know the Further Possibilities

11.1 INTRODUCTION

The philosophical investigation of human death has focused on two overarching questions: (1) What is human death? and (2) How can we determine that it has occurred? The first question is ontological or conceptual. An answer to this question will consist of a definition (or conceptualization). Examples include death as the irreversible cessation of organismic functioning and human death as the irreversible loss of personhood. The second question is epistemological. A complete answer to this question will furnish both a general standard (and criterion) for determining that death has occurred and specific clinical tests to show whether the standard has been met in a given case. Examples of standards for human death are the traditional cardiopulmonary standard and the whole-brain standard. Insofar as clinical tests are primarily a medical concern, the present entry will not address them.

The philosophical issues concerning the correct definition and standard for human death are closely connected to other questions. How does the death of human beings relate to the death of other living things? Is human death simply an instance of organismic death, ultimately a matter of biology? If not, on what basis should it be defined? Whatever the answers to these questions, does death or at least human death have an essence—either de re or de dicto—entailing necessary and jointly sufficient conditions? Or do the varieties of death reveal only "family resemblance" relations? Are life and death exhaustive categories of those things that are ever animated, or do some individuals fall into an ontological neutral zone between life and death? Finally, how do our deaths relate, conceptually, to our essence and identity as human persons?

For the most part, such questions did not clamor for public attention until well into the twentieth century. (For historical background, see Pernick 1999 and Capron 1999, 120–124.) Sufficient destruction of the brain, including the brainstem, ensured respiratory failure leading quickly to terminal cardiac arrest. Conversely, prolonged cardiopulmonary failure

inevitably led to total, irreversible loss of brain function. With the invention of mechanical respirators in the 1950s, however, it became possible for a previously lethal extent of brain damage to coexist with continued cardiopulmonary functioning, sustaining the functioning of other organs. Was such a patient alive or dead? The widespread dissemination in the 1960s of such technologies as mechanical respirators and defibrillators to restore cardiac function highlighted the possibility of separating cardiopulmonary and neurological functioning. Quite rapidly the questions of what constituted human death and how we could determine its occurrence had emerged as issues both philosophically rich and urgent.

Various practical concerns provided further impetus for addressing these issues. (Reflecting these concerns is a landmark 1968 report published by a Harvard Medical School committee led by physician Henry Beecher (Ad Hoc Committee of the Harvard Medical School 1968).) Soaring medical expenditures provoked concerns about prolonged, possibly futile treatment of patients who presented some but not all of the traditionally recognized indicators of death. Certainly, it would be permissible to discontinue life-supports if these patients were dead. Concurrent interest in the evolving techniques of organ transplantation motivated physicians not to delay unnecessarily in determining that a patient had died. Removing vital organs as quickly as possible would improve the prospect of saving lives. But removing vital organs of living patients would cause their deaths, violating both laws against homicide and the widely accepted moral principle prohibiting the intentional killing of innocent human beings (see the entry on doing vs. allowing harm). To be sure, there were—as there are now—individuals who held that procuring organs from, thereby killing, irreversibly unconscious patients who had consented to donate is a legitimate exception to this moral principle (see the entry on voluntary euthanasia), but this judgment strikes many as a radical departure from common morality. In any event, in view of concerns about the possibility of killing in the course of organ procurement, physicians wanted clear legal guidance for determining when someone had died.

The remainder of this entry takes a dialectical form, focusing primarily on ideas and arguments rather than on history and individuals. It begins with an approach that nearly achieved consensus status after these issues came under the spotlight in the twentieth century: the whole-brain approach. (Most of what are here referred to as "approaches" include a standard and a corresponding definition of death; a few offer more radical suggestions for how to understand human death.) The discussion proceeds, in turn, to the higher-brain approach, to an updated cardiopulmonary approach, and to several more radical approaches. The discussion of each approach examines its chief assertions, its answers to questions identified above, leading arguments in its favor, and its chief difficulties. The entry as a whole is intended to identify the main philosophical issues connected with the definition and determination of human death, leading approaches that have been developed to address these issues, and principal strengths and difficulties of these visions viewed as competitors.

11.2 THE CURRENT MAINSTREAM VIEW: THE WHOLE-BRAIN APPROACH

According to the whole-brain standard, human death is the irreversible cessation of functioning of the entire brain, including the brainstem. This standard is generally associated with an organismic definition of death (as explained below). Unlike the older cardiopulmonary standard, the whole-brain standard assigns significance to the difference between assisted and unassisted respiration. A mechanical respirator can enable breathing, and thereby circulation, in a "brain-dead" patient—a patient whose entire brain is irreversibly nonfunctional. But such a patient necessarily lacks the capacity for unassisted respiration. On the old view, such a patient counted as alive so long as respiration of any sort (assisted or unassisted) occurred. But on the whole-brain account, such a patient is dead. The present approach also maintains that someone in a permanent (irreversible) vegetative state is alive because a functioning brainstem enables spontaneous respiration and circulation as well as certain primitive reflexes.[1]

Before turning to arguments for and against the whole-brain standard, it may be helpful to review some basic facts about the human brain, "whole-brain death" (total brain failure), and other states of permanent (irreversible) unconsciousness. (The most important terms for our purposes appear in italics.) We may think of the brain as comprising two major portions: (1) the "higher brain," consisting of both the cerebrum, the primary vehicle of conscious awareness, and the cerebellum, which is involved in the coordination and control of voluntary muscle movements; and (2) the "lower brain" or brainstem. The brainstem includes the medulla, which controls spontaneous respiration, the reticular activating system, a sort of on/off switch that enables consciousness without affecting its contents (the latter job belonging to the cerebrum), as well as the midbrain and pons.

With these basic concepts in view, it may be easier to contrast various states of permanent unconsciousness. (For a helpful overview, see Cranford 1995.) "Whole-brain death" or total brain failure involves the destruction of the entire brain, both the higher brain and the brainstem. By contrast, in a permanent (irreversible) vegetative state (PVS), while the higher brain is extensively damaged, causing irretrievable loss of consciousness, the brainstem is largely intact. Thus, as noted earlier, a patient in a PVS is alive according to the whole-brain standard. Retaining brainstem functions, PVS patients exhibit some or all of the following: unassisted respiration and heartbeat; wake and sleep cycles (made possible by an intact reticular activating system, though destruction to the cerebrum precludes consciousness); pupillary reaction to light and eyes movements; and such reflexes as swallowing, gagging, and coughing. A rare form of unconsciousness that is distinct from PVS and tends to lead fairly quickly to death is permanent (irreversible) coma. This state, in which patients never appear to be awake, involves partial brainstem functioning. Permanently comatose patients, like PVS patients, can maintain breathing and heartbeat without mechanical assistance.

With this background, we turn to the advantages and disadvantages of the whole-brain approach. First, what considerations favor this approach over

the traditional focus on cardiopulmonary function in determining death? The most prominent and arguably the most powerful case for the whole-brain standard appeals to two considerations: (1) the organismic definition of death and (2) an emphasis on the brain's role as the primary integrator of overall bodily functioning. (Some who regard a general definition of death as unnecessary have focused on consideration (2) in defending the whole-brain standard. Some others, as discussed later, have retained consideration (1) but dropped consideration (2).) An additional consideration that has been influential, yet is logically separable from the other two, is (3) the thesis that the whole-brain standard updates, without replacing, the traditional approach to defining death.

According to the organismic definition, death is the irreversible loss of functioning of the organism as a whole (Becker 1975; Bernat, Culver, and Gert 1981). Proponents of this approach emphasize that death is a biological occurrence common to all organisms. Although individual cells and organs live and die, organisms are the only entities that literally do so without being parts of larger biological systems. (Ideas, cultures, and machines live and die only figuratively; cells and tissues are literally alive but are parts of larger biological systems.) So an adequate definition of death must be adequate in the case of all organisms. What happens when a paramecium, clover, tree, mosquito, rabbit, or human dies? The organism stops functioning as an integrated unit and breaks down, turning what was once a dynamic object that took energy from the environment to maintain its own structure and functioning into an inert piece of matter subject to disintegration and decay. In the case of humans, no less than other organisms, death involves the collapse of integrated bodily functioning.

The whole-brain standard does not follow straightforwardly from the organismic conception of death. One might insist, after all, that a human organism's death occurs upon irreversible loss of cardiopulmonary function. Why think the brain so important? According to the mainstream whole-brain approach, the human brain plays the crucial role of integrating major bodily functions so only the death of the entire brain is

necessary and sufficient for a human being's death (Bernat, Culver, and Gert 1981). Although heartbeat and breathing normally indicate life, they do not constitute life. Life involves integrated functioning of the whole organism. Circulation and respiration are centrally important, but so are maintenance of body temperature, hormonal regulation, and various other functions—as well as, in humans and other higher animals, consciousness. The brain makes all of these vital functions possible. Their integration within the organism is due to a central integrator, the brain.

This leading case for the whole-brain standard, then, consists in an organismic conception of death coupled with a view of the brain as the chief integrator of interdependent bodily functions. Another consideration sometimes advanced in favor of the whole-brain standard positions it as a part of time-honored tradition rather than a departure from tradition. (The argument may be understood either as an appeal to the authority of tradition or as an appeal to the practicality of not departing radically from tradition.) The claim is that the traditional focus on cardiopulmonary function is part and parcel of the whole-brain approach, that the latter does not revise our understanding of death but merely updates it with a more comprehensive picture that highlights the brain's crucial role:

Three organs—the heart, lungs, and brain—assume special significance … because their interrelationship is close and the irreversible cessation of any one very quickly stops the other two and consequently halts the integrated functioning of the organism as a whole. Because they were easily measured, circulation and respiration were traditionally the basic "vital signs." But [they] are simply used as signs—as one window for viewing a deeper and more complex reality: a triangle of interrelated systems with the brain at its apex. [T]he traditional means of diagnosing death actually detected an irreversible cessation of integrated functioning among the interdependent bodily systems. When artificial means of support mask this loss of integration as measured by the old methods, brain-oriented criteria and tests provide a new window on the same phenomena (President's Commission 1981, 33).

According to this view, when the entire brain is nonfunctional but cardiopulmonary function continues due to a respirator and perhaps other life-supports, the mechanical assistance presents a false appearance of life, concealing the absence of integrated functioning in the organism as a whole.

The whole-brain approach clearly enjoys advantages. First, whether or not the whole-brain standard really incorporates, rather than replacing, the traditional cardiopulmonary standard, the former is at least fairly continuous with traditional practices and understandings concerning human death. Indeed, current law in the American states incorporates both standards into disjunctive form, most states adopting the Uniform Determination of Death Act (UDDA) while others have embraced similar language (Bernat 2006, 40). The UDDA states that "… an individual who has sustained either (1) irreversible cessation of circulatory and respiratory functions, or (2) irreversible cessation of all functions of the entire brain, including the brainstem, is dead," (President's Commission 1981, 119). Similar legal developments have occurred in Canada (Law Reform Commission of Canada 1981; Canadian Congress Committee on Brain Death 1988). The close pairing of the whole-brain and cardiopulmonary standards in the law suggests that the whole-brain standard does not depart radically from tradition.

The present approach offers other advantages as well. For one, the whole-brain standard is prima facie plausible as a specification of the organismic definition of death in the case of human beings. Moreover, acceptance of whole-brain criteria for death facilitates organ transplantation by permitting a declaration of death and retrieval of still-viable organs while respiration and circulation continue, with mechanical assistance, in a "brain-dead" body. Another practical advantage is permitting, without an advance directive or proxy consent, discontinuation of costly life-support measures on patients who have incurred total brain failure. While most proponents of the whole-brain approach insist that such practical advantages are merely fortunate consequences of the biological facts about death, one might regard these advantages as part of the justification

for a standard whose defense requires more than appeals to biology (see subsection 4.2 below).

The advantages proffered by this approach contributed to its widespread social acceptance and legal adoption in the last few decades of the 20th century. As mentioned, every American state has legally adopted the whole-brain standard alongside the cardiopulmonary standard as in the UDDA. It is worth noting, however, that a close cousin to the whole-brain standard, the brainstem standard, was adopted by the United Kingdom and various other nations. According to the brainstem standard—which has the practical advantage of requiring fewer clinical tests—human death occurs at the irreversible cessation of brainstem function. One might wonder whether a person's cerebrum could function—enabling consciousness—while this standard is met, but the answer is no. Since the brainstem includes the reticular activating system, the on/off switch that makes consciousness possible (without affecting its contents), brainstem death entails irreversible loss not only of unassisted respiration and circulation but also of the capacity for consciousness. Importantly, outside the English-speaking world, many or most nations, including virtually all developed countries, have legally adopted either whole-brain or brainstem criteria for the determination of death (Wijdicks 2002). Moreover, most of the public, to the extent that it is aware of the relevant laws, appears to accept such criteria for death (ibid). Opponents commonly fall within one of two main groups. One group consists of religious conservatives—and, recently, a growing number of secular academics—who favor the cardiopulmonary standard, according to which one can be brain-dead yet alive if (assisted) cardiopulmonary function persists. The other group consists of those liberal intellectuals who favor the higher-brain standard (to be discussed), which, notably, has not been adopted by any jurisdiction.

The widespread acceptance in the U.S. of the whole-brain standard and the broader international acceptance of some sort of "brain death" criteria—whether whole-brain or brainstem—are remarkable considering the subtlety of issues surrounding the definition and determination of

death. Yet this near-consensus has been broader than it is deep. Increasingly, both in academic and clinical circles, doubts about "brain death" are being voiced. Following are several major challenges to the whole-brain standard—and, implicitly, to the brainstem standard. (Several additional challenges are implicit in arguments supporting the higher-brain approach.)

The first challenge is directed at proponents of the whole-brain approach who claim that its standard merely updates, without replacing, the traditional cardiopulmonary standard. A major contention that motivates this thesis is that irreversible cessation of brain function will quickly lead to irreversible loss of cardiopulmonary function (and vice versa). But extended maintenance on respirators of patients with total brain failure has removed this component of the case for the whole-brain standard (PCB 2008, 90). The remaining challenges to the whole-brain approach are not specifically directed to those who assert that its standard merely updates the traditional cardiopulmonary standard.

First, in the case of at least some members of our species, total brain failure is not necessary for death. After all, human embryos and early fetuses can die although, lacking brains, they cannot satisfy whole-brain criteria for death (Persson 2002, 22–23). An advocate could respond by introducing a modified definition: In the case of any human being in possession of a functioning brain, death is the irreversible cessation of functioning of the entire brain. While this may be practically useful in the world as we know it for the foreseeable future, this definition is not conceptually satisfactory if it is possible in principle for some human beings with brains (that is, who have functioning brains at any point in their existence) to die without destruction of their brains. The "in principle" is important here, for this is not possible in our world currently. But suppose we develop the ability to transplant brains. (The thought-experiment that follows appears in McMahan, 429.) Recall that the whole-brain standard is generally thought to receive support from an organismic definition of death. But such a conception of human death, one could argue, only makes sense on the assumption that we are

essentially human organisms (see discussion of the essence of human persons in section 2.1)—as some proponents explicitly acknowledge (see, e.g., Olson 1997). According to the present critique, the brain is merely a part of the organism. Suppose the brain were removed from one of us, and kept intact and functioning, perhaps by being transplanted into another, de-brained body. Bereft of mechanical assistance, the body from which the brain was removed would surely die. But this body was the living organism, one of us. So, although the original brain continues to function, the human being, one of us, would have died. Total brain failure, then, is not strictly necessary for human death. A possible rebuttal to this challenge from one who accepts that we are essentially organisms is to argue that the existence of a functioning brain is sufficient for the continued existence of the organism (van Inwagen 1990, 173–174, 180–181). If so, then in the imagined scenario the original human being would survive the brain transplant in a new body. Thus, the rebuttal concludes, it is false that a human being could die although her brain continued to live.

Perhaps more threatening to the whole-brain approach is the growing empirical evidence that total brain failure is not sufficient for human death —assuming the latter is construed, as whole-brain advocates generally construe it, as the breakdown of organismic functioning mediated by the brain. Many of our integrative functions, according to the challenge, are not mediated by the brain and can therefore persist in individuals who meet whole-brain criteria for death by standard clinical tests. Such somatically integrating functions include homeostasis, assimilation of nutrients, detoxification and recycling of cellular wastes, elimination, wound healing, fighting of infections, and cardiovascular and hormonal stress responses to unanesthetized incisions (for organ procurement); in a few cases, brain-dead bodies have even gestated a fetus, matured sexually, or grown in size (Shewmon 2001; Potts 2001). It has been argued that most brain functions commonly cited as integrative merely sustain an existing functional integration, suggesting that the brain is more an enhancer than an indispensable integrator of bodily functions (Shewmon 2001). Moreover, several studies have demonstrated that most patients diagnosed as brain dead continue to exhibit some brain functions

including the regulated secretion of vasopressin, a hormone critical to maintaining a body's balance of salt and fluid (Halevy 2001). This hormonal regulation is a brain function that represents an integrated function of the organism as a whole (Miller and Truog 2010).

Another, related problem for the sufficiency of total brain failure for human death arises from reflection on locked-in syndrome. People with locked-in syndrome are conscious, and therefore alive, but completely paralyzed with the possible exception of their eyes. With intensive medical support they can live. The interesting fact for our purposes is that some patients with this syndrome exhibit no more somatic functioning integrated by the brain than some brain-dead individuals. Whatever integration of bodily functions remains is maintained by external supports and by bodily systems other than the brain, which merely preserves consciousness (Bartlett and Youngner 1988, 205–6). If total brain failure is supposed to be sufficient for death, and if this is true only because the former entails the loss of somatic functioning integrated by the brain, then the loss of those functions should also be sufficient for death. But these patients, who are clearly alive, show that this is not so. Either the whole-brain definition must be rejected or this particular reason for accepting the whole-brain approach must be rejected and some other good reason for accepting it found.

Recently, a new rationale—distinct from the one that understands human death in terms of loss of organismic functioning mediated by the brain—has been advanced in support of the whole-brain standard (PCB 2008, ch. 4). According to this rationale, a human being dies upon irreversibly losing the capacity to perform the fundamental work of an organism, a loss that occurs with total brain failure. The fundamental work of an organism is characterized as follows: (1) receptivity to stimuli from the surrounding environment; (2) the ability to act upon the world to obtain, selectively, what the organism needs; and (3) the basic felt need that drives the organism to act as it must to obtain what it needs and what its receptivity reveals to be available (ibid, 61). According to a sympathetic reading of the ambiguous discussion in which this analysis is advanced,

any patient who meets even one of these criteria is alive and therefore not dead. A patient with total brain failure meets none of these criteria, even if a respirator permits the continuation of cardiopulmonary function. By contrast, PVS patients meet at least the second criterion through spontaneous respiration (a kind of acting upon the world to obtain what is needed: oxygen); and locked-in patients meet the first criterion if they can see or experience bodily sensation and certainly meet the third insofar as they are conscious. One difficulty with this "fundamental work" rationale for the whole-brain standard, a rationale that is intended to capture "what distinguishes every organism from non-living things" (ibid), is that present-day robots, which are certainly not alive, seem to satisfy the first two criteria. If one insisted, contrary to the reading deemed sympathetic, that a being must satisfy all three criteria—as robots do not since they lack felt needs—in order to qualify as living, the same may be asserted not only of insentient animal life but also of presentient human fetuses and of unconscious human beings of any age. Another difficulty of the "fundamental work" rationale for the whole brain standard is that it was intended to replace the idea that integrated functional unity within an organism is what constitutes life—but the latter idea is extremely plausible and helps to explain what any "fundamental work" would be working toward (cf. Thomas, 105). Whether any variation or modification of the present rationale for the whole-brain standard can survive critical scrutiny remains an open question.

Some traditional defenders of the cardiopulmonary approach believe that the insufficiency of whole-brain criteria for death is evident not only in exceptional cases, such as those described earlier, but in all cases in which patients with total brain failure exhibit respirator-assisted cardiopulmonary function. Anyone who is breathing and whose heart functions cannot be dead, they claim. The champion of whole-brain criteria may retort that such a body is not really breathing and circulating blood; the respirator is doing the work. The traditionalist, in response, will likely contend that what is important is not who or what is powering the breathing and heartbeat, just that they occur. Even complete dependence on external support for vital functions cannot entail that one is dead, the

traditionalist will continue, as is evident in the fact that living fetuses are entirely dependent on their mothers' bodies; nor can complete dependence on mechanical support entail that one is dead, as is evident in the fact that many living people are utterly dependent on pacemakers.

A third major criticism of the whole-brain approach—at least in its legally authoritative formulation in the United States—concerns its conceptual and clinical adequacy. The whole-brain standard, taken at its word, requires for human death permanent cessation of all brain functions. Yet many patients who meet routine clinical tests for this standard continue to have minor brain functions such as electroencephalographic activity, isolated nests of living neurons, and hypothalamic functioning (see, e.g., Potts 2001, 482; Veatch 1993, 18). Indeed, the latter, which controls neurohormonal regulation, is indisputably an integrating function of the brain (Brody 1999, 73). Now one could maintain the coherence of the whole-brain approach by insisting that the individuals in question are not really dead and that physicians ought to use more thorough clinical tests before declaring death (see, e.g., Capron 1999, 130–131). But whole-brain theorists tend to agree that these individuals are dead—that the residual functions are too trivial to count against a judgment of death (see, e.g., President's Commission 1981, 28–29; Bernat 1992, 25)—suggesting that the problem lies with the formulation of the whole-brain standard rather than with its spirit.

Within this spirit and in response to this challenge, a leading proponent of the whole-brain approach has revised both (1) the organismic definition of death to "the permanent cessation of the critical functions of the organism as a whole" and (2) the corresponding standard to permanent cessation of the critical functions of the whole brain (Bernat 1998, 17). The organism's critical functions may be identified by reference to its emergent functions—that is, properties of the whole organism that are not possessed by any of its component parts—as follows: "The irretrievable loss of the organism's emergent functions produces loss of the critical functioning of the organism as a whole and therefore is the death of the organism," (Bernat 2006, 38). The emphasis on critical functions, of course, allows

one to declare dead those patients with only trivial brain functions. According to this revised whole-brain approach, the critical functions of the organism are (1) the vital functions of spontaneous breathing and autonomic circulation control, (2) integrating functions that maintain the organism's homeostasis, and (3) consciousness. A human being dies upon losing all three. Whether this or some similar modification of the whole-brain approach adequately addresses the present challenge is a topic of ongoing debate (see, e.g., Brody 1999, Bernat 2006). What seems reasonably clear is that not all functions of the brain will count equally in any cogent defense of the whole-brain approach.

The judgment that some brain functions are trivial in this context invites a reconsideration of what is most significant about what the human brain does. According to an alternative approach, what is far and away most significant about human brain function is consciousness.

11.3 A PROGRESSIVE ALTERNATIVE: THE HIGHER-BRAIN APPROACH

According to the higher-brain standard, human death is the irreversible cessation of the capacity for consciousness. "Consciousness" here is meant broadly, to include any subjective experience, so that both wakeful and dreaming states count as instances. Reference to the capacity for consciousness indicates that individuals who retain intact the neurological hardware needed for consciousness, including individuals in a dreamless sleep or reversible coma, are alive. One dies on this view upon entering a state in which the brain is incapable of returning to consciousness. This implies, somewhat radically, that a patient in a PVS or permanent coma is dead despite continued brainstem function that permits spontaneous cardiopulmonary function. Although no jurisdiction has adopted the higher-brain standard, it enjoys the support of many scholars (see, e.g., Veatch 1975; Engelhardt 1975; Green and Wikler 1980; Gervais 1986; Bartlett and Youngner 1988; Puccetti 1988; Rich 1997; and Baker 2000). These scholars conceptualize, or define, human death in different ways—though in each case as the irreversible loss of some property for which the

capacity for consciousness is necessary. This discussion will consider four leading argumentative strategies in support of the higher-brain approach.

11.3.1 Appeals to the Essence of Human Persons

One strategy for defending the higher-brain approach is to appeal to the essence of human persons on the understanding that this essence requires the capacity for consciousness (see, e.g., Bartlett and Youngner 1988; Veatch 1993; Engelhardt 1996, 248; Rich 1997; and Baker 2000, 5). "Essence" here is intended in a strict ontological sense: that property or set of properties of an individual the loss of which would necessarily terminate the individual's existence. From this perspective, we human persons—more precisely, we individuals who are at any time human persons—are essentially beings with the capacity for consciousness such that we cannot exist at any time without having this capacity at that time. We go out of existence, it is assumed, when we die, so death involves the loss of what is essential to our existence.

Unfortunately, the use of terminology in these arguments can be confusing because the same term may be used in different ways and terms are frequently used without precise definition. It is sometimes claimed, for example, that we are essentially persons. But what, exactly, is a person? Some authors (e.g., Engelhardt 1996, Baker 2000) use the term to refer to beings with relatively complex psychological capacities such as self-awareness over time, reason, and moral agency. Then the claim that we are essentially persons implies that we die upon losing such advanced capacities. But this means that at some point during the normal course of progressive dementia the demented individual dies—upon losing complex psychological capacities, however these are defined—despite the fact that a patient remains, clearly alive, with the capacity for (basic) consciousness. This view is extraordinarily radical and appears inconsistent with the higher-brain approach, which equates death with the permanent loss of the capacity for (any) consciousness. A proponent of the view that we are essentially persons in the present sense, however, may hold that practical considerations—such as the impossibility of

drawing a clear line between sentient persons and sentient nonpersons, and the potential for abuse of the elderly—recommend the capacity for consciousness as the only safe line to draw, thereby vindicating the higher-brain view (Engelhardt 1996, 250). Meanwhile, other proponents of the view that we are essentially persons (e.g., Bartlett and Youngner 1988) apparently hold that any member of our species who retains the capacity for consciousness qualifies as a person. This view, unlike the previous one, straightforwardly supports the higher-brain standard. Still other authors (e.g., Veatch 1993) hold that we are essentially human beings where this term refers not to all members of our species but just to those judged to be persons by the previous group of authors: members of our species who have the capacity for consciousness. And some authors who defend the higher-brain standard (e.g., McMahan 2002) assert that we are essentially minds or minded beings, which is to say beings with the capacity for consciousness. In each case, an appeal to our essence is advanced to support the higher-brain standard.

Taking this collection of arguments together, the reasoning might be reconstructed as follows:

For humans, the irreversible loss of the capacity for consciousness entails (is sufficient for) the loss of what is essential to their existence;
For humans, loss of what is essential to their existence is (is necessary and sufficient for) death;
Therefore,

For humans, irreversible loss of the capacity for consciousness entails (is sufficient for) death.
We have noted that various commentators who advance this reasoning hold that we are essentially persons in a sense requiring complex psychological capacities. We have noted that this implies that for those of us who become progressively demented, we die—go out of existence—at some point during the gradual slide to permanent unconsciousness. Even if practical considerations recommend safely drawing a line at irreversible loss of the capacity of consciousness for policy purposes, the implication

that, strictly speaking, we go out of existence during progressive dementia will strike many as incredible. At the other end of life there is another problematic implication. For if we are essentially persons (in this sense), then inasmuch as human newborns lack the capacities that constitute personhood, each of us came into existence after what is ordinarily described as his or her birth.

For those attracted to the general approach of understanding our essence in terms of psychological capacities, a promising alternative thesis is that we are essentially beings with the capacity for at least some form of consciousness who die upon irreversibly losing that very basic capacity. Stated more simply, we are essentially minded beings, or minds, and we die when we completely "lose our minds." (Note that this thesis is consistent with the claim that we are also essentially embodied.)

What, then, about the human organism associated with one of us minded beings? Surely the fetus that gradually developed prior to the emergence of sentience or the capacity for consciousness—that is, prior to the emergence of a mind—was alive. On the other end of life, a patient in a PVS who is spontaneously breathing, circulating blood, and exhibiting a full range of brainstem reflexes appears to be alive. Consider also anencephalic infants, who are born without cerebral hemispheres and never have the capacity for consciousness: They, too, seem to be living organisms, their grim prognosis notwithstanding. In response to this challenge, a proponent of the higher-brain approach may either (1) assert that the presentient fetus, PVS patient, and anencephalic infant are not alive despite appearances (Puccetti 1988) or (2) allow that these organisms are alive but are not of the same fundamental kind as we are: minded beings (McMahan 2002, 423–6). Insofar as life is a biological concept, and the organisms in question satisfy commonly accepted criteria for life, option (1) seems at best hyperbolic. At best, the claim is really that these organisms, though alive, are not alive in any state that matters much, so we may count them as dead or nonliving for our purposes. This claim, in turn, may be understood as depending on option (2), on which we may focus. This option implies that for each of us minded beings,

there is a second, closely associated being: a human organism. The prospects of the present strategy for defending the higher-brain approach turn significantly on its ability to make sense of this picture of two closely associated beings: (1) the organism, which comes into existence at conception or shortly thereafter (perhaps after twinning is no longer possible) and dies when organismic functioning radically breaks down, and (2) the minded being, who comes into existence when sentience emerges and might—in the event of PVS or permanent coma—die before the organism does. (For doubts on this score, see DeGrazia 2005, ch. 2).

Appealing to the authority of biologists and common sense, some philosophers (e.g., Olson 1997) charge as indefensible the claim that we (who are now) human persons were never presentient fetuses. One might also find puzzling the thesis that there is one definition of death, appealing to the capacity for consciousness, for human beings or persons and another definition, appealing to organismic functioning, for nonhuman animals and the human organisms associated with persons. It is open to the higher-brain theorist, however, to allow that there are also two closely associated beings in the case of sentient nonhuman animals—the minded being and the organism—with the death of, say, Lassie (the minded dog) occurring at her irreversible loss of consciousness (McMahan 2002, ch. 1). But some will find unattractive the failure to furnish a single conception of death that applies to all living things. To be sure, not everyone finds these objections compelling.

One of the most significant challenges confronting the present approach is to characterize cogently the relationship between one of us and the associated human organism. The relationship is clearly not identity—that is, being one and the same thing—because the organism originates before the mind, might outlive the mind, and therefore has different persistence conditions. This strongly suggests, perhaps surprisingly, that we human persons are not animals. If you are not identical to the human organism associated with you, then since there is at most one animal sitting in your chair, you are not she and are therefore not an animal (Olson 1997). Yet many consider it part of educated common sense that we are animals.

Notes

Might you be part of the organism associated with you—namely, the brain (more precisely, the portions of the brain associated with consciousness) (McMahan 2002, ch. 1)? But the brain seems capable of surviving death, when you are supposed to go out of existence. Are you then a functioning brain, which goes out of existence at the irreversible loss of consciousness? But it seems odd to identify the functioning brain—as distinct from the brain—as you. How could you be some organ only when it functions? Presumably you are a substance (see the entry on substance), a bearer of properties, not a substance only when it has certain properties. One might reply that the functioning brain is itself a substance, a substance distinct from the brain, but that, too, strains credibility. Might you instead be not the brain, but the mind understood as the conscious properties of the brain? That would imply that you are a set of properties, rather than a substance, which is no less counterintuitive. Note that the charge of incredibility is not directed at the assertion that the mind is the functioning brain, or is a set of brain properties, and not a distinct substance—a thesis in good standing in the philosophy of mind (see the entries on identity theory of mind and functionalism). The charge of incredibility is directed at the assertion that you are a set of properties and not a substance.

Another possibility regarding the person/organism relationship is that the human organism constitutes the person it eventually comes to support (Baker 2000). One might even claim the legitimacy of saying—employing an "is" of constitution—that we are animals (or organisms), just as we can say that a statue constituted by a hunk of bronze, shaped in a particular way, is a hunk of bronze (ibid). Challenges to this reasoning includes doubts that we may legitimately speak of an "is" of constitution; if not, then the constitution view implies that we are not animals after all. Another challenge, which applies equally to the view that we minds are parts of organisms, concerns the counting of conscious beings. On either the constitution view or the part-whole view, you are essentially a being with the capacity for consciousness. Closely associated with you—without being (identical to) you, due to different persistence conditions—

is a particular animal. But that animal, having a functioning brain, would also seem to be a conscious being. Either of these views, then, apparently suggests that for each of us there are two conscious beings, seemingly one too many. Despite such difficulties as these, the thesis that we are essentially minded beings remains a significant basis for the higher-brain approach to human death.

11.3.2 Appeals to Personal Identity

A second argumentative strategy in defense of the higher-brain approach claims to appeal to our personal identity while remaining agnostic on the question of our essence (Green and Wikler 1980). The fundamental claim is that, whatever we are essentially, it is clear that one of us has gone out of existence once the capacity for consciousness has been irreversibly lost, supporting the higher-brain standard of death. Clearly, though, any view of our numerical identity over time—our persistence conditions—is conceptually dependent on a view of what we essentially are (DeGrazia 1999; DeGrazia 2005, ch. 4). If we are essentially human animals, and not essentially beings with psychological capacities, then, contrary to the above argument, it is not clear—indeed, it is false—that we go out of existence upon irreversible loss of the capacity for consciousness; rather, we die upon the collapse of organismic functioning. The appeal to personal identity in support of the higher-brain standard depends on the thesis that we are essentially minded beings and therefore inherits the challenges facing this view, as discussed in the previous subsection. Nevertheless, the appeal to personal identity, construed as a distinct argumentative strategy, has been somewhat influential (see, e.g., President's Commission 1981, 38–9).

11.3.3 The Claim that the Definition of Death is a Moral Issue

Another prominent argumentative strategy in support of the higher-brain approach contends that the definition of death is a moral issue and that confronting it as such vindicates the higher-brain approach (see, e.g., Veatch 1975, 1993; Gervais 1986, ch. 6). In asking how to determine that

a human has died, according to this argument, what we are really asking is when we ought to discontinue certain activities such as life-support efforts and initiate certain other activities such as organ donation, burial or cremation, grieving, change of a survivor's marital status, and transfer of property. The question, in other words, is when "death behaviors" are appropriate. This, the argument continues, is a moral question, so an answer to this question should be moral as well. Understood thus, the issue of defining human death is best addressed with the recognition that irreversible loss of the capacity for consciousness marks the time at which it is appropriate to commence death behaviors.

Is the definition of death really a moral issue? To say that someone has died does seem tantamount to saying that certain behaviors are now appropriate while certain others are no longer appropriate. But it hardly follows that the assertion of death is itself a moral claim. An alternative hypothesis is that the sense of moral import derives from the fact that certain moral premises—for example, that we shouldn't bury or cremate prior to death—are shared by virtually everyone. Moreover, the concept of death is (at least originally) at home in biology, which offers many instances in which a determination of death—say, of a gnat or a clover— seems morally unimportant. Rather than asserting that death itself is a moral concept, it might be more plausible to assert that death, a biological phenomenon, is generally assumed to be morally important—at least in the case of human beings—given a relatively stable background of social institutions and attitudes about "death behaviors." Furthermore, due to the moral salience of human death, discussions about its determination are often prompted by a moral or pragmatic agenda such as interest in organ transplantation or concerns about expensive, futile treatment. But these observations do not imply that death is itself a moral concept.

Even if it were, it would hardly follow that the higher-brain standard is preferable to other standards. A person with relatively conservative instincts might hold that death behaviors are morally appropriate only when the whole-brain or cardiopulmonary standard has been met. We need to ask, therefore, what grounds exist for the claim—advanced by

proponents of the higher-brain standard—that death behaviors are appropriate as soon as someone has irreversibly lost the capacity for consciousness. Perhaps the best possible grounds are that irreversible loss of consciousness entails an existence lacking in value for the unconscious individual herself. It appears, then, that the strongest specification of the present line of reasoning actually relies upon the next (and final) argumentative strategy to be considered—and may, as we will see, lead to the conclusion that we should permit individuals to select among several standards of death.

11.3.4 The Appeal to Prudential Value

The idea here is to defend the higher-brain approach on the basis of claims about prudential value (for a discussion, see DeGrazia 2005, 134–8). Conscious life, it is argued, is a precondition for virtually everything that we value in our lives. We have an enormous stake in continuing our lives as persons and little or no stake in continuing them when we are permanently unconscious. The capacity for consciousness is therefore essential not in a metaphysical sense connected to our persistence conditions, but in the evaluative sense of indispensable to us. One need not claim that the capacity for consciousness underlies everything of prudential value, just that it underlies the overwhelmingly greater part of what matters to us prudentially. And although, for many people, consciousness may not be sufficient for what matters prudentially—insofar as they find indispensable, say, some degree of self-awareness and meaningful interaction with others—it is certainly necessary; and the basic capacity for consciousness (as opposed to self-consciousness or personhood) is the only safe place to demarcate death for policy and social purposes. We should therefore regard irreversible loss of the capacity for consciousness as a human being's death—even if the original concept of death is biological and biological considerations favor some less progressive standard.

How persuasive is this case for the higher-brain approach? One might challenge the assumption that prudential, as opposed to moral, considerations ought to be decisive in adopting a standard for human

death. On the other hand, as suggested in our discussion of the previous argumentative strategy, moral considerations may not favor a particular standard of death except insofar as they rest on prudential considerations—our present concern. But even if we accept the claim that human death should be understood on the basis of prudential values, we confront the prospect of reasonable pluralism about prudential value. While supporters of the higher-brain approach (who tend to be liberal intellectuals) are likely to have prudential values in line with this approach, many other people do not. If a patient has a stake in his family's need for closure should he enter a PVS—an interest that may be self-regarding as well as other-regarding—this fact would count against allowing the PVS to constitute death in his case. If an Orthodox Jew or conservative Christian believes that (biological) life is inherently precious to its possessor, even if the individual cannot appreciate its value at a given time, this would count against the higher-brain standard in the case of the individual in question. Perhaps, then, the appeal to prudential value favors not the higher-brain standard for everyone but a pro-choice view about standards of death. A jurisdiction might, for example, have one default standard of death but permit conscientious exemption from that standard and selection of a different one within some reasonable range of options.

In reply to this argument, a proponent of the appeal to prudential value might contend that it is simply irrational to value biological existence without the possibility of returning to consciousness. But this reply assumes the experience requirement: that only states of affairs that affect one's experience can affect one's well-being (for a discussion, see Griffin 1986, 16–19). The experience requirement is not self-evident. Some people believe that they are worse off for being slandered even if they never learn of the slander and its repercussions never affect their experience. Some even believe, following Aristotle's suggestion, that the quality of one's life as a whole can be affected by posthumous states of affairs such as tragedy befalling a loved one. Although the intelligibility of this belief in posthumous interests might be challenged, the following is surely intelligible: States of affairs that don't affect one's experience but

connect importantly with one's values can affect one's interests at least while one exists. Desire-based accounts of well-being (see, e.g., Hare 1981) standardly accept this principle, for what is desired may occur without one's awareness of its occurrence and without affecting one's experience. These considerations illuminate the intelligibility of one's prudential values extending to a period of time when one is alive but irreversibly unconscious.

In view of apparently reasonable pluralism regarding prudential values, including reasonable disagreement about the experience requirement, it seems doubtful that appeal to prudential value alone can support the higher-brain standard for everyone. At the same time, and more generally, the higher-brain approach remains an important contender in the debate over the definition of death.

11.4 A PROPOSED RETURN TO TRADITION: AN UPDATED CARDIOPULMONARY APPROACH

Prior to the brain-death movement, death was traditionally understood along the lines of the cardiopulmonary standard: death as the irreversible cessation of cardiopulmonary function. In the supportive background of this consensus on the cardiopulmonary standard hovered several general definitions or conceptualizations of death. Some champions of the traditional standard (e.g., Becker 1975) have conceptualized death in the same organismic terms that proponents of the whole-brain standard invoke: death as the irreversible cessation of functioning of the organism as a whole. Other champions of tradition have conceptualized death in more spiritual terms such as the departure of the animating (or vital) principle or loss of the soul.

In determining whether someone was dead, one could check for a pulse, moisture on a mirror held in front of the mouth, or other indications that the heart and lungs were working. Before the development of respirators and other modern life-supports, a working heart and lungs indicated

continuing brainstem function. As we have seen, however, modern life-supports permitted cardiopulmonary function without brain function, setting up a competition between traditional and whole-brain criteria for determining death. Although, as noted above, the whole-brain approach achieved near-consensus status, this approach is increasingly questioned and faces significant difficulties. Its difficulties and those facing the more radical higher-brain alternative have contributed to renewed interest in the traditional approach.

Further contributing to renewed interest in the traditional approach—and warranting a brief digression—is an approach to organ donation that capitalizes on the fact that current American legal standards for death are disjunctive, permitting satisfaction of either the whole-brain standard or the cardiopulmonary standard, whichever applies first, for a declaration of death. This approach to organ donation, called donation after cardiac death (DCD) or non-heart-beating organ donation, was very rare until instituted with much publicity by the University of Pittsburgh in the early 1990s in response to a perception that awaiting a neurological determination of death for (heart-beating, respirator-maintained) organ donors was insufficient to meet the demand for viable organs. In the Pittsburgh program, a respirator-dependent patient who had previously agreed to forgo life supports and donate vital organs is taken to an operating room and disconnected from the respirator, leading predictably to cardiac arrest. Two minutes after cardiac arrest, the patient is declared dead on the basis of the cardiopulmonary standard: "irreversible cessation of circulatory and respiratory functions." This procedure allows organ procurement to commence very shortly after cardiac arrest, providing relatively fresh organs for transplant. (Organs, of course, would not be viable if medical staff awaited a declaration of total brain failure—which requires confirmatory tests hours after initial tests—in the patients in question, who will not incur total brain failure unless respirator support is discontinued.)

The practice of DCD, which has expanded to several medical centers, has provoked considerable controversy. Critics have charged that in DCD

vital organs are removed before patients are really dead, implying that organ procurement kills the patients. Some proponents of the whole-brain approach argue that the patients are not yet dead because only total brain failure (or perhaps that of the brainstem) constitutes human death. But current law in its disjunctive form suggests otherwise—at least for legal purposes. Other critics of DCD charge that a patient cannot be dead two minutes after cardiac arrest because the loss of cardiopulmonary functioning is not irreversible: Victims of heart attack are sometimes revived more than two minutes after the arrest. One might reply that the loss of functioning is irreversible because, the patient having requested removal of life supports, no one may violate the patient's rights by resuscitating him or her (Tomlinson 1993). It seems fair to reply, however, that a decision not to resuscitate does not mean that resuscitation is impossible as suggested by the concept of irreversibility. Has the latter concept been conflated in DCD with the concept of permanence? Permanent loss of function does not imply that resuscitation is impossible, just that it will not occur. These concerns about abandoning the standard of irreversible loss of cardiopulmonary function apply even to more modest proposals, such as that advanced by the Institute of Medicine (2000), in which a declaration of death and DCD proceed after a waiting period of five minutes: Resuscitation is sometimes possible more than five minutes after a heart attack. Proponents of DCD might reply that permanence, rather than irreversibility, is the appropriate standard in this context (see, e.g., Bernat 2006, 41) or that DCD represents an instance where it is permissible to remove vital organs from someone who is dying but not yet dead. Certainly, any proponent of DCD will see the current law's (disjunctive) acceptance of cardiopulmonary criteria for death as offering a major practical advantage over any policy that accepted only whole-brain criteria.

We return to the view of those who champion only the cardiopulmonary standard. Proponents of this approach believe that it correctly implies, contrary to competing standards, that a human body that is breathing and maintaining circulation is alive regardless of whether continuation of these functions requires external support (as with "brain-dead" patients,

locked-in patients, and normal fetuses) (Shewmon 2001; Potts 2001). At the same time, the usual characterization of the traditional approach is problematic in suggesting that the difference between human life and death comes down to the state of two organs: heart and lungs. This reductionistic picture arguably obscures the holistic nature of bodily functioning.

A more realistic picture, some argue, features integrative unity as existing diffusely throughout the organism. As a leading proponent puts it, "What is of the essence of integrative unity is neither localized nor replaceable: namely the anti-entropic mutual interaction of all the cells and tissues of the body, mediated in mammals by circulating oxygenated blood" (Shewmon 2001, 473). On this view, the brain, like the heart and lungs, is a very important component of the interaction among body systems, but is not the supremely important integrator as suggested by the (mainstream) whole-brain approach. Nor is the functioning of other organs and bodily systems passively dependent on the brain. The brain's capacity to augment other systems presupposes their preexisting capacity to function. This is true even of a brain function as somatically integrating as the maintenance of body temperature: the "thermostat" may be in the brain, but the "furnace" is the energy metabolism diffused throughout the body. If not covered with blankets, brain-dead bodies maintained on respirators will grow colder—but not comparably to corpses (ibid, 471).

Although a realistic picture of organismic functioning must be holistic, according to this updated traditional approach, it should also portray certain functions as central. Tradition is correct that respiration and circulation are especially crucial, but respiration is not simply lung function and circulation is not just a working heart. Both organs, after all, can be artificially replaced as the organism maintains integrated functioning. Respiration and circulation occur throughout the body as oxygenated blood circulates to different organs and bodily systems—a condition necessary and sufficient for the integrated organismic functioning that constitutes life. Unlike whole-brain and higher-brain death, loss of respiration and circulation leads relentlessly to the

breakdown of cells, tissues, organs, bodily systems, and eventually the organism as a whole. Hence an updated traditional standard, which we might call the circulatory-respiratory standard: death as the irreversible cessation of circulatory-respiratory function.

The chief advantage of such an updated traditional approach, according to proponents, is that it most adequately characterizes the difference between life and death—where the latter is understood in terms of organismic functioning—in a full range of cases. Such cases include several that the whole-brain and higher-brain standards handle less plausibly such as prenatal human organisms prior to brain development as well as locked-in patients and "brain-dead" individuals whose vital functions are maintained with mechanical assistance. The present approach also avoids some of the conceptual problems facing the higher-brain approach, as discussed earlier.

Nevertheless, the traditional approach, whether updated or not, faces significant issues. One concern is that the approach overemphasizes our biological nature, suggesting we are nothing more than organisms, and by demoting the brain from prominence underemphasizes the mental life that is generally thought to distinguish our species from others. We human beings are not merely organisms or animals, the argument continues; we are also (after normal development) conscious beings and persons whose nature, one might say, is to transcend nature with culture. Our conception of human death should be faithful to a species self-image that does justice not only to our animality but also to our personhood (cf. Pallis 1999, 96).

Whole-brain (or brainstem) theorists and higher-brain theorists will extend this line of argument in different directions. The higher-brain theorist will suggest that our capacity for consciousness, a precondition for higher capacities and personhood, is so important that permanent loss of the basic capacity should count as death. The whole-brain theorist who develops the present line of reasoning will maintain greater contact with the organismic conception of death, stressing the brainstem's role in integrating vital functions and claiming either that (a) consciousness is a

critical function of the organism, permitting it to interact adaptively to its environment (Bernat 1998), (b) consciousness is a characteristic aspect of the fundamental work of organisms like us, or (c) consciousness is crucial to our personhood, a feature no less important to what we are than our animality. The latter option, in effect, would move the whole-brain theorist to a dual-aspect understanding of human nature, as just discussed: human persons as essentially both persons and animals (cf. Schechtman 2014).

A second major challenge confronting any traditional approach is the specter of highly unpalatable practical consequences. Currently the whole-brain standard is enshrined in law. Suppose we reversed legislative course and returned to traditional criteria (whether updated in formulation or not). Then a patient who satisfied whole-brain criteria would count as alive. Unless we overturned the "dead-donor rule"—the policy of permitting extraction of vital organs only from dead bodies—then it would be illegal to procure organs from these living patients who have incurred total brain failure; yet the viability of their organs would require maintaining respiration and circulation with life-supports. There is broad agreement that having to wait until traditional criteria are met to harvest organs would constitute a great setback to organ transplantation (even if donation after cardiac death, which invokes traditional criteria, is permitted). Moreover, a legal return to traditional criteria for death might lead physicians to feel they had lost the authority to discontinue treatment unilaterally—when a family requests continued treatment—upon a determination of total brain failure despite what many would consider the futility of further treatment. Furthermore, laws for determining death would have to be revised.

A defender of tradition might respond that we can avoid most of these unsavory consequences while legally adopting traditional criteria for determining death (see, e.g., DeGrazia 2005, 152–8). We could, for one thing, abandon the dead-donor rule, permitting the harvesting of vital organs when authorized by appropriate prospective consent of the donor even though taking the organs, by causing the donor's death, would

instantiate killing (Truog and Robinson 2003; Sade 2011). We could also authorize physicians—through hospital policies, professional guidelines, or laws—unilaterally to withdraw life-supports upon a declaration of total brain failure (perhaps even upon a determination of irreversible unconsciousness) in cases where continued treatment is unnecessary for organ procurement and appears otherwise futile. Not all of what are traditionally considered "death behaviors" need to be permanently anchored to a declaration of death. Thus we currently use advance directives and other considerations to justify withdrawal of life-supports in some circumstances, although several decades ago such withdrawal had to await a determination of death. There is no reason to regard further reforms of our practices surrounding death as beyond responsible consideration. Thus, despite rowing against the tide of the brain-death movement, the traditional approach has reclaimed the status of a serious contender in the debate over the definition of death.

11.5 FURTHER POSSIBILITIES

In recent decades, the debate over the definition of death has generally been understood as a competition between the approaches discussed here: traditional, whole-brain (or brainstem), and higher-brain standards and their corresponding conceptualizations. Each of these approaches, however, makes certain assumptions that might be contested: (1) that death is more or less determinate, more event-like than process-like, (2) that there is a uniquely correct definition of death, which can be formulated in terms of necessary and jointly sufficient conditions, and (3) that human death is morally a very important marker. Now we will consider three nonstandard ways of thinking about death, each of which directly challenges one of these assumptions.

11.5.1 Death as a Process, Not a Determinate Event

Each of the approaches considered so far asserts the correctness of a single standard of death. Might different standards be appropriate for different purposes? If so, then the debate characterized in previous sections has reflected, to some extent, an exercise in futility: a search that

wrongly seeks a determinate event, which can be captured by a single standard, rather than a process.

According to two authors who develop this line of reasoning, the nearly simultaneous emergence of organ transplantation and mechanical ventilators provoked three practical questions: (1) When may doctors take organs for transplantation? (2) When may doctors unilaterally discontinue treatment? (3) When is a patient dead for legal purposes and appropriately transferred to an undertaker? (Halevy and Brody 1993). Rather than assuming that one standard for death will adequately answer these three questions—a possibility rendered doubtful by the interminable debate over standards—we should answer each question on its merits, disaggregating death accordingly.

Providing one example of how these practical questions might be answered, the authors argue that organ procurement is appropriate when the whole-brain standard has been met (apparently precluding DCD), unilateral discontinuation of treatment is appropriate when the higher-brain standard has been met, and a patient should legally count as dead when traditional criteria have been met (ibid). (Here we need not consider the authors' specific arguments for these determinations.)

But why must each answer invoke a standard of death? An alternative would be to adopt an updated traditional standard, which would supply legal criteria for death, while denying that unilateral discontinuation of treatment and organ procurement must await death. To be sure, harvesting vital organs from living patients would require an exception to the dead-donor rule, the social risks of which might well be avoided if death were disaggregated along the lines suggested. But the alternative possibility of separating death from particular "death behaviors" motivates the question of whether there are further grounds for disaggregating death into a process.

A possible further ground is the thesis that life and death, although mutually exclusive states, are not exhaustive: "Although no organism can

fully belong to both sets [life and death], organisms can be in many conditions (the very conditions that have created the debates about death) during which they do not fully belong to either. … Death is a fuzzy set," (Brody 1999, 72). What are we to think of this proposal?

It seems undeniable that the boundary between life and death is not perfectly sharp. The specification of any standard will require some arbitrary line-drawing. Operationalizing the whole-brain standard requires a decision about which brain functions are too trivial to count and need not be tested for. Making a traditional standard clinically useful requires a cut-off point of some number of minutes without heartbeat or respiration for the loss of functioning to count as irreversible. A higher-brain approach needs criteria for determining what sorts of brain damage constitute irreversible loss of the capacity for consciousness and which count as reversible. Yet, while some arbitrariness is inevitable, and highlights a blurred boundary, the blurring in each instance concerns very specific criteria and clinical tests for determining that a standard has been met, not the standard itself. None of the blurred boundaries just considered is inconsistent with the claim that some standard is uniquely correct. Moreover, if essentialism regarding human persons is true—that is, if we human persons have an essence locating us in our most basic kind (e.g., animal, minded being)—this would strengthen the case for a uniquely correct standard by suggesting a foundation for one.

But we must consider the possibility that there is no correct standard. Perhaps death is no more determinate than adulthood. Some people are clearly adults and some people are clearly not adults. But, as any college professor knows, many people are ambiguously adults—mature enough to count as adults in some ways but not in others. Socially and legally, we treat 16-year-olds as adults for purposes of driving, 18-year-olds as adults for purposes of voting and bearing the full weight of criminal law, 21-year-olds as adult enough to drink alcoholic beverages, and so on. Nor is this disaggregation of adulthood incoherent or even particularly awkward; rather, it seems to fit the facts about the gradual development of maturity, acquisition of experience, and accumulation of birthdays. Disaggregating

death, one might argue, would be similarly faithful to facts about the frequently very gradual demise of human persons.

Even if this argument persuades us that death is more process-like than event-like—and to do this it must persuade us that it is death itself, not dying, that is process-like—it does not follow that we ought to draw several lines for the determination of death. Consider the confusion that would likely result from such statements as "Grandmother is partly dead, but less dead than Grandfather, although he's not fully dead." People are so accustomed to thinking of life and death as mutually exclusive, exhaustive sets that there would be considerable practical advantage in insisting on some sensible line that demarcates death in this way. It is true that disaggregating adulthood poses no insuperable practical difficulties, but death is importantly different. For we generally assume that one goes out of existence (at least in this world) at death, a rather momentous change with—at least in the status quo—far-reaching social and legal ramifications. Confusion as a result of plural lines for death may be more troubling and more likely, for the idea of someone's only partly existing is of questionable intelligibility. On the other hand, a proponent of disaggregating death might reply that (1) we could either reserve the language of death for the traditional standard or get used to the language of someone's being partially dead, and (2) we should appreciate that existence is sometimes partial as in the case of a half-assembled car.

11.5.2 Death as a Cluster Concept not Amenable to Classical Definition

Most discussions of the definition and determination of death assume that there is a uniquely correct definition of death. Definitions, classically understood, are supposed to state necessary and jointly sufficient conditions for the correct application of a word or concept. They may be thought to capture de re essences existing independently of human thought, language, and interests, or de dicto essences determined solely by linguistic meaning. The major approaches we have considered have tried both to define death by capturing its essence and to advance a standard for

determining human death that coheres with the definition. But what if the term "death" cannot be defined in any such way?

One might insist that death can be defined, as the competing definitions demonstrate. But, of course, the trick is to define the term adequately. For example, the organismic definition—death as the irreversible cessation of functioning of the organism as a whole—makes no reference to consciousness. Yet surely, one might argue, any organism that maintains consciousness should count as alive even if the organism as a whole has irreversibly ceased to function (whether or not this possibility is merely theoretical). Definitions associated with the higher-brain approach—such as human death as the irreversible loss of mind—implausibly imply that a PVS patient is dead despite exhibiting spontaneous breathing and circulation, brainstem-mediated reflexes, and the like. The best explanation for the shortcomings of leading efforts to define death, the argument continues, is that death is not amenable to definition in terms of necessary and sufficient conditions (Chiong 2005). Let's consider two distinct ways this thesis might be developed.

First, one might argue that the concept of death exhibits only "family resemblance" relations among its instances, as Wittgenstein argued was the case for the concepts of game, language, and many others (Wittgenstein 1953). There are various features of an organism that count towards its being dead, yet there is no authoritative list of features all of which must be satisfied for it to be dead. Each of the following, for example, seems relevant: unconsciousness, absence of spontaneous efforts to breathe, absence of heartbeat, inertness, lack of integrated bodily functions, incapacity to grow, and physical decay. If all of these conditions are present, an organism has surely died. But producing an authoritative shortlist of necessary and sufficient conditions seems futile. One scholar has advanced a parallel claim about the concept of life:

When some property is central to the cluster—as I've argued consciousness is—then possessing only this one property may be sufficient for membership in [the class of living things]. However, merely

possessing one or several properties that are peripheral to the cluster may not be sufficient for membership. [S]ome robots are organizationally complex and functionally responsive, though intuitively not alive (Chiong 2005, 26).

Another direction in which to take the thesis that death is not amenable to classical definition is to argue that death is a natural kind whose essence may be obscure. Kripke influentially argued that natural kinds—kinds determined by nature rather than by human thinking, language, or interests—often resist adequate definition because their essential features may be entirely unknown to those referring to the kind in question (Kripke 1970). To define a term by reference to the features people originally used to pick out the kind in question won't do, because those features may be accidental, not essential, and speakers may even be mistaken about them. Those naming the kind whale might have thought whales were the largest fish in the ocean, but whales aren't fish and their size relative to other creatures is a contingent matter. We can refer meaningfully to whales, to the creatures picked out by the term whale (the name for the kind), without knowing the essential features of whales, features likely to involve subtle biological details. Perhaps death, too, is a natural kind whose essence is obscure (a possibility entertained in Chiong 2005, 24–25). A likely challenge to this argument is that we already know a great deal about the physical processes involved in death, making it unlikely that death has a hidden essence the failure to discover which impedes adequate definition.

Importantly, though, one can claim that death is a natural kind without accepting any kind of essentialism. An alternative to the essentialist conception is the homeostatic property cluster theory of natural kinds (Millikan 1999). On this view, natural kinds do not, or at least need not, share essential properties. They are comprised by members sharing a stable cluster of similarities, which are brought about by "homeostatic causal mechanisms" (such as, in the case of species, common developmental programs and selective pressures). On this view, X (e.g., a fetus) might be a member of a natural kind (e.g., our species) despite lacking one of the properties (e.g., potentiality for rationality) among the

cluster of similarities. Death and its opposite, life, might similarly be natural kinds lacking essences, each kind being associated with a cluster of properties that tend to go together and support one another without being necessarily coinstantiated (see, e.g., Chiong 2005). If so, death cannot be defined in a set of necessary and sufficient conditions—in which case no such definition can justify a particular standard.

If death has no essence and resists definition, what is the upshot? One possible inference—that the boundaries of death are vague—would partially merge this approach with the previous one, which construed death as a process. We have noted that one response to the claim of vague boundaries (the response favored in the previous approach) is to embrace several lines, each for a different purpose, in determining death. Another possibility is to understand the vague boundaries as inviting discretion in the matter of producing a single standard of death. So long as a particular standard does not have clear and highly implausible implications, it is admissible for consideration on this view. Society may then select, among admissible standards, whichever is most attractive for practical purposes. It has been argued, along these lines, that the higher-brain standard is inadmissible for implying that those in PVS are dead while the traditional cardiopulmonary standard is inadmissible for implying (in principle) that a still-conscious individual might be dead, clearing the ground for the whole-brain standard, which has no fatal implications and is attractive from a practical standpoint (Chiong 2005).

Having already explored difficulties (and strengths) of each standard, how might we evaluate the more general thesis that death is not amenable to classical definition? One strategy open to critics of this reasoning, of course, is to argue that some definition is adequate. Another is to defend the disaggregation of death, as previously discussed. A third strategy would be to argue that our failure thus far to produce an adequate definition does not mean that none is possible. Some concepts can be adequately captured by classical definitions even if it is difficult to produce them. It would appear premature, therefore, to render a judgment on the success of the present approach to understanding human death.

11.5.3 Death as Separable from Moral Concerns

A final assumption underlying the mainstream discussion of the definition of death is that human death is a morally crucial marker. Were it not, then accuracy in the definition of death would be of purely ontological, conceptual, or scientific interest. This attitude, of course, is not the prevailing one. Not only do we tend to regard many behaviors as appropriate only if an individual has died; the criminal law treats as momentous the question of whether one person has killed—that is, caused the death of—another person, even if such considerations as motive, deliberation, and special circumstances are also important.

It is not difficult to see, though, how one might challenge this presumption of death's moral salience. After all, we have already begun to remove certain behaviors from the class of death behaviors. For example, in many circumstances termination of life supports need not await a patient's death. And, as we have noted, there are calls to abandon the dead-donor rule in the context of organ transplantation. We might go further in separating death from the cluster of moral concerns traditionally associated with it. For example, without embracing the higher-brain approach to death, we could hold that irreversible loss of the capacity of consciousness entails a loss of moral status, at which point traditional death behaviors are appropriate (Persson 2002). We might even overhaul the criminal law with respect to killing:

It is then the irrevocable loss of the capacity for consciousness that is the great loss; so it is for the causing of it that criminal law should mete out the severest punishment. Killing, or the causing of (biological) death, should be punished to this degree only if, as is normally the case, it brings along the irrevocable loss of the capacity for consciousness (ibid, 32).

One implication of this proposal is that harvesting organs from PVS patients, thereby killing them, would not be punishable insofar as these patients, having irrevocably lost the capacity for consciousness, have already suffered "the great loss" and no longer possess moral status. Some attracted to this approach will want to argue further that the crime of

murder is really that of causing the irrevocable loss of the capacity for consciousness without first obtaining voluntary, informed consent from the person to be killed. The italicized qualification would create conceptual space for a justification of active euthanasia (see the entry on voluntary euthanasia).

Check Your Progress 1

Note: a) Use the space provided for your answer

b) Check your answers with those provided at the end of the unit

1.   What do you know about the Current Mainstream View: The Whole-Brain Approach?

    _____

    _____

    _____

2.   To discuss A Progressive Alternative: The Higher-Brain Approach.

    _____

    _____

    _____

11.6 LET US SUM UP

The present proposal to separate the issue of death from what is morally important is somewhat radical. Yet its chief ground for doing so, the claim that the capacity for consciousness is what underlies moral status, cannot be dismissed. On the other hand, this claim apparently relies on the thesis (which we considered in connection with the higher-brain approach) that only what affects one's experience can affect one's interests. As we saw, this thesis is far from self-evident. For those who disagree with it, the time of death—the time at which one no longer exists (at least in this world)—is likely to retain some of the moral importance traditionally accorded to it. Moreover, even if the philosophical case for demoting the moral importance of death were airtight, we cannot responsibly dismiss widely held sensibilities, including those at odds with the present approach, in constructing public policies concerning death.

Certainly it is contestable to what extent the public could embrace further demotion of the moral importance of death, and to what extent its limited ability to do so matters for public policy.

11.7 KEY WORDS

Cardiopulmonary: relating to the heart and the lungs.

11.8 QUESTIONS FOR REVIEW

1. Highlight a Proposed Return To Tradition: An Updated Cardiopulmonary Approach
2. What do you know the Further Possibilities?

11.9 SUGGESTED READINGS AND REFERENCES

- Ad Hoc Committee of the Harvard Medical School, 1968, "A Definition of Irreversible Coma—Report of the Ad Hoc Committee of the Harvard Medical School to Examine the Definition of Brain Death," Journal of the American Medical Association 205 (6): 337–40.
- Baker, L. R., 2000, Persons and Bodies, Cambridge: Cambridge University Press.
- Bartlett, E. and S. Youngner 1988, "Human Death and the Destruction of the Neocortex," in R. Zaner (ed.), 1988, Death: Beyond Whole-Brain Criteria, Dordrecht, Netherlands: Kluwer.
- Becker, L., 1975, "Human Being: The Boundaries of the Concept," Philosophy and Public Affairs, 4: 334–59.
- Bernat, J., 2006, "The Whole-Brain Concept of Death Remains Optimum Public Policy," Journal of Law, Medicine & Ethics, 34 (1): 35–43.
- ——, 1998, "A Defense of the Whole-Brain Concept of Death," Hastings Center Report, 28 (2): 14–23.
- ——, 1992, "How Much of the Brain Must Die on Brain Death," Journal of Clinical Ethics, 3 (1): 21–26.

- Bernat, J., C. Culver, and B. Gert, 1981, "On the Definition and Criterion of Death," Annals of Internal Medicine, 94: 389–94.

- Brody, B. 1999, "How Much of the Brain Must be Dead?" in S. Youngner, R. Arnold, and R. Shapiro (eds.), The Definition of Death: Contemporary Controversies, Baltimore, MD: Johns Hopkins University Press: 71–82.

11.10   ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

1. See Section 11.2
2. See Section 11.3

UNIT 12: DEATH

STRUCTURE

12.0 OBJECTIVES

After this unit, we can able to know:

1. To know the concept of life and death
2. To discuss about the misfortune.
3. To highlight The Harm Theses
4. Is Death Always a Misfortune?
5. Can Death's Harmfulness be Reduced?

12.1 INTRODUCTION

First, what constitutes death? It is clear enough that people die when their lives end, but less clear what constitutes the ending of a person's life.

Second, in what sense might death or posthumous events harm us? To answer this question, we will need to know what it is for something to be in our interests.

Third, what is the case for and the case against the harm thesis, the claim that death can harm the individual who dies, and the posthumous harm thesis, according to which events that occur after an individual dies can still harm that individual?

Fourth, how might we solve the timing puzzle? This puzzle is the problem of locating the time during which we incur harm for which death and posthumous events are responsible.

A fifth controversy concerns whether all deaths are misfortunes or only some. Of particular interest here is a dispute between Thomas Nagel, who says that death is always an evil, since continued life always makes good things accessible, and Bernard Williams, who argues that, while premature death is a misfortune, it is a good thing that we are not immortal, since we cannot continue to be who we are now and remain meaningfully attached to life forever.

A final controversy concerns whether or not the harmfulness of death can be reduced. It may be that, by adjusting our conception of our well-being, and by altering our attitudes, we can reduce or eliminate the threat death

poses us. But there is a case to be made that such efforts backfire if taken to extremes.

12.2 DEATH

Death is life's ending. To clarify death further, we will need to say a bit about the nature of life, and ask whether life can be suspended or restored, and how it relates to our continued existence. We can also distinguish between the concept of death and criteria by which death can be detected.

12.2.1 Life and Death

It is not easy to clarify the nature of life. Suppose we could construct a machine, the HAL 1.01, with (nearly) all of the psychological attributes of persons: would HAL 1.01 be alive? Probably not, given the nature of HAL's hardware. It seems that being conscious does not entail being alive. Still, to the extent that we are puzzled about the nature of life, we will be puzzled about what is entailed by the ending of life, that is, death.

Things that are alive have a distinctive capacity to develop or maintain themselves by engaging in various processes including chemosynthesis, photosynthesis, cellular respiration, cell generation, and maintenance of homeostasis. Let us call these vital processes. It is one thing to have the capacity to deploy these processes and another to actually deploy them, just as there is a difference between having the ability to run and actually running. For something to have the property 'alive' seems to be a matter of its having the capacity to sustain itself using processes that are saliently similar to these. (For accounts of life, see Van Inwagen 1990 and Bedau 2014.)

By contrast, the property 'dead' seems applicable to something that has lost this capacity. We can call this the loss of life account of death. The event by which the capacity to employ vital processes is lost is one thing and the condition of having lost it is another. 'Death' can refer to either.

Let us add that 'the ending of life' is itself potentially ambiguous. On one hand it might be a process wherein our lives are progressively extinguished, until finally they are gone. On the other it might be a momentary event. This event might be understood in three ways. First, it might be the ending of the dying process—the loss of the very last trace

of life. Call this 'denouement death'. Second, it might be the point in the dying process when extinction is assured, at least given the resources available to prevent it. Call this moment 'threshold death'. A third possibility is that life ends when the physiological systems of the body have lost the capacity to function as an integrated whole, or when this loss becomes irreversible (Belshaw 2009; DeGrazia 2014). Call this 'integration death'.

Thus death can be a state (being dead), the process of extinction (dying), or one of three events that occur during the dying process. Death in all of these senses can be further distinguished from events—such as being shot with an arrow—that cause death.

12.2.2 Death and Suspended Animation

The loss of life account of death has been challenged by theorists who claim that things placed in suspended animation are not alive (Feldman 1992, Christopher Belsaw 2009, Cody Gilmore 2013, and David DeGrazia 2014). When zygotes and embryos are frozen for later use in the in vitro fertilization procedure, their vital processes are brought to a stop, or very nearly so. The same goes for water bears that are dehydrated, and for seeds and spores. It seems clear that the zygotes and water bears are not dead, since their vital processes can easily be restarted—by warming the zygote or by wetting the water bear. They are not dead, but are they alive? If we deny that they are alive, presumably we would do so on the grounds that their vital processes are halted. If something's life can be ended by suspending its vital processes without its dying, then we must reject the loss of life account of death.

However, the loss of life account is thoroughly established in ordinary usage, and is easily reconciled with the possibility of suspended animation. In denying that frozen embryos are dead, it is clear that we mean to emphasize that they have not lost the capacity to deploy their vital processes. When we say that something is dead, we mean to emphasize that this capacity has been lost. Having used 'dead' to signal this loss, why would we want to use the word 'alive' to signal the fact that something is making active use of its vital processes? Our best option is to use a pair of contrasting terms. We can use 'viable' to indicate that

something has the capacity to deploy vital processes and 'unviable' to indicate that it has lost this capacity. When instead we are concerned about whether or not something is engaging its vital processes, we can use different contrasting terms, say 'vital' and 'nonvital', the former to characterize something that is employing its capacity for vital processes and the latter to characterize something that is not making use of its capacity for vital processes. What seems relatively uncontroversial is that being dead consists in unviability. To retain the loss of life account, we have only to add that being alive consists in viability. We can then say that a frozen embryo is viable and hence alive despite its lack of vitality, and it will die if its life ends (it will die if it ceases to be viable). Of course, if we are willing to abandon the loss of life account, we could instead use 'alive' to characterize something that is both viable and vital. We would then say that a frozen embryo is not alive (since it lacks vitality) but also that it is not dead (since it remains viable).

12.2.3 Resurrection

It will be useful to sharpen the loss of life account if, as seems conceivable, it is possible to restore life to something that has died.

Restoration in this sense is quite different from the revival of something, such as a frozen embryo, whose vital processes have been halted. Something can be revived only if it is alive—only if it has the capacity to deploy vital processes, as in the case of a frozen zygote. It is revived when it regains vitality. Something's life can be restored only if it has lost its capacity for vital processes. Life is restored when this capacity is regained.

To bring the possibility of restoration into view, imagine a futuristic device, the Disassembler-Reassembler, that reduces me to small cubes, or individual cells, or disconnected atoms, which it stores and later reassembles just as they were before. Many of us will say that I would survive—my life would continue—after Reassembly, but it is quite clear that I would not live during intervals when my atoms are stacked in storage. I would not even exist during such intervals. If I can be Reassembled, my life would be restored, not revived. Restoration, not revival, is a way of bringing a creature back from the dead. Now imagine

a device that repairs corpses: it moves molecules back to where they were prior to the death of the creature that left the corpse, and restarts its vital processes. Like the Disassembler-Reassembler, the corpse reanimator would resurrect the dead—it would restore the lives of people who have died.

Given the possibilities of restoration and revivification, it seems best to refine the loss of life account, as follows:

Dying is the loss of a thing's life—the loss of its capacity to perpetuate itself using vital processes. A thing dies at the time it loses this capacity. It is dead at all times afterwards, except while that capacity is regained.

12.2.4 Death and What We Are

Death for you and me is constituted by the loss of our capacity to sustain ourselves using vital processes. This characterization of death could be sharpened if we had a clearer idea of what we are, and the conditions under which we persist. However, the latter is a matter of controversy.

There are three main views: animalism, which says that we are human beings (Snowdon 1990, Olson 1997, 2007); personism, which says that we are creatures with the capacity for self-awareness; and mindism, which says that we are minds (which may or may not have the capacity for self-awareness) (McMahan 2002). Animalism suggests that we persist over time just in case we remain the same animal; mindism suggests that we persist just when we remain the same mind. Personism is usually paired with the view that our persistence is determined by our psychological features and the relations among them (Locke 1689, Parfit 1984).

If we are animals, with the persistence conditions of animals, our deaths are constituted by the cessation of the vital processes that sustain our existence as human beings. If we are minds, our deaths are constituted by the extinction of the vital processes that sustain our existence as minds. And if persistence is determined by our retaining certain psychological features, then the loss of those features will constitute death.

These three ways of understanding death have very different implications. Severe dementia can destroy a great many psychological features without destroying the mind, which suggests that death as understood by personists can occur even though death as understood by mindists has not.

Moreover, human beings sometimes survive the destruction of the mind, as when the cerebrum dies but the brainstem does not, leaving an individual in a persistent vegetative state. It is also conceivable that the mind can survive the extinction of the human being: this might occur if the brain is removed from the body, kept alive artificially, and the remainder of the body is destroyed (assuming that a bare brain is not a human being). These possibilities suggest that death as understood by mindists can occur even though death as understood by animalists has not (and also that the latter sort of death need not be accompanied by the former.)

12.2.5 Death and Existence

What is the relationship between existence and death? May people and other creatures continue to exist after dying, or cease to exist without dying?

Take the first question: may you and I and other creatures continue to exist for some time after our lives end? Fred Feldman (1992, p. 91) coins the term termination thesis to refer to the view that "when a person dies, he or she. . .goes out of existence; subsequently, there is no such thing as that person." (A version of the thesis applies to any living thing.) The position that we can indeed survive death we might call the dead survivors view. The dead survivors view has been defended by various theorists, most notably Feldman (1992, 2000, 2013). One point cited in its favor is that we commonly refer to 'dead animals' (and 'dead plants') which may suggest that we believe that animals continue to exist, as animals, while no longer alive. The idea might be that an animal continues to count as the same animal if enough of its original components remain in much the same order, and animals continue to meet this condition for a time following death (Mackie 1997). On this view, if you and I are animals (as animalists say) then we could survive for a time after we are dead, albeit as corpses. In fact, we could survive indefinitely, by arranging to have our corpses preserved.

However, this way of defending the dead survivors view may not be decisive. The terms 'dead animal' and 'dead person' seem ambiguous. Normally, when we use 'dead people' or 'dead animal' we mean to speak

of persons or animals who lived in the past. One dead person I can name is Socrates; he is now a 'dead person' even though his corpse surely has ceased to exist. However, in certain contexts, such as in morgues, we seem to use the terms 'dead animal' and 'dead person' to mean "remains of something that was an animal" or "remains of something that was a person." On this interpretation, even in morgues calling something a dead person does not imply that it is a person.

What about the second question: can creatures cease to exist without dying? Certainly things that never were alive, such as bubbles and statues, can be deathlessly annihilated. Arguably, there are also ways that living creatures can be deathlessly annihilated (Rosenberg 1983, Feldman 1992, Gilmore 2013). Perhaps an amoeba's existence ends when it splits, replacing itself with two amoebas, and the existence of chlamydomonas ends when pairs of them fuse to form a zygote. Yet when amoebas split, and chlamydomonas fuse, the vital processes that sustain them do not cease. If people could divide like amoebas, perhaps they, too could cease to exist without dying. (For a famous discussion of division, fusion, and their implications, see Parfit 1981.) If such 'deathless exits' are possible, we would have to modify the loss of life account of death.

However, proponents of that account can hold their ground. They can say that division, fusion, and other apparent examples of deathless exits are unusual ways of dying, because nonexistence is not brought about via the destruction of vital processes, but they are not ways of escaping death altogether. Proponents of the loss of life account might also turn the tables on its critics, and argue as follows: nothing can be alive unless it exists, so if something ceases to exist it ceases to be alive, but to cease to be alive is to die. So there are no deathless exits after all.

12.2.6 Criteria for Death

Defining death is one thing; providing criteria by which it can be readily detected or verified is another. A definition is an account of what death is; when, and only when its definition is met, death has necessarily occurred. A criterion for death, by contrast, lays out conditions by which all and only actual deaths may be readily identified. Such a criterion falls short of

a definition, but plays a practical role. For example, it would help physicians and jurists determine when death has occurred.

In the United States, the states have adopted criteria for death modeled on the Uniform Determination of Death Act (developed by the President's Commission, 1981), which says that "an individual who has sustained either (1) irreversible cessation of circulatory and respiratory functions, or (2) irreversible cessation of all functions of the entire brain, including the brain stem, is dead. A determination of death must be made in accordance with accepted medical standards." In the United Kingdom, the accepted criterion is brain stem death, or the "permanent functional death of the brain stem" (Pallis 1982).

These current criteria are subject to criticism, even if we put aside reservations concerning the qualifier 'irreversible'. Animalists might resist the criteria since the vital processes of human beings whose entire brains have ceased to function can be sustained artificially using cardiopulmonary assistance. Mindists and personists might also resist the criteria, on the grounds that minds and all psychological features can be destroyed in human beings whose brain stems are intact. For example, cerebral death can leave its victim with an intact brain stem, yet mindless and devoid of self-awareness.

12.3 MISFORTUNE

May death or posthumous events harm us? Might they benefit us? Perhaps; in order to decide, we will need an analysis of welfare, which tells us what well-being is and how well off we are. We will also need an account of personal interests, which tells us what it is for something to be in our interests or against our interests.

12.3.1 Comparativism

The most widely accepted account of our interests is *comparativism*. In order to clarify comparativism, it is best to distinguish different senses in which an event can have value.

Some events are intrinsically good (or bad) for a subject; such events are good (bad) for their own sakes, rather than in virtue of their contingent effects. By contrast, some events are extrinsically good (bad) for a subject; they are good (bad) because of their contingent effects. For example, many people count their pleasure as intrinsically good and their pain as intrinsically bad; aspirin would be extrinsically good, since it eliminates pain, and really bad puns would be extrinsically bad in that they are painful.

Events can have value in a different way: they can be overall good (bad) for a subject; that is, they can be good (bad) all things considered. Events are overall good (bad) for me when (and to the extent that) they make my life better (worse) than it would be if those events had not occurred. Contrast events that are partially good (bad) for me: these make my life better (worse) only in some respects. Partial goods may be overall bad for me. For example, playing video games every day gives me pleasure, and is hence partially good for me, but if it also causes me to neglect my job, health and family, it might well be overall bad for me.

According to comparativism, the value an event EE has for me is roughly EE's overall value for me. But let us attempt to formulate the comparativist account a bit more precisely.

To assess the value for me of an event EE, we begin by distinguishing two possible situations, or possible worlds. One of these is the actual world, which is the world as it actually is, past present and future. The other is the possible world that is the way things would be if EE had not occurred. We can assume that this is the world that is as similar to the actual world as possible, and in that way 'closest' to the actual world, except that EE does not occur, and various other things are different because of EE's nonoccurrence. We can call the actual world $WEWE$, in this way reminding ourselves that world $WEWE$, in this way reminding ourselves that EE actually occurred. And by $W{\sim}EW{\sim}E$ we can indicate the closest world to the actual world in which EE does not occur. Here the tilde, '$\sim\sim$', stands for 'not'.

The next step is to assess my welfare level in $WEWE$ and my welfare level in $W{\sim}EW{\sim}E$. My welfare level in $WEWE$ is the intrinsic value for me of my life in $WEWE$; it is the value my life actually has for me,

measured in terms of intrinsic goods and intrinsic evils. To calculate my welfare level in W_E, we start by assigning a value to my intrinsic goods in W_E. This will be a positive value representing the sum of these goods. Next we assign a value to my intrinsic evils in W_E; this will be a negative value. Next we sum these values; the goods will raise this sum, while the evils will lower it. Some symbolism might help fix ideas (although it may give the false impression that our subject matter admits of more precision that is actually possible). Let $G(S,W)$ stand for the sum of the values of S's intrinsic goods in world W, and let $B(S,W)$ stand for the sum of the values of S's intrinsic evils in W. So far we have said that S's welfare level in W_E equals $G(S,W_E)+B(S,W_E)$. If we let $IV(S,W)$ stand for the intrinsic value of world W for subject S, the claim is that

$$IV(S,W_E)=G(S,W_E)+B(S,W_E).$$

My welfare level in $W{\sim}E$ is assessed similarly; it is the sum of my intrinsic goods and evils in $W{\sim}E$.

Finally, we subtract the value for me of my life in $W{\sim}E$ from the value for me of my life in W_E. According to comparativism, this is the value E has for me. Letting $V(S,E)$ stand for the value of E for subject S, comparativism says that

$$V(S,E)=IV(S,W_E)-IV(S,W{\sim}E).$$

This value determines whether an event is overall bad (good) for a subject S. If E's value for S is negative, that is, if $V(S,E)<0$, then E is overall bad for S. If E's value is positive, then E is overall good for S. The more negative (positive) E's value is, the worse (better) E is for S.

Consider an example. Suppose that we are looking to identify the value for me of drinking this cup of coffee. Call this event *Drink*. Then the first step is to distinguish the actual world, W_{Drink}, in which I drank the coffee, from the closest world in which I did not, $W_{\sim Drink}$. Then we calculate my welfare level in W_{Drink}, $IV(Luper,W_{Drink})$ and in $W_{\sim Drink}$, $IV(Luper,W_{\sim Drink})$. The former, $IV(Luper,W_{Drink})$, equals the value of the

intrinsic goods I will enjoy in my life plus the value of the intrinsic evils I will endure. For simplicity, let us pull a number out of the hat to indicate the value of the goods I will enjoy in my life before I drink the coffee, say 100, and another number to indicate the value of the evils I will endure in my life before I drink the coffee, say −50−50. Let us also assume that drinking the coffee will give me some pleasure for one hour, which has a value of 10, and drinking the coffee will not cause me to endure any evils. Finally, let us assume that after that hour of savoring my coffee, I will go on to enjoy goods with a value of 50, and evils with a value of −10−10. Then

$$IV(Luper, WDrink) = 100 + 10 + 50 + -50 + 0 + -10 = 100.IV(Luper, WDrink) = 100 + 10 + 50 + -50 + 0 + -10 = 100.$$

Assuming that my life one hour *after* drinking my coffee would be just like my life would have been were I not to drink my coffee, more or less, so that drinking my coffee benefits me only during the hour I savor it, we can say that

$$IV(Luper, W{\sim}Drink) = 100 + 0 + 50 + -50 + 0 + -10 = 90.IV(Luper, W{\sim}Drink) = 100 + 0 + 50 + -50 + 0 + -10 = 90.$$

Given these assumptions,

$$V(Luper, WDrink) = IV(Luper, WDrink) - IV(Luper, W{\sim}Drink) = 100 - 90 = 10.V(Luper, WDrink) = IV(Luper, WDrink) - IV(Luper, W{\sim}Drink) = 100 - 90 = 10.$$

Drinking the coffee, then, was good for me, as 10 is a positive value.

We can now offer a rough statement of the comparativist account of interests.

An event EE is in SS's interests just in case EE overall benefits (is good for) SS, making SS's life better than it would have been if EE had not occurred, which EE does just when its value for SS is positive. An event EE is against SS's interests just in case EE overall harms (is bad for) SS, making SS's life worse than it would have been if EE had not occurred, which EE does just when its value for SS is negative. How much EE benefits (harms) SS depends on how much better (worse) SS's life is in the actual world than it would have been if EE had not occurred: the better (worse) SS's life is, the more beneficial (harmful) EE is.

In order to refine the comparativist account, we will need to distinguish between event *tokens* and event *types*. Event tokens are concrete events, such as the bombing of the World Trade Center. Event types are abstract entities such as bombings, leapings and burials. One token of the type *bombing* is the bombing of the World Trade Center. Earlier we used the letter 'EE' to refer to event tokens rather than event types. What is more, we assumed that the event tokens to which 'EE' referred were actual events, not merely possible events. But perhaps we can also offer a comparativist account of the value of the occurrence of an EE-type event; that is, a comparativist account of how valuable it would be for a subject SS if an event of type EE *were* to occur.

To this end, we might assess the value for SS of the occurrence of an EE-type event by working out SS's welfare level in the actual world (where presumably an EE-type event did not occur) then SS's welfare level in the closest world in which an EE-type event does occur, and subtracting the first from the second. When the result is a positive value, the occurrence of an EE-type event would be good for SS; when negative, the occurrence would be bad for SS.

In sum, the comparativist view may be stated as follows:

Comparativist Account of Interests:

1. An event EE is in SS's interests just in case EE overall benefits (is good for) SS, making SS's life better than it would have been if EE did not occur, which EE does just when its value for SS is positive.

2. An event EE is against SS's interests just in case EE overall harms (is bad for) SS, making SS's life worse than it would have been if EE did not occur, which EE does just when its value for SS is negative.

3. The occurrence of an EE-type event is in SS's interests just in case it would overall benefit (be good for) SS. The occurrence of an EE-type event would benefit SS if and only if its value for SS is positive.

4. The occurrence of an EE-type event is against SS's interests just in case it would overall harm (be bad for) SS. The occurrence of an EE-type event would harm SS if and only if its value for SS is negative.

5. How much EE benefits (harms) SS depends on how much better (worse) SS's life is in the actual world than it would have been

if EE had not occurred: the better (worse) SS's life is, the more beneficial (harmful) EE is. Similarly, how much the occurrence of an E-type event would benefit (harm) SS depends on how much worse (better) SS's life is in the actual world than it would have been if an EE-type event had occurred: the worse (better) SS's life is, the more beneficial (harmful) the occurrence of an EE-type event would have been.

We sometimes say things that suggest that we can have interests at particular times which we lack at others. For example, we might say that having a tooth drilled by a dentist is not in our interests while we are undergoing the procedure, even though it is in our long-term interests. The idea seems to be that what makes a subject SS better off at time tt is in SS's interests-at-time-tt. But it is important to distinguish *interests-at-tt* from *interests*. What is in our interests-at-time-t1t1 need not be in our interests-at-time-t2t2. This is not true of our interests. Whatever interests we have we have at all times. If something is in our interests, it is timelessly in our interests.

12.3.2 Welfare

Comparativism analyses our interests in terms of our welfare, and is compatible with any number of views of welfare. There are three main ways of understanding welfare itself: positive hedonism, preferentialism, and pluralism. Let us briefly consider each of these three views.

Positive hedonism is the following position:

Positive Hedonism: for any subject S, experiencing pleasure at t is the one and only thing that is intrinsically good for S at t, while experiencing pain at t is the one and only thing that is intrinsically bad for S at t. The more pleasure (pain) S experiences at t, the greater the intrinsic good (evil) for S at t.

Positive hedonism has been defended (by J.S.Mill 1863) on the grounds that it resolves the problem of commensurability. The difficulty arises when we attempt to equate units of different sorts of goods. For example, how do we decide when one unit of love is worth one unit of achievements, assuming that both love and achievements are intrinsically

good? The problem does not arise for hedonists, who evaluate all things in terms of the pleasure and pain that they give us.

However, most theorists consider positive hedonism to be implausible. Nagel argues that I would harm you if I were to cause you to revert to a pleasant infantile state for the rest of your life, yet by hedonist standards I have not harmed you at all. Similarly, it would be a grave misfortune for you if your spouse came to despise you, but for some reason pretended to love you, so that you underwent no loss of pleasure. And Nozick notes that we would refuse to attach ourselves to an Experience Machine that would give us extremely pleasant experiences for the rest of our lives. By hypothesis, the Machine would give us far more pleasure (and less pain) than is otherwise possible. Our reluctance to use the Machine suggests that things other than pleasure are intrinsically good: it is because we do not wish to miss out on these other goods that we refuse to use the Machine.

Preferentialism assesses welfare in terms of desire fulfillment. To desire is to desire that some proposition P hold; when we desire P,P is the object of our desire. According to preferentialism, our welfare turns on whether the objects of our desires hold:

Preferentialism: for any subject S, it is intrinsically good for S at t that, at t,S desires P and P holds; it is intrinsically bad for S at t that, at t,S desires P and ~P holds. The stronger S's desire for P is, the better (worse) it is for S that P holds (~P holds).

In this, its unrefined form, preferentialism is implausible. Many of the things we desire do not appear to contribute to our welfare. Consider, for example, Rawls' famous example of the man whose main desire is to count blades of grass. In response to the grass counter case, Rawls (1971) adopts critical preferentialism, which says that welfare is advanced by the fulfillment of rational aims. Assuming that counting grass blades is irrational because it is pointless, fulfilling the desire to count grass blades is not intrinsically good. However, even critical preferentialism seems vulnerable to attack, since the fulfillment of rational desires need not advance one's welfare. Parfit (1984) illustrates the point by supposing that you have the (rational) desire that a stranger's disease be overcome: the fulfilment of this desire advances the stranger's welfare, not yours. This

example can be handled by egocentric preferentialism, which says that only desires that make essential reference to the self can advance our welfare when fulfilled (Overvold 1980). Thus the fulfilment of my desire that I be happy is intrinsically good for me, but the fulfilment of my desire that somebody or other be happy is not. A further variant of preferentialism might be called achievement preferentialism. This view says that subject S's accomplishing one of S's goals (or ends) is intrinsically good for S, and that being thwarted from accomplishing such a goal is intrinsically bad for S (Scanlon 1998; Keller 2004; Portmore 2007).

Pluralism is the third main account of welfare. Pluralists can agree with the hedonist position that a person's pleasure is intrinsically good for that person, and with the preferentialist's view that the fulfilment of a person's desire is intrinsically good for that person. However, pluralism says that various other sorts of things are intrinsically good, too. Some traditional examples are wisdom, friendship and love, and honor. Another example might be engaging in self-determination.

12.4 THE HARM THESES

Typically, those who value life accept the harm thesis: death is, at least sometimes, bad for those who die, and in this sense something that 'harms' them. (Several theorists, including Barbara Levenbook (2013), emphasize that, in one sense of the term 'harm', events that are only partially bad for me might be said to harm me. In what follows the term 'harm' will be restricted to events that are overall bad for me.) It is important to know what to make of this thesis, since our response itself can be harmful. This might happen as follows: suppose that we love life, and reason that since it is good, more would be better. Our thoughts then turn to death, and we decide it is bad: the better life is, we think, the better more life would be, and the worse death is. At this point, we are in danger of condemning the human condition, which embraces life and death, on the grounds that it has a tragic side, namely death. It will help some if we remind ourselves that our situation also has a good side. Indeed, our condemnation of death is here based on the assumption that more life would be good. But such consolations are not for everyone. (They are

unavailable if we crave immortality on the basis of demanding standards by which the only worthwhile projects are endless in duration, for then we will condemn the condition of mere mortals as tragic through and through, and may, as Unamuno (1913) points out, end up suicidal, fearing that the only life available is not worth having.) And a favorable assessment of life may be a limited consolation, since it leaves open the possibility that, viewing the human condition as a whole, the bad cancels much of the good. In any case it is grim enough to conclude that, given the harm thesis, the human condition has a tragic side. It is no wonder that theorists over the millennia have sought to defeat the harm thesis. We will examine their efforts, as well as the challenges to the posthumous harm thesis, according to which events occurring after we die can harm us. First, however, let us see how the harm theses might be defended.

12.4.1 The Main Defense

Those theorists who defend the harm theses typically draw upon some version of comparativism (e.g., Nagel 1970, Quinn 1984, Feldman 1991). According to comparativism, a person's death may well harm that person. Death may also be harmless. To decide whether a person's death is bad for that person, we must compare her actual welfare level to the welfare level she would have had if she had not died. Suppose, for example, that Hilda died on December 1, 2008 at age 25 and that, had she not died, she would have prospered for 25 years and suffered during her final five years. To apply comparativism, we must first select an account of welfare with which to assess Hilda's well-being. For simplicity, let us adopt positive hedonism. The next step is to sum the pleasure and pain she had over her lifetime. Suppose that she had considerably more pleasure than pain. We can stipulate that her lifetime welfare level came to a value of 250. Next we sum the pleasure and pain she would have had if she had not died on December 1, 2008. The first 25 years of her life would be just as they actually were, so the value of these would be 250. We can suppose that her next 25 years would also receive the value of 250. And let us stipulate that her final 5 years, spent mostly suffering, carry a value of −50. Then, had she not died, her lifetime welfare level would have been 250+250−50=450. Subtracting this value from her actual lifetime welfare

level of 250 gives us −200. This is the value for her of dying on December 1, 2008. According to comparativism, then, her death was quite bad for her. Things would have been different if the last 30 years of her life would have been spent in unrelenting agony. On that assumption, her death would have been good for her.

Our example concerned a particular death at a specific time. Comparativism also has implications concerning whether dying young is bad for the one who dies, and whether it is bad for us that we die at all. In both cases the answer depends on how our lives would have gone had we not died. Usually dying young deprives us of many years of good life, so usually dying young is bad for us. As for whether or not it is bad to be mortal, that depends on whether the life we would lead as an immortal being would be a good one or not.

According to comparativism, when death is bad for us, it is bad for us because it precludes our coming to have various intrinsic goods which we would have had if we had not died. We might say that death is bad for us because of the goods it deprives us of, and not, or at least not always, because of any intrinsic evils for which it is responsible.

So much for the harm thesis. Now let us ask how the posthumous harm thesis might be defended.

Note first that we must reject the posthumous harm thesis if we adopt positive hedonism and combine it with comparativism, for nothing that happens after we die can boost or reduce the amounts of pleasure or pain in our lives.

However, posthumous events might well be bad for us on other accounts of welfare. Suppose that I want to be remembered after I die. Given preferentialism, something could happen after I die that might be bad for me, namely my being forgotten, because it thwarts my desire.

These ways of defending the harm theses seem quite plausible. Nevertheless, there are several strategies for criticizing the harm theses. Let us turn to these criticisms now, starting with some strategies developed in the ancient world by Epicurus and his follower Lucretius.

12.4.2 The Symmetry Argument

One challenge to the harm thesis is an attempt to show that the state death puts us in, nonexistence, is not bad. According to the symmetry argument, posed by Lucretius, a follower of Epicurus, we can prove this to ourselves by thinking about our state before we were born:

Look back at time … before our birth. In this way Nature holds before our eyes the mirror of our future after death. Is this so grim, so gloomy? (Lucretius 1951)

The idea is clear to a point: it is irrational to object to death, since we do not object to pre-vital nonexistence (the state of nonexistence that preceded our lives), and the two are alike in all relevant respects, so that any objection to the one would apply to the other. However, Lucretius' argument admits of more than one interpretation, depending on whether it is supposed to address death understood as the ending of life or death understood as the state we are in after life is ended (or both).

On the first interpretation, the ending of life is not bad, since the only thing we could hold against it is the fact that it is followed by our nonexistence, yet the latter is not objectionable, as is shown by the fact that we do not object to our nonexistence before birth. So understood, the symmetry argument is weak. Our complaint about death need not be that the state of nonexistence is ghastly. Instead, our complaint might be that death brings life, which is a good thing, to an end, and, other things being equal, what deprives us of good things is bad. Notice that the mirror image of death is birth (or, more precisely, becoming alive), and the two affect us in very different ways: birth makes life possible; it starts a good thing going. Death makes life impossible; it brings a good thing to a close.

Perhaps Lucretius only meant to argue that the death state is not bad, since the only thing we could hold against the death state is that it is nonexistence, which is not really objectionable, as witness our attitude about pre-vital nonexistence. So interpreted, there is a kernel of truth in Lucretius' argument. Truly, our pre-vital nonexistence does not concern us much. But that is because pre-vital nonexistence is followed by existence. Nor would we worry overly about post-vital nonexistence if it, too, were followed by existence. If we could move in and out of existence, say with the help of futuristic machines that could dismantle us,

then rebuild us, molecule by molecule, after a period of nonexistence, we would not be overly upset about the intervening gaps, and, rather like hibernating bears, we might enjoy taking occasional breaks from life while the world gets more interesting. But undergoing temporary nonexistence is not the same as undergoing permanent nonexistence. What is upsetting is the permanence of post-vital nonexistence—not nonexistence per se.

There is another way to use considerations of symmetry to argue against the harm thesis: we want to die later, or not at all, because it is a way of extending life, but this attitude is irrational, Lucretius might say, since we do not want to be born earlier (we do not want to have always existed), which is also a way to extend life. As this argument suggests, we are more concerned about the indefinite continuation of our lives than about their indefinite extension. (Be careful when you rub the magic lamp: if you wish that your life be extended, the genie might make you older!) A life can be extended by adding to its future or to its past. Some of us might welcome the prospect of having lived a life stretching indefinitely into the past, given fortuitous circumstances. But we would prefer a life stretching indefinitely into the future.

Is it irrational to want future life more than past life? No; it is not surprising to find ourselves with no desire to extend life into the past, since the structure of the world permits life extension only into the future, and that is good enough. But what if life extension were possible in either direction? Would we still be indifferent about a lengthier past? And should our attitude about future life match our attitude about past life?

Our attitude about future life should match our attitude about past life if our interests and attitudes are limited in certain ways. If quantity of life is the only concern, a preference for future life is irrational. Similarly, the preference is irrational if our only concern is to maximize how much pleasure we experience over the course of our lives without regard to its temporal distribution. But our attitude is not that of the life- or pleasure-gourmand.

According to Parfit, we have a far-reaching bias extending to goods in general: we prefer that any good things, not just pleasures, be in our future, and that bad things, if they happen at all, be in our past. He argues

that if we take this extensive bias for granted, and assume that, because of it, it is better for us to have goods in the future than in the past, we can explain why it is rational to deplore death more than we do our not having always existed: the former, not the latter, deprives us of good things in the future (he need not say that it is because it is in the past that we worry about the life-limiting event at the beginning of our lives less than the life-limiting event at the end). This preference for future goods is unfortunate, however, according to Parfit. If cultivated, the temporal insensitivity of the life- or pleasure-gourmand could lower our sensitivity to death: towards the end of life, we would find it unsettling that our supply of pleasures cannot be increased in the future, but we would be comforted by the pleasures we have accumulated.

Whether or not we have the extensive bias described by Parfit, it is true that the accumulation of life and pleasure, and the passive contemplation thereof, are not our only interests. We also have active, forward-looking goals and concerns. Engaging in such pursuits has its own value; for many of us, these pursuits, and not passive interests, are central to our identities. However, we cannot make and pursue plans for our past. We must project our plans (our self-realization) into the future, which explains our forward bias. (We could have been devising and pursuing plans in the past, but these plans will not, I assume, be extensions of our present concerns.) It is not irrational to prefer that our lives be extended into the future rather than the past, if for no other reason than this: only the former makes our existing forward-looking pursuits possible. It is not irrational to prefer not to be at the end of our lives, unable to shape them further, and limited to reminiscing about days gone by. As Frances Kamm (1998) emphasizes, we do not want our lives to be all over with.

Nevertheless, it does not follow that we should be indifferent about the extent of our pasts. Being in the grip of forward-looking pursuits is important, but we have passive interests as well, which make a more extensive past preferable. Moreover, having been devising and pursuing plans in the past is worthwhile. If fated to die tomorrow, most of us would prefer to have a thousand years of glory behind us rather than fifty. We want to have lived well.

12.4.3 Epicurean Challenges: Death Cannot Affect Us

Further challenges to the harm theses are offered by Epicurus (341–270) in his *Letter to Menoeceus*:

Death …, the most awful of evils, is nothing to us, seeing that, when we are, death is not come, and, when death is come, we are not.

We might restate Epicurus' brief argument as follows: if death harms the individual who dies, there must exist a *subject* who is harmed by death, a clear *harm* that is received, and a *time* when that harm is received. As to the timing issue, there seems to be two possible solutions, given that death follows immediately upon life: either death harms its victims while they are alive or later. If we opt for the second solution we appear to run head on into the problem of the subject, for assuming that we do not exist after we are alive, no one is left to incur harm. We also encounter the problem of specifying a harm that might be accrued by a nonexistent person. If we opt for the first solution—death harms its victims while they are alive— we have a ready solution to the problem of the subject but we face the problem of supplying a clear way in which death is bad: death seems unable to have *any* ill effect on us while we are living since it will not yet have occurred. Seeing that there is no coherent solution to all three issues, Epicurus rejects the harm thesis.

Epicurus focuses on death, but if his argument is good, it applies more generally, to include all events that follow death.

In some respects Epicurus's argument is not clear. One problem is that what he means by 'death' is unclear. For now let us assume that he meant to refer to the process by which our lives are ended. Another interpretive problem arises as well: his intent might be to show that neither death nor posthumous events can *affect* us *at all*. From this claim it would follow that death and posthumous events are harmless, assuming that an event harms us only if it somehow affects us at some time (perhaps well after it occurs).

Let us see if it is possible to show that death and posthumous events do not affect us. Then we can try out (in the next section) a weaker thesis: that death and posthumous events cannot affect us *in a way that is bad for*

us. This weaker claim is easier to defend, but the stronger claim is worth exploring.

We can start with some assumptions about when an event can affect us. To this end, let us adopt the *causal account of responsibility*:

    a. An event (or state of affairs) can affect some subject (person or thing) SS only by having a causal effect on SS (the *causal impact* thesis).

    b. A subject SS cannot be causally affected by an event while SS is nonexistent.

    c. A subject cannot be causally affected by an event before the event occurs (the *ban on backwards causation*).

From the causal account, together with some plausible assumptions, it follows that a *post*-mortem event, such as the burning of one's corpse, cannot affect us after we are dead, since, by (a), to be affected is to be affected causally, but, by (b), nonexistent people cannot be causally affected by any event. It also follows that the state of being dead cannot affect us while we are dead. Here we are assuming that people cease to exist when they die (the termination thesis). From the causal account it also follows that neither being dead, nor any events that follow, can affect us while we are alive, given the ban on backwards causation:

    1. An event can affect us only by causally affecting us (the causal impact thesis).

    2. We cannot be causally affected by an event while we are nonexistent.

    3. We do not exist while dead (the termination thesis).

    4. So neither being dead, nor any posthumous event, can affect us while we are dead (by 1–3).

    5. We cannot be causally affected by an event before the event occurs (the ban on backwards causation).

    6. So neither being dead, nor any posthumous event, can affect us while we are alive (by 1 and 5).

    7. So neither being dead, nor any posthumous event, can ever affect us (by 4 and 6).

So far so good: neither the state of being dead nor any post-mortem event can ever affect us. However, it has not been shown that we cannot be

affected by the dying process. Of course, the thesis that we must exist to be affected, together with the termination thesis, rule out the possibility that death affects us *after* it occurs (after we are nonexistent). And the ban on backwards causation rules out the possibility that death affects us before it occurs. Thus:

8. Death cannot affect us after it occurs (by 1–3).

9. Death cannot affect us before it occurs (by 1 and 5).

10. So death can affect us, if at all, only when it occurs (by 8 and 9).

But nothing said so far rules out the possibility that death affects us exactly *when* it occurs. In particular, the problem of the subject does not arise since it is a living, existing person who is harmed by death while it occurs. Is there any way to establish that death cannot affect us even at the time it occurs? There might be two ways. First, we might claim that death occurs only after we are nonexistent. This assumption has the odd consequence that death can affect us only if posthumous events can. It will follow from (7) that death cannot ever affect us. Second, we might claim that death is instantaneous; it happens too quickly to affect us.

Some theorists have indeed defined 'death'—the ending of life—in such a way as to imply that it occurs only after we are nonexistent. For example, Feinberg (1984), following Levenbook (1984), defines death as "the first moment of the subject's nonexistence." Perhaps this definition is motivated by the awkwardness of attaching 'death' to a moment in the dying process when a spark of life persists. However, it is at least as awkward to attach 'death' to a moment after the dying process is over—to suggest that the ending of life occurs while we are in a state of death. It is also to concede too much to the Epicurean, who could then establish that death is no evil merely by showing that posthumous events are innocuous. What about the suggestion that death happens too quickly to affect us?

Recall that 'death' can be used in the process as well as the denouement sense (Section 1). Death, in the process sense, unfolds over a period of time, and it obviously affects us while it occurs—even if instantly.

What if we opt for the denouement sense of 'death'? Is it plausible to say that losing the very last of life can have no affect on us? It is difficult to see why. If we were correct when we said that the complete destruction of

our vital capacities affects us, surely we are also affected, albeit less, by losing the very last of the vital capacities that sustain us.

Let's review. Granting them some leeway, Epicureans can show:

    11. Neither being dead, nor any posthumous event, can ever affect us, and the dying process itself can affect us, if at all, only while it occurs (by 7 and 10).

They can then argue as follows:

    12. An event harms us only if it somehow affects us at some time.

    13. So neither being dead, nor any posthumous event, can harm us, and the dying process can harm us, if at all, only while it occurs (by 11 and 12).

For a more rigorous presentation of the above argument, see the supplementary document:

But Epicureans lack a convincing argument against the possibility that the dying process and some of its effects overlap in time; hence they cannot refute the harm thesis. We have a subject, harm, and time: the subject of death is a live creature who endures its effects at the very time the creature dies.

12.4.4 Epicurean Challenges: Death Is Harmless

Instead of trying to establish that death cannot affect us at all, Epicureans might argue that death cannot affect us *in a way that is bad for us.* To that end, they can provide a condition for something's being bad for us and argue that death fails to meet it.

The condition which Epicurus himself supplied is this: an event (or state of affairs) harms us only if it causes in us the presence of some condition we find unpleasant. For simplicity, we can call all such conditions pain or suffering. That condition, the suffering, need not occur at the same time as the event that causes its presence in us. An event may occur long before it has any direct impact on us; it may occur even before we exist, as when someone times a bomb to go off 150 years later, killing everyone around. Epicurus himself did not spell out a complete view of welfare. He did not make it entirely clear when things are, overall, beneficial or harmful to a person. But he surely did think that something harms us only if it causes us to suffer.

Some theorists prefer to phrase Epicurus's condition in terms of experience, thusly: we are harmed only by what we experience. Given that one never experiences one's death, it would follow that it cannot harm those who die. A variant of this *experience condition* was proposed by Rosenbaum (1986): we are harmed only by what we *can* experience. Other theorists state Epicurus's condition in terms of existence, thus attributing to him the *existence condition*: something harms us only at times when we exist. (For good discussions of the experience condition and its plausibility, see Nussbaum 2013, Silverstein 2013, and Fischer 2014.)

On the Epicurean view, clearly neither the state nor process of death is *inherently* harmful—it is, in itself, not bad for us. For death is not necessarily painful. One can die painlessly, as when one dies while unconscious. But Epicurus did not say merely that death *need* not be harmful; he claimed that death was never harmful; on his criterion, this means that death never causes the subject to suffer.

To show that death can have no salient effect on us, Epicureans might argue that death cannot be responsible for *any* condition's presence in us, salient or otherwise. It can only be responsible for our *ceasing* to be in a condition. However, this thesis is clearly false on the process sense of 'death:' moving from being wholly alive to completely lacking life might well introduce the presence of some bad condition in us, such as pain. No doubt Epicureans gravitate to the denouement sense of death since the ending of the final trace of life might occur extremely quickly, perhaps so quickly that it has no salient effect on us while it happens. Nevertheless, Epicureans may argue, with some degree of plausibility, that denouement death cannot harm us:

1. Denouement death occurs too quickly to be responsible for the presence of any unpleasant condition in us at the time it occurs.
2. Only something responsible for the presence of an unpleasant condition in us is harmful to us.
3. So denouement death cannot harm us at the time it occurs (by 14 and 15).

By combining 16 with 13, established earlier, Epicureans may conclude that:

17. Neither posthumous events nor the state of death nor denouement death may ever harm us, and process death may harm us only while it occurs.

However, this conclusion will disappoint people who wonder whether dying is a misfortune: they want to know whether *losing their lives* is a bad thing, not just whether, having nearly completely lost life, it is bad to lose the very last of it (Luper 2004). Even for Epicurus himself this conclusion is not entirely adequate. For it leaves in place the possibility that the dying process can be harmful.

So why did Epicurus say that death is nothing to us? He surely knew that the dying process can be harmful to us. One possibility is that he did not really intend to show that death is innocuous. Many commentators insist that he wanted only to show that *being dead*, that is, the state of death, is nothing to us, and that he realized that dying is often a misfortune. It is also possible that Epicurus did not believe that what we have called 'process death' is part of death; instead, death is what we have called 'denouement death'. This line of thought would position him to admit that 'process death' is bad for us, but it is only the precursor to death.

However, if Epicurus meant to show only that denouement death is harmless, or that the state of being dead is harmless, his efforts are disappointing given his own goal, which was to enable us to achieve ataraxia, or complete tranquility. He cannot reach this goal if he does not free us from our concern about the dying process or the events leading up to the dying process.

The best Epicurus could do is to downplay the painfulness of process death and its cause, and this he appears to do:

Continuous pain does not last long in the flesh; on the contrary, pain, if extreme, is present a very short time. … Illnesses of long duration even permit of an excess of pleasure over pain in the flesh (Principal Doctrines, Doctrine 4)

Unfortunately, Epicurus was wrong; the dying process and its cause can be excruciating.

There are things other than death that seem bad for us. To prepare us for complete tranquility in the fact of these things, Epicurus would need to

address them as well. Let us consider some examples, and what Epicurus might say about them.

One example is obvious: we suffer when we anticipate death. Epicurus would probably admit that anticipating death is a bad thing if it upsets us. But he emphasizes that our (present) anticipatory fear is not caused by our (future) death, since future events are powerless to affect the past. Hence, by the painfulness criterion, the fear of death is not grounds for saying that death is harmful. Moreover, fear is irrational unless its object is genuinely evil in some way, which death is not:

He speaks idly who says the he fears death, not because it will be painful when present but because it is painful in anticipation. For if something causes no distress when present, it is fruitless to be pained by the expectation of it (Letter to Menoeceus).

Something else that is related to death seems bad for us: namely, the grief others experience when we die. But Epicurus would urge us to distinguish what is bad for us from what is bad for others. At most, the fact that your family grieves at your death supports the claim that your demise harms them, not that it harms you. (Too, your distress at anticipating your family's grief over your death is not grounds for you to regard your death as a bad thing: the suffering your death brings them cannot affect you, and your anticipatory grief is irrational.) Furthermore, their grief should be mitigated by the fact that your death is not bad for you. Their grief is entirely self-centered, exactly like the self-pity a stamp collector might feel at the destruction of a treasured stamp, in that the stamp is not harmed by its own destruction.

These examples illustrate that Epicurus can address some death-related concerns by showing that they are misguided, if we grant him his claim that we can be harmed only by what causes us to suffer. However, some death-related concerns cannot be handled this way. For example, the fact that everyone dies causes us distress and is therefore harmful to us even on Epicurus' criterion. At most Epicurus can say that mortality need not be harmful to us, and that it will not be if we can manage not to be distressed by it (Luper 2009).

12.4.5 Further Objections to the Harm Theses

Epicurus' case against the harm theses hinged on the assumption that we can be harmed only by what causes us to suffer. However, in section 3.1 we supported the harm theses by combining comparativism with one of the three leading accounts of welfare. We noted that death and posthumous events seem harmful because they deprive us of goods we would otherwise have had. If our argument was correct, then Epicurus' assumption must be mistaken. It must be false that harm requires incurring pain. Instead, harm can consist in being deprived of goods.

Are there ways Epicureans could resist the view that being deprived of goods can be bad for us? Perhaps. Epicureans could criticize comparativism. They could also defend some view of welfare that is more congenial to their position. Let us consider each strategy, starting with the second.

Epicurus may have accepted the following view of welfare:

Negative Hedonism:

for any subject S,SS,S's experiencing pain is the one and only thing that is intrinsically bad for SS, and nothing is intrinsically good for SS.

When paired with comparativism, this view has implications that Epicurus would have welcomed. It implies that harm is limited to what increases our pain, and benefit is limited to what reduces it. Consequently death is harmless to those who die painlessly, no matter how good the life they would have had would have been. Moreover, death can be beneficial: it can preclude our suffering.

However, the implications of negative hedonism are quite absurd. For example, it implies that we never have reason to endure pain for the sake of pleasure or any other good. It also implies that we should end our lives as quickly and painlessly as we can since living on will harm us and cannot possibly benefit us. The negative hedonist account of welfare is clearly false.

Perhaps Epicureans would have better success if they were to reject comparativism itself. To that end, they might adopt one of four strategies which we will discuss in turn.

Bifurcated Comparativism

It is quite possible that Epicurus himself rejected comparativism, as formulated above. Perhaps he thought that the harmfulness of an

event EE is not a matter of the good it deprives us of, but rather a matter of how much intrinsic harm it causes, and the goodness of EE is a matter of how much intrinsic good it causes. Earlier we let $B(S,W)$ stand for the sum of the values of the things which are intrinsically bad for S in world W, and we let $G(S,W)$ stand for the sum of the values of the things which are intrinsically good for S in world W. Using this symbolism, we can state the following alternative to comparativism:

Bifurcated Comparativism:

EE harms S if and only

if $B(S,WE)<B(S,W{\sim}E)$; EE benefits S if and only

if $G(S,WE)>G(S,W{\sim}E)$.

Bifurcated comparativism implies that goods do not offset evils, but might eliminate them: that is, the goods EE brings do not reduce the harmfulness of EE unless they cause us to have less pain or less of some other evil. Similarly, evils do not offset goods. Combined with positive hedonism, bifurcated comparativism implies that we are harmed only by what increases our suffering, and benefitted only by what increases our pleasure; all else is a matter of indifference. Epicurus might have been drawn to this combination because it implies that death can neither harm nor benefit us, ignoring the pain it can cause while it occurs.

However, bifurcated comparativism is implausible. One problem is its implication that that goods and evils do not offset each other. Another worry is that surely some events or states of affairs harm us without causing us pain or some other intrinsic evil, and benefit us without giving us pleasure or some other intrinsic good. It is better to be anaethetized before surgery, but not if bifurcated comparativism is true. Moreover, if I slip into a temporary coma, which precludes my suffering from injuries inflicted upon me in a car crash, the coma benefits me, even though it does not give me pleasure or other goods. Similarly, a coma that precludes my enjoying a week's worth of good life harms me, yet gives me no pain or other evils.

Surely death is capable of benefitting us the same way that anesthetization and unconsciousness can. It can preclude our enduring great suffering. Similarly, like anesthetization and unconsciousness, death can harm us by

precluding our living well. Comparativism gets things right and bifurcated comparativism gets things wrong in all of these examples.

Temporal Relativism

Comparativism assesses our interests in a temporally neutral way. It implies that, at each point in my life, it is in my interests that my welfare be as high as possible across my entire life, so that it is prudent for me now to do what will boost my welfare later, other things being equal. Famously, Derek Parfit (compare McMahan 2002) supplies grounds for assessing our interests in a temporally relative way instead of a temporally neutral way. Assessing our interests in a temporally relative way may help Epicureans to undermine the harm theses.

Consider that sometimes we have no reason whatever to satisfy a desire. Parfit gives two examples. First, a desire might be implicitly conditional on its own persistence, in the sense that we want to satisfy it only on condition that we still have it. The desire to play cards is like this. We lose all reason to satisfy such desires as soon as we cease to have them. Compare desires, mentioned earlier, that are conditional on our persistence. We might have reason to satisfy these right up until our last day, even if we cease to have them much earlier. Second, Parfit notes, we might change our values or ideals, which might lead us to condemn some of our desires. In this case it is reasonable to forego any opportunity to satisfy them. When a property, such as conditionality, undermines the importance of satisfying a desire for PP, so that PP's holding is not intrinsically good for us (and ~P~P's holding is not intrinsically bad for us), let us say that it is an undermining feature.

When we no longer want something, we may speak of a past desire. Perhaps a desire is undermined by being past, as Parfit has claimed (compare Suits 2001). Then Epicureans may be able to revive their attack on the posthumous harm thesis: dying ensures that we cannot be harmed by posthumous events, since we are without desires long before these occur (Vorobej, 1998). This strategy does not seem to vindicate death itself, since death may preclude the fulfillment of some of the very desires it destroys. However, the die-hard Epicurean might suggest that a desire is undermined, in passing, at the very moment of its destruction; if it is later thwarted, no harm is done.

In any case, it is far from clear that our interests should be assessed in a temporally relative way. The matter is quite controversial. Consider Parfit's claim that our desires are undermined by their pastness: neutralists, who assess our interests in the temporally neutral way prescribed by comparativism, can resist Parfit's claim by finding a feature other than pastness that tends to undermine desires that we no longer have. One possibility becomes evident once we notice that most of our aims are tentative in the sense that we adopt them in the expectation that we may later revise them. An extreme way to revise a desire for PP is to stop wanting PP altogether—to end the desire for PP, say on the grounds that it conflicts with other, more pressing interests. We defer to future exercises of our own autonomy, realizing that we may reassess our priorities, until our life plan matures. In particular, we are always prepared to revise desires in light of the projects and commitments with which we identify, and loath to abandon projects and commitments which have become parts of our identities. We favor some of the ways our desires change, and take what steps we can to coax them in preferred directions. As a rough approximation, we may say that, unless our desires change in ways we (do or) would oppose, the changes are voluntary (Cf. Harry Frankfurt 1971). For our purposes we can even count, as voluntary, the intentional elimination of a desire using artificial means, as when we take pills to remove the desire to smoke cigarettes. If we voluntarily stop wanting P,~PP,~P can no longer harm us. It will not harm us during the time we wanted PP, or later, when our desire is thwarted. So we undermine a desire when we voluntarily abandon it (Luper 1987). On this view, Epicureans cannot show that desires are harmlessly thwarted by death and posthumous events on the grounds that such desires are past at the time death or posthumous events thwart them.

Actualist Comparativism

Comparativism says that the value of my dying at time tt depends on the intrinsic goods (and evils) I would have accrued after tt had I not died, even though I am actually dead after tt. Being dead, I am incapable of accruing any intrinsic goods or evils after tt, and in that sense I am unresponsive after tt. *Interest actualism* denies that the value of my dying at tt can depend on these goods. It says that the value for SS of

event EE is not affected by the intrinsic goods or evils SS would have accrued had EE not occurred, if SS would have accrued them after SS has actually become unresponsive.

Accepting interest actualism would force us to modify comparativism. The actualist view would be this:

Actualist Comparativism: EE's value for SS equals the intrinsic value for SS of SS's life in WEWE, the actual world in which EE occurs, minus the intrinsic value for SS of SS's life in the closest world, W∼EW∼E, in which EE does not occur excluding any intrinsic value SS would accrue in W∼EW∼E after SS ceases to be responsive in WEWE.

However, actualist comparativism does not appear to be more plausible than standard comparativism. If I die at tt, accruing goods after tt is not in my interests-after−t−t, but it does not follow that it is not in my interests. If developing and fulfilling certain desires is entailed in making my life as a whole as good as possible, then it is in my interests to develop and fulfill those desires. Even though I will die before I develop and fulfill the desires, it is in my interests to develop and fulfill them, and bad for me not to develop and fulfill them.

Moot Preclusion

One other line of thought might be pressed against the comparativist account of interests. Comparativism says that something harms me when it makes my life worse than it would have been. However, there seem to be events and states of affairs that do not harm me even though their value for me is negative. I am not harmed, it seems, by failing to be a genius, or rich and beautiful. But compare my life as it is, with my unimpressive IQ, income and looks, to my life as it would be were I brilliant or rich or beautiful: the former is considerably worse than the latter. By failing to be brilliant, rich and beautiful, I am precluded from having many goods, but we might say that the preclusion is moot, in the sense that it is harmless to me. Epicureans might renew their attack on the harm thesis by exploiting examples like these. The examples appear to show that things can have enormous negative value for me without harming me. Similarly, Epicureans might insist, the preclusion of goods by death is moot: cut short, my life is worse than it would be were I not to die, but this comparative difference does not show that I am harmed.

It seems that the comparative criteria work well when we evaluate losses, such as the loss of my arms, and also when we evaluate some lacks, such as the inability to see or to feel pleasure. But the criteria have worrisome implications when we evaluate certain other lacks, such as my lack of genius. It is relatively clear that a person is harmed by the inability to see but less clear that he is harmed by the lack of genius. Why is that?

There are various responses to the problem of moot preclusion. One is to deny that it makes any sense to speak of 'negativities', or events that consist in things not happening. This does not stop us from evaluating the event or process of dying (as opposed to the state of death) which is not a 'negativity'. Comparativists are right to claim that things harm us by making our lives worse than they would have been otherwise; negativities are not counterexamples, since they do not exist. Another response is that moot preclusion involves cases in which the events or states of affairs that would be good for us if they held are highly improbable (Draper 1999). A further explanation involves the relative importance of having some goods rather than others. In some moods, we may consider it harmful to be deprived of a good just when it is important for us to have it. The troublesome lacks we have been discussing might be lacks of goods it is unimportant to have; such lacks would not be harmful even though we would be better off without them.

12.5 IS DEATH ALWAYS A MISFORTUNE?

Are all deaths misfortunes? Perhaps, but there is a strong case to the contrary.

12.5.1 Only Premature Death Is a Misfortune

To support the conclusion that death is not always a misfortune, we might adopt some version of preferentialism. Perhaps it is not bad to die at an advanced enough age, for people who live long enough may be ground down by life until they give up many of their goals. Also, they will have attained many of their aspirations. If already satisfied, or given up, a desire cannot be thwarted, even by death, so as we lose our motivation for living, death ceases to be objectionable to us. Perhaps death is bad for us

only if premature in the sense that it comes when we still have interests such as salient desires that propel us forward in life, and only if meeting these interests is a real prospect.

12.5.2 Immortality Is a Misfortune

We are left to wonder whether death would ever cease to be objectionable were we not ravaged by bad health and other hardships. Bernard Williams (and others, such as Kagan 2012) argues that it would be bad to live forever, even under the best of circumstances. His view is based on an assumption about the relationship between our identities and the desires that motivate us to live.

Consider a woman who wants to die. She might still take the view that if she is to live on, then she should be well fed and clothed. She wants food and clothing on condition she remain alive. In this sense her desires are conditional, and do not give her reason to live. Contrast a father who is committed to rearing a beloved daughter: he desires unconditionally that the child do well, and his desire gives him reason to live, because he can rear his child only if he survives. In this sense, his desire is categorical, or unconditional. Williams thinks that categorical desires are essential to identity, and give meaning to life. Through categorical desires, we are attached to projects or relationships that are definitive of the self; faced with their destruction, we would feel our lives are meaningless, and that in an important sense we cannot survive as the persons we once were.

The bearing on death, according to Williams, is, first, that people have good reason to condemn a death that is premature in the sense that it thwarts their categorical desires. Second, mortality is good, since people who live long enough eventually will lose the categorical desires with which they identify. Life will lose its novelty, and oppressive boredom will set in. To avoid ennui, superseniors would have to replace their fundamental desires, again and again. But this is to abandon their identities; it is tantamount to death.

As Williams says, lives of unimaginative routine will eventually grow stale if extended long enough. Of course, this is not supposed to comfort ordinary mortals, most of whom will die long before routine undermines

the joy in living. However, as several theorists, including Nagel (1986, p. 224, n. 3) Glover (1977, p. 57), and Fischer (1994) have suggested, it is not obvious that life must become dull. Williams may have overlooked how rich and complex life can be, especially for superseniors who pursue multiple open-ended projects in the company of other superseniors. His response to this kind of criticism is that even rich and open-ended projects eventually will become routine (say after a few billion years), so our pursuits must be replaced periodically if we are to remain interested in life. But to phase in wholly new projects is to lose our identity.

Williams's response faces objections. First, we might avoid boredom by adding to our pursuits, and varying the way we approach them, without abandoning certain core interests that define us. Second, Williams is working with a view of identity that may be too narrow. Many of us would welcome a possibility that he downplays: gradually transforming our interests and projects over time. Transformation is not death. It is distinct from, and preferable to, annihilation. Transformation would be death only if identity were wholly a matter of retaining (most of) our psychological features over time. However, it is questionable that persistence requires this kind of connectedness. Even if our persistence hinges on our psychological features, transformation need not be death, since transformation is consistent with the gradual, continuous change of our psychological features. If we could live endlessly, the stages of our lives would display reduced connectedness, yet they could be continuous, which is a property that is important in the kind of survival most of us prize. Even after drinking at the fountain of eternal youth, we would tend to focus on relatively short stretches of our indefinitely extensive lives, and over these periods we would prize connectedness, since we are animated by specific projects and relationships that can be developed only if there are strong interconnections among the temporal stages of our lives. However, sometimes we would turn our attention to relatively long stretches of life, and then, prizing continuity, we would phase in new and worthwhile undertakings that build upon, and do not wholly replace, the old.

12.6 CAN DEATH'S HARMFULNESS BE REDUCED

Even if death is usually bad for those who die, perhaps it need not be bad for us, if we prepare ourselves suitably. This might be possible if some form of preferentialism is true, and if, by altering our desires, we could cease to have any interests that dying would impair. For then we might be able to thanatize our desires, in this sense: abandon all desires that death might thwart. Among these are desires we can satisfy only if we live on for a few days, but also desires we cannot possibly satisfy within the span of a normal lifetime, and the desire for immortality itself. Instead of desiring that some project of mine succeed, which is a desire that might be thwarted by my death, I might instead adopt a conditionalized version of this desire, namely: should I live on, let my project succeed. If all goes well, thanatizing would insulate us from harm from death by leaving us with no interests with which dying interferes.

Check Your Progress 1

Note: a) Use the space provided for your answer

b) Check your answers with those provided at the end of the unit

1.   What do you know the concept of life and death?

_____

_____

_____

2.   Discuss about the misfortune.

_____

_____

_____.

3.   Highlight The Harm Theses.

_____

_____

_____

12.7 LET US SUM UP

Unfortunately, this strategy will backfire. The main problem is that death can interfere with desire fulfillment not just by falsifying the objects of our desires but also by precluding our having desires (Luper 2013). So even if we resolve, from now on, to limit ourselves to desires whose objects cannot be falsified by death, we are still vulnerable to the harm death will do us if it precludes our having and fulfilling desires. Hence thanatizing would force us to avoid having any desires whose fulfillment would have benefitted us, and to deny ourselves such desires would be as bad for us as the harm we are trying to avoid.

However, the core idea of adapting our desires is useful, if not taken to an extreme. It is prudent to avoid taking on goals we cannot possibly attain, and hence prudent to eschew projects that cannot possibly be completed during the course of a normal lifetime.

12.8 KEY WORDS

Misfortune: an unfortunate condition or event.

Harm: physical injury, especially that which is deliberately inflicted.

Reduce: make smaller or less in amount, degree, or size.

12.9 QUESTIONS FOR REVIEW

1. Is Death Always a Misfortune?
2. Can Death's Harmfulness be reduced?

12.10 SUGGESTED READINGS AND REFERENCES

- Bedau, M., 2014. "The Nature of Life" in S. Luper (ed.), The Cambridge Companion to Life and Death, Cambridge: Cambridge University Press.

- Belshaw, C., 2009, Annihilation: The Sense and Significance of Death, Dublin: Acumen Press.

- Braddock, G., 2000. "Epicureanism, Death, and the Good Life," Philosophical Inquiry, 22 (1–2): 47–66.

- Bradley, B., 2004. "When Is Death Bad for the One Who Dies?" Noûs, 38: 1–28.

- Bradley, B., 2009. Well Being and Death, Oxford: Oxford University Press.

- Bradley, B., Feldman, F., and Johansson, J., eds., 2013. The Oxford Handbook of Philosophy of Death, Oxford: Oxford University Press.

- DeGrazia, D., 2014. "The Nature of Human Death" in S. Luper (ed.), The Cambridge Companion to Life and Death, Cambridge: Cambridge University Press.

- Draper, K., 1999. "Disappointment, Sadness, and Death," Philosophical Review, 108: 387–414.

- Epicurus, 1966a. Principal Doctrines, in J. Saunders (ed.), Greek and Roman Philosophy after Aristotle, New York: Free Press.

12.11 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

3. See Section 12.2
4. See Section 12.3
5. See Section 12.2

UNIT 13: TEMPORALITY

STRUCTURE

13.0 OBJECTIVES

After this unit, we can able to know:

- To know about the Temporal reasoning from antiquity to modern days
- To discuss the Interval temporal logics
- To describe other variants of temporal logics
- To know the Logical deduction and decision methods for temporal logics
- To discuss the Applications of temporal logics

13.1 INTRODUCTION

In philosophy, temporality is traditionally the linear progression of past, present, and future. However, some modern-century philosophers have interpreted temporality in ways other than this linear manner. Examples would be McTaggart's The Unreality of Time, Husserl's analysis of internal time consciousness, Martin Heidegger's Being and Time, George Herbert Mead's Philosophy of the Present, and Jacques Derrida's criticisms of Husserl's analysis, as well as Nietzsche's eternal return of the same, though this latter pertains more to historicity, to which temporality gives rise.

The term Temporal Logic has been broadly used to cover all approaches to reasoning about time and temporal information, as well as their formal representation within a logical framework, and also more narrowly to refer specifically to the modal-logic type of approach introduced around 1960 by Arthur Prior under the name Tense Logic and subsequently developed further by many logicians and computer scientists. Applications of Temporal Logic include its use as a formalism for clarifying philosophical issues about time, as a framework within which to define the semantics of temporal expressions in natural language, as a language for encoding temporal knowledge in artificial intelligence, and as a tool for specification, formal analysis, and verification of the executions of computer programs and systems.

Here we provide a broadly representative, yet concise and inevitably incomplete, overview of the rich variety of temporal models and logics introduced and studied over the past 50 years.

13.2 TEMPORAL REASONING FROM ANTIQUITY TO MODERN DAYS

Discussions of temporality and reasoning about time go back to antiquity, and examples can be found even in the Bible (Boyd 2014). Zeno's famous flying arrow paradox refers to the nature of time and broaches the corresponding notion of change. Much of the early temporal discussion, however, centered around the problem of future contingents, that is, the question whether statements about future events that are neither necessary

nor impossible can have definite truth values. The most widely known and probably most cited example is the sea-fight scenario discussed by Aristotle in On Interpretation (Chapter 9). Aristotle argued that statements such as "There will be a sea-fight tomorrow", as well as the contrary prediction "There will not be a sea-fight tomorrow", do not hold of necessity and hence lack definite truth values at present, while conceding that it is necessary that either there will be a sea-fight tomorrow or not. A few decades later, the philosopher Diodorus Chronus demonstrated the problem of future contingents in his famous Master Argument, which led him to define the possible as "what is or will be the case". A detailed discussion of Diodorus' argument is provided in e.g. Rescher and Urquhart (1971, Chapter XVII) and the entry on future contingents.

Philosophical discussions concerning time and the contingent future continued in the Middle Ages, where the theme was taken up by writers such as Peter Aureole, William of Ockham, and Luis Molina. In the center of focus here was the question how to reconcile God's foreknowledge with the idea of human freedom. Ockham, for example, embraced the idea of a true or actual future, holding that future contingent statements are either true or false even though only God knows their truth values. According to Ockham, this is not to say, however, that future contingents are necessary, meaning that there are alternative possibilities for humans to choose from. Later, several philosophers and logicians engaged in the problem of relating temporality with free will, indeterminism, and the open future, proposing various different solutions. C.S. Peirce objected to the idea that future contingents can have definite truth values. He advanced the view that only the present and the past are actual whereas the future is the realm of possibility and necessity. In a similar spirit, J. Łukasiewicz devised a three-valued logic, treating the truth values of future contingent statements as undetermined. For a more recent philosophical discussion on free will, indeterminism, and the open future, see e.g. Belnap et al. (2001) and Müller (2014).

The modern era of formal temporal logic was initiated by the seminal work of Arthur N. Prior, with important precursors such as H.

Reichenbach, J. Findlay, J. Łukasiewicz, and J. Łoś.[1] From the early 1950s, Prior introduced and analyzed in detail over more than a decade several different versions of Tense Logic, many of which are discussed below. Prior's invention of Tense Logic was largely driven by philosophical considerations. In particular, the Master Argument of Diodorus Chronus and the intricate relationship between time, (in)determinism, God's foreknowledge, and human freedom played a pivotal role in his work. Prior was convinced that a proper logical approach could help to clarify and solve such philosophical problems. He introduced temporal operators, studied metric tense logic, was a pioneer in hybrid temporal logic, devised two versions of branching time temporal logic, which he took to reflect the views of Ockham and Peirce, respectively, etc. His work paved the way for the development of the vast and diverse field of temporal logic, with numerous important applications not only in philosophy, but also in computer science, artificial intelligence, and linguistics. For more on Prior's views and work, see Hasle et al. (2017); Blackburn et al. (2019); and the entry on Arthur Prior.

13.3 INTERVAL TEMPORAL LOGICS

Instant-based and interval-based models of time are two different kinds of temporal ontologies and, even though they are technically reducible to each other, this does not solve the main semantic issue arising when developing a logical formalism to capture temporal reasoning: should propositions about time, and therefore formulae in the given logical language, be interpreted as referring to time instants or to intervals?

There have been various proposals and developments of interval-based temporal logics in the philosophical logic literature. Important early contributions include Hamblin (1972); Humberstone (1979); Röper (1980); and Burgess (1982b). The latter provides an axiomatization for an interval-based temporal logic involving the temporal precedence relation between intervals on the rationals and the reals. The interval-based approach to temporal reasoning has been very prominent in Artificial Intelligence. Some notable works here include Allen's logic of planning (Allen 1984), Kowalski and Sergot's calculus of events (Kowalski and

Sergot 1986), and Halpern and Shoham's modal interval logic (Halpern and Shoham 1986). But, it also features in some applications in computer science, such as *real-time logics* and *hardware verification*, notably Moszkowski's Interval Logic (Moszkowski 1983) and Zhou, Hoare, and Ravn's Duration Calculus (see Hansen and Zhou 1997).

Here we will briefly present the propositional modal interval logic proposed by Halpern and Shoham (1986), hereafter called HS.HS. The language of HSHS includes a family of unary interval operators of the form $\langle X\rangle, \langle X\rangle$, one for each of Allen's interval relations over linear time. The respective notations are listed in Given a set of atomic propositions PROPPROP, formulae are recusively defined by the following grammar:

$$\varphi := p \in PROP | \bot | \neg\varphi | (\varphi \wedge \varphi) | \langle X\rangle\varphi. \varphi := p \in PROP | \bot | \neg\varphi | (\varphi \wedge \varphi) | \langle X\rangle\varphi.$$

The interval logic HSHS starts from instant-based models over linear time, and intervals are considered defined elements. So let $T = \langle T, \prec\rangle T = \langle T, \prec\rangle$ be a temporal frame and assume that the temporal precedence relation $\prec \prec$ induces a strict linear order on the set of time instants TT. An *interval* in TT is defined as an ordered pair $[a,b][a,b]$ such that $a,b \in Ta, b \in T$ and $a \leq b.a \leq b$. The set of all intervals in TT is denoted by $I(T).I(T)$. Note that the definition allows for 'point intervals' whose beginning and end points coincide, following the original proposal by Halpern and Shoham (1986). Sometimes, only 'strict' intervals are considered, excluding point-intervals.

In interval-based temporal logic, formulae are evaluated relative to time intervals rather than instants. An *interval model* is a triple $M = \langle T, \prec, V\rangle M = \langle T, \prec, V\rangle$ consisting of a temporal frame $T = \langle T, \prec\rangle T = \langle T, \prec\rangle$ and a valuation VV that assigns to each atomic proposition $p \in PROPp \in PROP$ the set of time intervals $V(p) \subseteq P(I(T))V(p) \subseteq P(I(T))$ at which pp is considered true. The *truth of an arbitrary formula* $\varphi\varphi$ *with respect to a given*

interval [a,b][a,b] *in an interval model* MM is defined by structural induction on formulae as follows:

- $\mathcal{M}, [a, b] \models p$ iff $[a, b] \in V(p)$, for $p \in PROP$;
- $\mathcal{M}, [a, b] \not\models \bot$;
- $\mathcal{M}, [a, b] \models \neg\varphi$ iff $\mathcal{M}, [a, b] \not\models \varphi$;
- $\mathcal{M}, [a, b] \models \varphi \wedge \psi$ iff $\mathcal{M}, [a, b] \models \varphi$ and $\mathcal{M}, [a, b] \models \psi$;
- $\mathcal{M}, [a, b] \models \langle X \rangle \varphi$ iff $\mathcal{M}, [c, d] \models \varphi$ for some interval $[c, d]$ such that $[a, b] R_X [c, d]$, where R_X is Allen's interval relation corresponding to the modal operator $\langle X \rangle$ (cf. Table 1).

That is, the new interval operators are given a Kripke-style semantics over the associated Allen relations. E.g., for the Allen relation "meets", we have:

M,[t0,t1]⊨⟨A⟩φ iff M,[t1,t2]⊨φ for some interval [t1,t2].M,[t0,t1]⊨⟨A⟩φ iff M,[t1,t2]⊨φ for some interval [t1,t2].

For each diamond modality ⟨X⟩,⟨X⟩, the corresponding box modality is defined as its dual: [X]φ≡¬⟨X⟩¬φ.[X]φ≡¬⟨X⟩¬φ. Sometimes it is useful to include an additional modal constant for point intervals, denoted π,π, with the following truth definition:

M,[a,b]⊨π iff a=b.M,[a,b]⊨π iff a=b.

Some of the HSHS modalities are definable in terms of others, and for each of the strict and non-strict semantics, minimal fragments that are expressive enough to define all other operators have been identified. Complete sets of equivalences that allow one to define certain HSHS modalities in terms of others are presented in the supplementary document:

The logic HSHS has over a thousand expressively non-equivalent fragments involving only some of the modal operators, which have been studied extensively (see Della Monica et al. 2011 for a recent survey). HSHS and most of its fragments are very expressive, and the respective notions of validity are usually undecidable (under some

additional assumptions even highly undecidable, viz. Π_1^1-complete). However, some quite non-trivial decidable fragments of HS have been identified. Probably the best studied one is the *neighborhood interval logic*, which involves the operators $\langle A \rangle$ and $\langle \bar{A} \rangle$ (Goranko et al. 2003). One specific axiom for $\langle A \rangle$ (and symmetrically for $\langle \bar{A} \rangle$)

is $\langle A \rangle \langle A \rangle \langle A \rangle p \rightarrow \langle A \rangle \langle A \rangle p$, saying that any two consecutive right-neighboring intervals can be joined into one right-neighboring interval.

In addition to the unary HS interval modalities associated with Allen's binary interval relations, there is a natural and important operation of chopping an interval into two subintervals, which gives rise to the ternary interval relation 'chop', proposed and studied in Moszkowski (1983). The framework was later extended in Venema (1991) to the logic CDT, which involves next to 'chop' (C), the two residual 'chop' operators D and T. The logic CDT was completely axiomatized in Venema (1991); see also Goranko et al. (2004) and Konur (2013).

There is a natural spatial interpretation of interval temporal logics, based on the idea that the pairs of points that define an interval on a linear order L can be considered coordinates of a point in the $L \times L$-plane. Relations between intervals are then interpreted as spatial relations between the corresponding points. This interpretation has been fruitfully explored to transfer various technical results between spatial and interval logics, such as undecidability, see e.g. Venema (1990) and Marx and Reynolds (1999).

Lastly, a few words about the relationship between interval temporal logics and first-order logic. The standard translation of Prior's basic tense logic TL into first-order logic extends naturally to interval logics, where atomic propositions are represented in the first-order language by binary relations. It turns out that some fragments of HS can be translated into the two-variable fragment FO^2 of first-order logic, which eventually implies their decidability. The expressively strongest such interval logic is

the neighborhood interval logic, which was proven to be expressively complete for FO22 (Bresolin et al. 2009). Other fragments of HSHS require at least three distinct variables for the standard translation. Still, even the full logic HSHS is less expressive than the three-variable fragment FO33 of first-order logic, for which Venema (1991) showed that the logic CDT is expressively complete.

For more on interval temporal logics, see Halpern and Shoham (1986); Venema (1990); Goranko et al. (2003; 2004); Della Monica et al. (2011); the survey Konur (2013), and the references therein.

13.4 OTHER VARIANTS OF TEMPORAL LOGICS

So far we have discussed the traditional family of temporal logics, but there are numerous variations and alternative developments that provide useful formalisms for various applications. We briefly present some of them here: hybrid temporal logics, metric and real-time temporal logics, and quantified propositional logics.

13.4.1 Hybrid temporal logics

A notable family of temporal logics, enriching the traditional framework, are *hybrid temporal logics*, which combine propositional temporal logic with elements of first-order logic and thereby considerably increase the expressive power of the language.

The most prominent notion in hybrid temporal logics is that of a *nominal*. Nominals are special atomic propositions in that they are considered to be true at exactly one instant of the temporal model. Hence, one can think of a nominal aa as saying "It is aa o'clock now". For this reason, nominals are sometimes also called 'clock variables'. The idea of nominals can be traced back to Prior (1967, Chapter V; 1968, Chapter XI), who considered the possibility of identifying instants with instant-propositions: an instant can be conceived of as the conjunction of all those propositions that are true at that instant. A first systematic treatment of hybrid temporal logic

was given in Bull (1970). In addition to nominals, hybrid languages are often augmented by further syntactic mechanisms, such as a satisfaction operator, nominal quantifiers, and reference pointers, which we briefly discuss below. The former two mechanisms can already be found in Prior's work (see Blackburn 2006) and were reinvented independently in Passy and Tinchev (1985). Reference pointers were introduced only much later in Goranko (1996), and a similar referencing mechanism is found in Alur and Henzinger (1994).

- *Satisfaction operator*: The satisfaction operator @i@i allows one to express that a given formula is true in a model at the time instant denoted by the nominal ii. That is, $M, t \vDash @i\varphi M, t \vDash @i\varphi$ iff $M, V(i) \vDash \varphi, M, V(i) \vDash \varphi$, where $V(i)V(i)$ is the unique instant where ii is true. The notion of truth at an instant of a temporal model is imported into the object-language.

- *Quantifiers over nominals*: By means of the nominal quantifier ∀i∀i one can express that a given formula is true at a given time instant in a temporal model under each possible assignment of time instants to ii. More formally, $M, t \vDash \forall i\varphi M, t \vDash \forall i\varphi$ iff $M[i \to s], t \vDash \varphi M[i \to s], t \vDash \varphi$ for any instant ss in M, M, where $M[i \to s]M[i \to s]$ is the model obtained from MM by re-assigning the denotation of ii to be s.s. The full power of first-order quantification is brought into the temporal language, while many of its propositional virtues are preserved.

- *Reference pointers*: Reference pointers ↓i↓i are often also referred to as 'binders', for they bind the value of the nominal ii to the current instant of evaluation. A formula ↓iφ↓iφ is true at a given instant tt in a temporal model iff φφ is true at tt whenever the nominal ii denotes tt. That is, $M, t \vDash \downarrow i\varphi M, t \vDash \downarrow i\varphi$ iff $M[i \to t], t \vDash \varphi M[i \to t], t \vDash \varphi$. Reference pointers provide a mechanism for referring to the current time instant, i.e. saying 'now'. For a systematic logical treatment of 'now', see Kamp (1971).

Other operators that can be considered hybrid temporal logic operators are the *universal modality*, the *difference modality*, and *propositional quantifiers*. The universal modality AA says that a given formula is true at

every instant of the temporal modal and hence captures the global notion of truth in a model: M,t⊨AφM,t⊨Aφ iff M⊨φM⊨φ. The difference modality DD, on the other hand, states that the given formula is true at some other instant. Note that both modalities abstract away from the underlying accessibility relation. Propositional quantifiers ∀p∀p introduce second-order quantification into the propositional language

Hybrid languages are very expressive. Here are just two examples:

- Irreflexivity of the precedence relation, which is not expressible in TL, can be expressed in a language with nominals and the satisfaction operator @i@i as @iG¬i.@iG¬i.
- The operators SS and UU are definable in a language with nominals and binders: φUψ:=↓iF(ψ∧H(Pi→φ))φUψ:=↓iF(ψ∧H(Pi→φ)) and likewise for S.S.

While the weaker versions of hybrid logics — with nominals, satisfaction operators, universal modality, and difference modality — are still decidable, the more expressive ones — with quantifiers over nominals or reference pointers — are usually undecidable. For details, see Goranko (1996); and Areces and ten Cate (2006).

Branching time versions of hybrid temporal logics have been investigated as well. For an overview of varieties of hybrid temporal logics and their historical development, see Blackburn and Tzakova (1999) and the entry on hybrid logic.

13.4.2 Metric and real-time temporal logics

Metric temporal logics go back to Prior, too (see Prior 1967, Chapter VI). He used the notation PnφPnφ for "It was the case the interval nn ago that φφ" (i.e. φφ was the case nn time units ago) and FnφFnφ for "It will be the case the interval nn hence that φφ" (i.e φφ will be the case nn time units hence). These operators presuppose that time has a certain metric structure and can be carved up into temporal units, which may be

associated with clock times (e.g. hours, days, years, etc.). If the relevant units are days, for example, the operator F1F1 reads 'tomorrow'.

Prior noted that PnφPnφ can be defined as F(−n)φ.F(−n)φ. The case n=0n=0 accordingly amounts to the present tense. The metric operators validate combination principles such as:

FnφFmφ→F(n+m)φ.FnφFmφ→F(n+m)φ.

The interrelation of the metric and non-metric versions of the temporal operators is captured by the following equivalences:

Pφ≡∃n(n<0∧Fnφ)Hφ≡∀n(n<0→Fnφ)Fφ≡∃n(n>0∧Fnφ)Gφ≡∀n(n>0→Fnφ).Pφ≡∃n(n<0∧Fnφ)Fφ≡∃n(n>0∧Fnφ)Hφ≡∀n(n<0→Fnφ)Gφ≡∀n(n>0→Fnφ).

Instant-based temporal logics for metric time are studied in e.g. Rescher and Urquhart (1971, Chapter X); Montanari (1996); and Montanari and Policriti (1996). For metric interval logics, see Bresolin et al. (2013).

Various metric extensions of temporal logics over the structure of the real numbers have been proposed as well, giving rise to so-called *real-time* logics. These logics introduce additional operators, such as the following, which allow for different formalizations of the example sentence "whenever pp holds in the future, qq will hold within three time units later":

- time-bounded operators, e.g.: G(p→F≤3q);G(p→F≤3q);
- freeze quantifiers (similar to hybrid logic reference pointers), e.g.: Gx.(p→Fy.(q∧y≤x+3));Gx.(p→Fy.(q∧y≤x+3));
- quantifiers over time variables, e.g.: ∀xG(p∧t=x→F(q∧t≤x+3)).∀xG(p∧t=x→F(q∧t≤x+3)).

Such real-time extensions are usually very expressive and often lead to logics with undecidable decision problems. A way to regain decidability

is to relax "punctuality" requirements involving precise time durations, by requirements involving time intervals. For details, see e.g. Koymans (1990); Alur and Henzinger (1992; 1993; 1994) as well as Reynolds (2010; 2014) on the real-time linear temporal logic RTL, and the survey Konur (2013).

13.4.3 Quantified propositional temporal logics

Propositional temporal logics can be extended with quantifiers over atomic propositions (see Rescher and Urquhart 1971, Chapter XIX). Semantically, these quantify over all valuations of the respective atomic propositions and hence are tantamount to monadic second-order quantifiers. The resulting languages are very expressive, and the respective logics are usually undecidable (often not even recursively axiomatizable). Notable extensions include the logic QPTL, the quantified propositional version of LTL (which is decidable albeit with non-elementary complexity), as well as the extension of CTL* (see French 2001). Complete axiomatic systems and decidability results for the quantified propositional temporal logic QPTL (with and without past operators) have been presented in Kesten and Pnueli (2002) and French and Reynolds (2003).

Check Your Progress 1

Note: a) Use the space provided for your answer

b) Check your answers with those provided at the end of the unit

1. What do you know about the Temporal reasoning from antiquity to modern days?

   _____

   _____

   _____

2. Discuss the Interval temporal logics.

   _____

   _____

   _____

3.   Describe other variants of temporal logics.

_____

_____

_____

13.5 LOGICAL DEDUCTION AND DECISION METHODS FOR TEMPORAL LOGICS

Extensive research and numerous publications over the past 50 years have developed a variety of logical deduction systems and decision methods for the temporal logics mentioned here and many more. Hilbert style axiomatic systems are the most common logical deduction systems for temporal logics, but many complete systems of semantic tableaux, sequent calculi, and resolution-based systems have been proposed as well. Some general references on deductive systems for temporal logics (in addition to the more specific references mentioned elsewhere in this text) include: Rescher and Urquhart (1971); McArthur (1976); Burgess (1984); Emerson (1990); Goldblatt (1992); Gabbay et al. (1994); van Benthem (1995); Bolc and Szalas (1995); Gabbay and Guenthner (2002); Gabbay et al. (2003); Fisher et al. (2005); Blackburn et al. (2006); Baier and Katoen (2008); Kröger and Merz (2008); Fisher (2011); Demri et al. (2016).

One of the most important logical decision problems is to determine whether a given formula of a given logic is valid (resp. satisfiable) in the semantics provided for that logic. Particularly efficient and practically useful for deciding satisfiability are the tableaux-based methods, originating from pioneering work of Beth, Hintikka, Smullyan, and Fitting. These methods are based on a systematic search of a satisfying model (resp. falsifying countermodel) if an input formula that is tested for satisfiability is provided, and they are guaranteed to find such a model whenever it exists. Tableaux-based methods have been successfully developed for constructive satisfiability testing for a variety of temporal logics. See Goré (1999) for a survey on tableaux systems for many

temporal logics and more specifically: Ben-Ari et al. (1983) for the branching time logic UB; Emerson and Halpern (1985) for the computation tree logic CTL; Wolper (1985) for the linear time temporal logic LTL; Kontchakov et al. (2004) on temporalizing tableaux; Reynolds (2007) for CTL with bundled tree semantics; Goranko and Shkatov (2010) for ATL; Reynolds (2011) for the full computation tree logic CTL*; Reynolds (2014) for the real-time temporal logic RTL, etc.

Other methods that have proven practically fruitful for deciding satisfiability as well as for model checking of temporal logics in computer science are the automata-based methods, which have been actively developing since the early 1990s. These methods transform temporal formulae into automata on infinite words (for linear time logics) or infinite trees (for branching time logics) and represent models for the logics as input objects (infinite words or trees) for their associated automata. Thus, satisfiability of a formula becomes equivalent to the language of the associated automaton being non-empty. The methods are based on classical results about decidability of the monadic second-order theories of the natural numbers (by Büchi) and of the infinite binary tree (by Rabin). For instance, in Emerson and Sistla (1984), automata on infinite trees and Rabin's theorem were used to obtain a decision procedure for CTL*. For further details see Vardi (2006).

Important references on decidability results and decision procedures for various temporal logics include: Burgess (1980) and Gurevich and Shelah (1985) for branching time logics; Burgess and Gurevich (1985) for linear temporal logics; Goldblatt (1992) for both linear and branching time logics; Montanari and Policriti (1996) for metric and layered temporal logics; French (2001) for some quantified propositional branching time logics.

While most propositional temporal logics are decidable, adding some syntactic or semantic features can make them explode computationally and become undecidable. The most common causes of undecidability of temporal logics, besides combinations with other expressive logics,

include: grid-like models; temporal operators along multiple time-lines; products of temporal logics; interval-based logics with no locality assumptions; time reference mechanisms, such as hybrid reference pointers and freeze quantifiers; arithmetic features, such as time addition, exact time constraints, etc. However, there are various ways to tame temporal logics and restore decidability, such as adding syntactic and parametric restrictions (e.g. on the number of propositional variables or the depth of nesting), imposing suitable semantic restrictions (e.g. locality for interval logics), identifying decidable fragments (e.g. the two-variable fragment FO2 of classical first-order logic, guarded fragments, monodic fragments), etc.

13.6 APPLICATIONS OF TEMPORAL LOGICS

Temporal logic is a field whose development has been heavily driven by philosophical considerations. At the same time, the logical formalisms and technical systems developed over the years have found application in various different disciplines, ranging from computer science, artificial intelligence, and linguistics, to natural, cognitive, and social sciences. In this section, we briefly discuss some pertinent applications of temporal logics in computer science, artificial intelligence, and linguistics.

13.6.1 Temporal logics in Computer Science

The idea to apply temporal reasoning to the analysis of deterministic and stochastic transition systems was already present in the theory of processes and events in Rescher and Urquhart (1971, Chapter XIV). However, it was with the seminal paper of Pnueli (1977) that temporal logic became really prominent in computer science. Pnueli proposed the application of temporal logics to the specification and verification of reactive and concurrent programs and systems. In order to ensure correct behavior of a reactive program, in which computations are non-terminating (e.g. an operating system), it is necessary to formally specify and verify the acceptable infinite executions of that program. In addition, to ensure correctness of a concurrent program, where two or more

processors are working in parallel, it is necessary to formally specify and verify their interaction and synchronization.

Key properties of infinite computations that can be captured by temporal patterns are liveness, safety, and fairness (see Manna and Pnueli 1992):

Liveness properties or eventualities involve temporal patterns of the forms Fp, q→Fp, or G(q→Fp), which ensure that if a specific precondition (q) is initially satisfied, then a desirable state (satisfying p) will eventually be reached in the course of the computation. Examples are "If a message is sent, it will eventually be delivered" and "Whenever a printing job is activated, it will eventually be completed".

Safety or invariance properties involve temporal patterns of the forms Gp, q→Gp, or G(q→Gp), which ensure that if a specific precondition (q) is initially satisfied, then undesirable states (violating the safety condition p) will never occur. Examples are: "No more than one process will be in its critical section at any moment of time", "A resource will never be used by two or more processes simultaneously", or, to give more practical examples: "The traffic lights will never show green in both directions", "A train will never pass a red semaphore".

Fairness properties involve combinations of temporal patterns of the forms GFp ("infinitely often p") or FGp ("eventually always p"). Intuitively, fairness requires that whenever several processes that share resources are run concurrently, they must be treated 'fairly' by the operating system, scheduler, etc. A typical fairness requirement says that if a process is persistent enough in sending a request (e.g. keeps sending it over and over again), its request will eventually be granted.

An infinite computation is formally represented by a model of the linear time temporal logic LTL. Non-deterministic systems are modeled by branching time structures. Thus, both LTL and the computation tree logics CTL and CTL* have been very important for specification and verification of reactive and concurrent systems.

The following example combines liveness and safety properties of a single computation: "Whenever a state of alert is reached, the alarm is activated and remains activated until a safe state is eventually reached". This property is expressible in LTL as

$$G(\text{alert} \rightarrow (\text{alarm } U \text{ safe})).$$

Another example, referring to all computations in the system, is: "If the process σ is eventually enabled on some computation starting from the current state, then on every computation starting there, whenever σ is enabled, it will remain enabled until the process τ is disabled". This property can be formalized in CTL* as

$$\Diamond F \text{ enabled}_\sigma \rightarrow \Box G(\text{enabled}_\sigma \rightarrow (\text{enabled}_\sigma U \text{ disabled}_\tau)).$$

A variation of LTL with useful applications for specifying and reasoning about concurrent systems is Lamport's (1994) temporal logic of actions TLA. Other applications of temporal logics in computer science include: temporal databases, real-time processes and systems, hardware verification, etc. Further information on such applications can be found in e.g. Pnueli (1977); Emerson and Clarke (1982); Moszkowski (1983); Galton (1987); Emerson (1990); Alur and Henzinger (1992); Lamport (1994); Vardi and Wolper (1994); Bolc and Szalas (1995); Gabbay et al. (2000); Baier and Katoen (2008); Kröger and Merz (2008); Fisher (2011); Demri et al. (2016).

13.6.2 Temporal logics in Artificial Intelligence

Artificial Intelligence (AI) is one of the major areas of application of temporal logics. Relating temporal reasoning to AI was suggested already in the early philosophical discussion on AI by McCarthy and Hayes (1969), the theory of processes and events in Rescher and Urquhart (1971, Chapter XIV), and the period-based theories of Hamblin (1972); see Øhrstrøm and Hasle (1995) for an overview of these early developments. In the 1980s, temporal representation and reasoning gradually became an increasingly prominent theme in AI with several influential works,

including McDermott's temporal logic for reasoning about processes and plans (McDermott 1982); Allen's general theory of action and time (Allen 1984); the Event Calculus of Kowalski and Sergot (1986); the reified temporal logic by Shoham (1987); the logic of time representation by Ladkin (1987); and the work on temporal database management by Dean and McDermott (1987). Galton (1987) provides a systematic account of these and other important developments in that period; see also Vila (1994) and Pani and Bhattacharjee (2001) for comprehensive reviews. Influential works in the 1990s include the introduction of interval-based temporal logics by Halpern and Shoham (1991) and by Allen and Ferguson (1994), with representation of actions and events; the Situation Calculus of Pinto and Reiter (1995); and Lamport's Action Theory (Lamport 1994), etc. Further important developments relating temporal reasoning and AI since then include: temporal reasoning in natural language, temporal ontologies, temporal databases and constraint solving, temporal planning, executable temporal logics, spatial-temporal reasoning, temporal reasoning in agent-based systems, etc. The field has gradually grown so rich and broad — as witnessed by the 20-chapter handbook edited by Fisher et al. (2005) — that it is impossible to even briefly survey it here, so we only bookmark a few of the main issues of philosophical relevance that have been in the focus of temporal reasoning and logics in AI.

The prevailing logical approach to temporal representation and reasoning in AI, especially in the 1980s-1990s, has traditionally been based on temporalized variations of first-order logic rather than on Prior-style temporal logic. The approach is best illustrated by the so-called method of temporal arguments (McCarthy and Hayes 1969; Shoham 1987; Vila 1994). According to this method, the temporal dimension is captured by augmenting propositions and predicates with 'time stamp' arguments, as for example "Publish(A. Prior, Time and Modality, 1957)". An alternative, yet closely related approach, is the one of reified temporal logics (McDermott 1982; Allen 1984; Shoham 1987; see Ma and Knight 2001 for a survey). This approach makes use of reifying meta-predicates, such as 'TRUE' and 'FALSE', but also 'HOLDS', 'OCCURS',

'BEFORE', 'AFTER', interval relations such as 'MEETS', 'OVERLAPS', etc., which are applied to propositions of some standard logical language (e.g. classical first-order logic). An example of a reified expression is "OCCUR(Born(A. Prior), 1914)". Associated with theories of time are theories of temporal incidence (cf. Vila 2005). Still, the prominence of the modal logic based approach has always been strong, and it has more recently resurged, e.g. in the context of agent-based temporal reasoning (cf. Fisher and Wooldridge 2005).

Theories of temporal reasoning in AI distinguish between fluents, which are propositions describing states of the world that may change over time, and events, representing what happens in the world and causes changes between states. Philosophical issues arising here concern the nature of fluents and events, the meaning of instantaneous events, the distinction between homogeneous states and inhomogeneous events, the dividing instant problem, the frame problem, etc. For further discussion, see Shoham (1987); Galton (1990); and Vila (2005). See also the related discussion on reasoning about action and change in Section 4 of the entry on logic and artificial intelligence.

Both fluents and events can be considered in discrete or continuous time and they can be instantaneous or durational. This keeps the debate on instant-based versus interval-based formal models of time alive, with various theories following and comparing both approaches, including van Benthem (1983); Allen (1983); Allen and Hayes (1989); Allen and Ferguson (1994); Galton (1995); Vila (2005); etc.

For further reading and discussion on temporal reasoning and logics in AI, see Vila (1994); Galton (1995); the comprehensive handbook Fisher et al. (2005); and the more concise handbook chapter Fisher (2008).

13.6.3 Temporal logics in Linguistics

Tense is an important feature of natural languages. It is a linguistic device that allows one to specify the relative location of events in time, usually with respect to the speech time. In most languages, including English, tense becomes manifest in a system of different verbal tenses. English allows for a distinction between past, present, and future tense ('will'

future) and, traditionally, the respective perfect and progressive forms are referred to as tenses as well.

As laid out above, Prior's invention of tense logic was largely motivated by the use of tense in natural language. An alternative early logical approach to tense was provided by Reichenbach (1947), who suggested an analysis of the English verbal tenses in terms of three points in time: speech time, event time, and reference time, where the reference time is a contextually salient point in time, which, intuitively, captures the perspective from which the event is viewed. Using the notion of reference time, Reichenbach was able to distinguish, for example, between the simple past ("I wrote a letter") and the present perfect ("I have written a letter"), which are conflated in Prior's account. In both cases, of simple past and of present perfect, the event time precedes the speech time; but, in the former case the reference time coincides with the event time whereas in the latter case the reference time is simultaneous with the speech time.

Neither Prior's nor Reichenbach's frameworks can account for the difference between, for instance, the simple past ("I wrote a letter") and the past progressive ("I was writing a letter"). The relevant distinction here is one of aspect rather than of tense and naturally requires an interval-based or event-based setting to be adequately dealt with. For accounts along these lines, see e.g. Dowty (1979); Parsons (1980); Galton (1984); and van Lambalgen and Hamm (2005).

Whereas Reichenbach's analysis makes reference to a contextually salient point in time, on Prior's account tenses are construed as temporal operators, which are interpreted as quantifiers over instants in time. This raises the general question: are tenses in natural language to be treated as quantifiers, or do they refer to specific points in time? In an influential paper, Partee (1973) provided the following counterexample against a quantifier treatment of tenses: the sentence "I didn't turn off the stove" means neither (1) there is an earlier time instant at which I do not turn off the stove, nor does it mean (2) there is no earlier instant at which I turn off

the stove. The first requirement is too weak, the second too strong. Partee suggested an analogy between tenses and referential pronouns. According to this proposal, tenses refer to specific, contextually given points in time (e.g. 8 o'clock this morning), which are presupposed to stand in appropriate temporal relations to the speech time. Subsequently, accounts that restrict quantification to a contextually given time interval (e.g. this morning) have become popular. On these accounts, Partee's example sentence has the intuitive meaning: there is no earlier time instant in the contextually salient time interval at which I turn off the stove. Formally, this idea is compatible with both a quantifier and a referential treatment of tenses; for details see Kuhn and Portner (2002) and Ogihara (2011).

Check Your Progress 2

Note: a) Use the space provided for your answer

b) Check your answers with those provided at the end of the unit

1. What do you know the Logical deduction and decision methods for temporal logics?

_____

_____

_____

2. Discuss the Applications of temporal logics.

_____

_____

_____

13.7 LET US SUM UP

In philosophy, temporality is traditionally the linear progression of past, present, and future. However, some modern-century philosophers have interpreted temporality in ways other than this linear manner. Examples would be McTaggart's The Unreality of Time, Husserl's analysis of internal time consciousness, Martin Heidegger's Being and Time (1927), George Herbert Mead's Philosophy of the Present (1932), and Jacques Derrida's criticisms of Husserl's analysis, as well as Nietzsche's eternal

return of the same, though this latter pertains more to historicity, to which temporality gives rise.

In social sciences, temporality is also studied with respect to human's perception of time and the social organization of time. The perception of time undergoes significant change in the three hundred years between the Middle Ages and Modernity.

13.8 KEY WORDS

Temporality: In philosophy, temporality is traditionally the linear progression of past, present, and future.

13.9 QUESTIONS FOR REVIEW

1.   Discuss about Temporal logics in Computer Science.
2.   Discuss about Temporal logics in Artificial Intelligence.
3.   Discuss about Temporal logics in Linguistics.

13.10 SUGGESTED READINGS AND REFERENCES

- Allen, J.F., 1983, "Maintaining Knowledge about Temporal Intervals", Communications of the ACM, 26(11): 832–843.
- ——, 1984, "Towards a General Theory of Action and Time", Artificial Intelligence, 23: 123–154.
- Allen, J.F., and G. Ferguson, 1994, "Actions and Events in Interval Temporal Logic", Journal of Logic and Computation, 4(5): 531–579.
- Allen, J.F., and P. Hayes, 1989, "Moments and Points in an Interval-Based Temporal Logic", Computational Intelligence, 5(4): 225–238.
- Alur, R., and T. Henzinger, 1992, "Logics and Models of Real-Time: A Survey", in Real-Time: Theory in Practice, Proceedings of the REX Workshop 1991 (Lecture Notes in Computer Science: Volume 600), Berlin: Springer, pp. 74–106.

- ——, 1993, "Real-Time Logics: Complexity and Expressiveness", Information and Computation, 104: 35–77.

- ——, 1994, "A Really Temporal Logic", Journal of the ACM, 41: 181–204.

- Alur, R., T. Henzinger, and O. Kupferman, 2002, "Alternating-Time Temporal Logic", Journal of the ACM, 49(5): 672–713.

- Andréka, H., V. Goranko, S. Mikulas, I. Németi, and I. Sain, 1995, "Effective First-Order Temporal Logics of Programs", in Bolc and Szalas (1995), pp. 51–129.

- Andréka, H., J. Madarász, and I. Németi, 2007, "Logic of Space-Time and Relativity Theory", in M. Aiello, J. van Benthem, and I. Pratt-Hartmann (eds.), Handbook of Spatial Logics, Dordrecht: Springer, pp. 607–711.

- Areces, C., and B. ten Cate, 2006, "Hybrid Logics", in Blackburn et al. (2006), pp. 821–868.

- Aristotle, Organon, II - On Interpretation, Chapter 9. See https://archive.org/stream/AristotleOrganon/AristotleOrganoncolle ctedWorks.

13.11 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

1. See Section 13.2
2. See Section 13.3
3. See Section 13.4

Check Your Progress 2

1. See Section 13.2
2. See Section 13.3

UNIT 14: AUTHENTIC AND NON-AUTHENTIC

STRUCTURE

14.0 OBJECTIVES

After this unit, we can able to know:

- To know the Origins and Meaning of the Concept of Authenticity
- To criticise of Authenticity
- To find out the Conceptions of Authenticity
- To understand the Recent Accounts of Authenticity

14.1 INTRODUCTION

The term 'authentic' is used either in the strong sense of being "of undisputed origin or authorship", or in a weaker sense of being "faithful to an original" or a "reliable, accurate representation". To say that something is authentic is to say that it is what it professes to be, or what it is reputed to be, in origin or authorship. But the distinction between authentic and derivative is more complicated when discussing authenticity as a characteristic attributed to human beings. For in this case, the question arises: What is it to be oneself, at one with oneself, or truly representing one's self? The multiplicity of puzzles that arise in conjunction with the conception of authenticity connects with metaphysical, epistemological, and moral issues (for recent discussion, see Newman and Smith 2016; Heldke and Thomsen 2014). On the one hand, being oneself is inescapable, since whenever one makes a choice or acts, it is oneself who is doing these things. But on the other hand, we are sometimes inclined to say that some of the thoughts, decisions and actions that we undertake are not really one's own and are therefore not genuinely expressive of who one is. Here, the issue is no longer of metaphysical nature, but rather about moral-psychology, identity and responsibility.

When used in this latter sense, the characterization describes a person who acts in accordance with desires, motives, ideals or beliefs that are not only hers (as opposed to someone else's), but that also express who she really is. Bernard Williams captures this when he specifies authenticity as "the idea that some things are in some sense really you, or express what you are, and others aren't" (quoted in Guignon 2004: viii).

Besides being a topic in philosophical debates, authenticity is also a pervasive ideal that impacts social and political thinking. In fact, one distinctive feature of recent Western intellectual developments has been a shift to what is called the "age of authenticity" (Taylor 2007; Ferrarra 1998). Therefore, understanding the concept also involves investigating its historical and philosophical sources and on the way it impacts the socio-political outlook of contemporary societies.

14.2 ORIGINS AND MEANING OF THE CONCEPT OF AUTHENTICITY

14.2.1 Sincerity and Authenticity

A number of significant cultural changes in the seventeenth and eighteenth centuries led to the emergence of a new ideal in the Western world (Trilling 1972). During this period, human beings came to be thought of more as individuals than as placeholders in systems of social relations. This emphasis on the importance of the individual is seen in the prevalence of autobiographies and self-portraits, where the individual becomes the centre of attention not because of extraordinary feats or access to special knowledge, but because he or she is an individual.

In the same period, society comes to be seen not as an organic whole of interacting components, but as an aggregate of individual human beings, a social system with a life of its own, which presents itself to the individual as not itself quite human but rather as artificial, the result of a "social contract". Being human is understood as being best achieved through being unique and distinctive, even when these collide with certain social norms. At the same time, there is an increasing awareness of what Charles Taylor (1989) calls "inwardness" or "internal space". The result is a distinction between one's private and unique individuality, and one's public self (Taylor 1991; Trilling 1972).

With these social changes there is a sharp shift in the conceptions of approbation and disapproval that are commonly used in judging others and oneself. For instance, concepts like sincerity and honor become obsolescent (Berger 1970). In earlier times, a sincere person was seen as someone who honestly attempts to neither violate the expectations that follow from the position he holds in society, nor to strive to appear otherwise than he ought to. However, by the time of Hegel, the ideal of sincerity had lost its normative appeal. Hegel polemically refers to sincerity as "the heroism of dumb service" (Hegel 2002 [1807]: 515) and launches an attack on the bourgeois "honest man," who passively internalizes a particular conventional social ethos. In the condition of

sincerity, the individual is uncritically obedient to the power of society—a conformity that for Hegel leads to subjugation and a deterioration of the individual (Hegel 2002 [1807]; Golomb 1995: 9; Trilling 1972). For Hegel, in the progress of "spirit", the individual consciousness will eventually move from this condition of sincerity to a condition of baseness, in which the individual becomes antagonistic to external societal powers and achieves a measure of autonomy. Hegel shows this clearly in a comment on Diderot's Rameau's Nephew, a story in which the narrator (supposedly Diderot himself) is portrayed as the reasonable, sincere man who respects the prevailing order and who has achieved bourgeois respectability. In contrast, the nephew is full of contempt for the society in which he figures as a worthless person. However, he is in opposition to himself, because he still aspires to a better standing in a society, which he believes has nothing but emptiness to offer (Despland 1975: 360; Golomb 1995: 13–15). For Hegel, the narrator is an example of the sincere, honest soul, while the nephew figures as the "disintegrated," alienated consciousness. The nephew is clearly alienated, but for Hegel this alienation is a step in the progression towards autonomous existence (Williams 2002: 190).

In the midst of this conceptual change, the term 'authenticity' becomes applicable in demarcating a somewhat new set of virtues. The older concept of sincerity, referring to being truthful in order to be honest in one's dealings with others, comes to be replaced by a relatively new concept of authenticity, understood as being true to oneself for one's own benefit. Earlier, the moral advice to be authentic recommended that one should be true to oneself in order thereby to be true to others. Thus, being true to oneself is seen as a means to the end of successful social relations. In contrast, in our contemporary thinking, authenticity as a virtue term is seen as referring to a way of acting that is choiceworthy in itself (Ferrara 1993; Varga 2011a; Varga 2011b).

14.2.2 Autonomy and Authenticity

The growing appeal of the idea of authenticity has led to the emergence of a highly influential modern "ethic of authenticity" (Ferrara 1993; Ferrara

Notes

2017). This ethic acknowledges the value of the dominant "ethic of autonomy" that shapes modern moral thought (Schneewind 1998; Dworkin 1988). The idea of autonomy emphasizes the individual's self-governing abilities, the independence of one's deliberation from manipulation and the capacity to decide for oneself. It is connected to the view that moral principles and the legitimacy of political authority should be grounded in the self-governing individual who is free from diverse cultural and social pressures. According to the ethic of autonomy, each individual should follow those norms he or she can will on the basis of rational reflective endorsement. To some extent, authenticity and autonomy agree in supposing that one should strive to lead one's life according to one's own reasons and motives, relying on one's capacity to follow self-imposed guidelines. In both cases, it is crucial that one has the ability to put one's own behavior under reflexive scrutiny and make it dependent on self-determined goals (Honneth 1994).

One crucial difference is that the ethic of authenticity introduces the idea that there are motives, desires and commitments that sometimes should outweigh the restrictions of rational reflection. This is because those motives are so fundamental to the cohesion of one's own identity that overriding them would mean disintegrating the very self which is necessary to be a moral agent. The point is that there are types of moral philosophical reasoning that can be repressive if they arise from "an autonomous moral conscience not complemented by sensitivity to the equilibrium of identity and by authenticity" (Ferrara 1993: 102). Besides leading an autonomous life, guided by one's own, non-constrained reasons and motives, authenticity requires that these motives and reasons should be expressive of one's self-identity. Authenticity guides the moral agent to follow only those "moral sources outside the subject [that speak in a language] which resonate[s] within him or her", in other words, moral sources that accord with "an order which is inseparably indexed to a personal vision" (Taylor 1989: 510). Hence, authenticity entails an aspect that lies beyond the scope of autonomy, namely, a "language of personal resonance" (Taylor 1991: 90). This points to the gap between (Kantian)

autonomy and authenticity: one can lead an autonomous life, even if this way of living fails to express a person's self-understanding.

In recent years, more attention has been devoted to highlighting how autonomy and authenticity can come apart (e.g. Oshana 2007; Roessler 2012; MacKay forthcoming). Some argue that authenticity demands more than is necessary for autonomy: a person does not have to reflectively endorse key aspects of her identity in order to qualify as autonomous (Oshana 2007). If she acknowledges that aspects of her identity contradict her self-conception, she might still be autonomous, even if this acknowledgement injects ambivalence into her life.

In all, the ideal of authenticity does not object to the importance of the self-given law, but disagrees that full freedom consists in making and following such a law (Menke 2005: 308). It is not just about being involved in the authorship of such a law, but about how this law fits with the wholeness of a person's life, and how or whether it expresses who the person is. In this sense, the idea of autonomy already represents a counterposition to an ethic that is solely concerned with strict adherence to social norms.

14.2.3 Authenticity and the self

Another decisive factor in the development of the ideal of authenticity was that it emerged together with a distinctively modern conception of the self. This is visible in the work of Rousseau, who argues that the orientation toward life that should guide the conduct one chooses should come from a source within. This led to questions about inwardness, self-reflection and introspection, many of them addressed in his Confessions (1770). When the space of interiority becomes a guiding authority, the individual must detect and distinguish central impulses, feelings and wishes from ones that are less central or conflict with one's central motives. In other words, interiority must be divided into what is at the core and what is peripheral. In this picture, the measure of one's actions is whether they spring from and express essential aspects of one's identity or whether they come from a peripheral place.

Such a conception of the self exhibits decisive parallels to the tradition of "religious individualism" that centers religious life on the individual and stresses the importance of inwardness and the introspective examination of one's inner motives, intentions and conscience. Investigating the characteristics of the modern subject of inwardness, Foucault (1980: 58–60) suggests that "it seems to us that truth, lodged in our most secret nature, 'demands' only to surface." For Foucault, confession—the look inward to monitor one's interior life and to tell certain "truths" about oneself—has become a part of a cultural life, reaching from religious contexts to psychological therapy. The radicalization of the distinction between true and false interiority has led to new possibilities; inner states, motivations and feelings are now increasingly thought of as objectifiable and malleable in different contexts.

Rousseau also adds that acting on motives that spring from the periphery of the self, while ignoring or denying essential aspects of one's self, simply amounts to self-betrayal and annihilation of the self. Rousseau's The New Heloise (1997 [1761]) emphasizes this aspect by showing how the novel accentuates the significant costs and the potential self-alienation involved in suppressing one's deepest motivations. But, in addition, in the Discourse on the Origin of Inequality, Rousseau argues that, with the emergence of a competitive public sphere, the ability to turn inward is increasingly compromised, because competitive relations require intense role-playing, which Rousseau calls an "excessive labor" (Rousseau 1992 [1754]: 22). The ongoing instrumental role-playing not only causes alienation, but ultimately inequality and injustice, since it destroys the immanent moral understanding with which, according to Rousseau, humans are hard-wired. Social life requires identification with social roles, but because role identity is determined by other people's normative expectations, role-playing leads to a tension that might be understood as a matter of politics more than anything else (Schmid 2017).

14.3 CRITIQUE OF AUTHENTICITY

The idea of autonomy—the view that each individual must decide how to act based on his or her own rational deliberations about the best course of action—has in many ways paved the way for the idea of authenticity. However, authenticity goes beyond autonomy by holding that an individual's feelings and deepest desires can outweigh both the outcome of rational deliberation in making decisions, and our willingness to immerse ourselves into the reigning norms and values of society. Whereas sincerity generally seems to accept a given social order, authenticity becomes an implicitly critical concept, often calling into question the reigning social order and public opinion. In Rousseau's optic, one of our most important projects is to avert from the social sphere and to unearth what is truly us underneath the 'masks' that society forces on us. But when authenticity comes to be regarded as something like sincerity for its own sake (Ferrara 1993: 86), it becomes increasingly hard to see what the moral good is that it is supposed to bring into being.

A frequently mentioned worry with the ideal of authenticity is that the focus on one's own inner feelings and attitudes may breed a self-centered preoccupation with oneself that is anti-social and destructive of altruism and compassion toward others. Christopher Lasch (1979) points out similarities between the clinical disorder referred to as Narcissistic Personality Disorder and authenticity. According to Lasch, narcissism and authenticity are both characterized by deficient empathic skills, self-indulgence and self-absorbed behavior. Similarly, Allan Bloom (1987: 61) maintains that the culture of authenticity has made the minds of the youth "narrower and flatter," leading to self-centeredness and the collapse of the public self. While Lasch and Bloom worry about the threat that the self-centeredness and narcissism of the "culture of authenticity" poses to morality and political coherence, Daniel Bell voices worries about its economic viability. What Bell fears is that the "megalomania of self-infinitization" that comes with the culture of authenticity will erode the foundations of market mechanisms that are "based on a moral system of reward rooted in the Protestant sanctification of work" (Bell 1976: 84). More recently, critics have argued that when properly analyzed, authenticity demands positing the existence of a "true self." It requires

positing an essentialist structure leading to metaphysical problems that current accounts of authenticity fail to solve (Bialystok 2014). Correspondingly, Feldman (2014) argues in favor of abandoning the ideal of authenticity because it builds on confused assumptions about the self, the value of one's "gut feelings" in revealing one's values, and the supposedly corrupting influence of the "external" social realm (for a critique of this position, see Bauer 2017; Ferrara 2009)

However, one might argue that this only becomes a problem if one thinks of authenticity as entirely a personal virtue. In other words, there is only a clash between morality and social life and being authentic if the "true" self is regarded as fundamentally prone to anti-social behaviour. But many thinkers at this time understood human nature as fundamentally disposed toward beneficence, so that evil was seen as arising from socialization and upbringing rather than from deep drives within the human being. For instance, Rousseau holds that certain immoral characteristics are immanent in man but were produced by the dynamics of modern society, which is characterized by a competitive way of relating to others and striving for acknowledgement in the public sphere. Rousseau thus externalizes the origins of societal evil and alienation from the original nature of man. The undistorted self-relation of natural man inspires sympathy and considerate relations with others, sensitive to "seeing any sentient being, especially our fellow-man, perish or suffer, principally those like ourselves" (Rousseau 1992 [1754]: 14). In somewhat the same way, economic theorists of the time supposed that unregulated markets are self-correcting, as human beings are naturally inclined to engage in mutually advantageous commercial activities (Taylor 2007: 221–269). On this view, authenticity does not amount to egoism or self-absorption. On the contrary, the prevailing view seems to have been that, by turning inward and accessing the "true" self, one is simultaneously led towards a deeper engagement with the social world. This is why Taylor (1989: 419–455) describes the trajectory of the project of authenticity is "inward and upward".

It might however be objected that supposing that the "inner" is a morally worthy guide is deeply misguided and builds upon an overly optimistic idea of human nature. It may be argued that once the idea of rational deliberation is set aside, the powerful impact of the non-rational becomes apparent. Thinkers such as Nietzsche and Freud have put in question the conception of human nature, and especially of our "inner" nature, as fundamentally good. Following their "hermeneutics of suspicion" (Ricoeur 1970), human nature comes to be seen as including forces of violence, disorder and unreason as well as tendencies toward beneficence and altruism. In that case, any idea of an ethic based primarily on the ideal of authenticity is simply untenable.

Others have expressed serious concerns not about the optimistic view of human nature, but about the conception of the self that underlies the idea of authenticity. Some argued that the dichotomies that the concept authenticity was built on, like conformity vs. independence, individual vs. society, or inner-directedness vs. other-directedness, were entirely misguided. The underlying assumption that considers the individual separate from the environment is an absurd assumption that erodes that bond between the individual and community, which ultimately is the source of the authentic self (Slater 1970: 15; Sisk 1973). In agreement with Slater (1970) and Yankelovich (1981), Bellah et al. (1985) and Fairlie (1978) contend that such a pursuit of authenticity is self-defeating, for with the loss of the bond with community, the sense of self is also diminished.

Additionally, in The Jargon of Authenticity, Adorno contended that the "liturgy of inwardness" is founded on the flawed idea of a self-transparent individual who is capable of choosing herself (Adorno 1973: 70). The doubtful picture of the self-centered individual covers up the constitutive alterity and mimetic nature of the self. In the concluding part of The Order of Things, Foucault maintained that present society was witnessing a crisis, not only of authenticity but also of the whole idea of the subject in its temporary historically contingent constitution, foreseeing that "man would be erased, like a face drawn in sand at the edge of the sea"

Notes

(Foucault 1994: 387). Foucault clearly opposed the idea of a hidden authentic self, which he critically referred to as the "Californian cult of the self" (1983: 266). The recognition that the subject is not given to itself in advance leads him to the practical consequence that it must create itself as a work of art (Foucault 1983: 392). Rather than searching for a hidden true self, one should attempt to shape one's life as a work of art, proceeding without recourse to any fixed rules or permanent truths in a process of unending becoming (Foucault 1988: 49). In a similar vein, Richard Rorty has argued that the idea of coming to "know a truth which was out there (or in here) all the time" (Rorty 1989: 27) is simply a myth. Postmodern thought raises questions about the existence of an underlying subject with essential properties accessible through introspection. The whole idea of the authentic as that which is "original", "essential", "proper", and so forth now seems doubtful. If we are self-constituting beings who make ourselves up from one moment to the next, it appears that the term "authenticity" can refer only to whatever feels right at some particular moment.

Yet others have based their criticism of authenticity especially on the emergence of a pervasive "culture of authenticity". Cultural critics have argued that the ostensible "decline" of modern society might not primarily be a result of economical or structural transformations, but as the outcome of an increasingly ubiquitous ideal of authenticity. Before we turn to these critiques, it is helpful to understand how the ideal of authenticity became so widespread. First, we should mention that Rousseau's work, made a significant contribution to the popularization of authenticity. Indeed, some argue that authenticity can be seen as a "keystone" in Rousseau's work, giving unity to his reflections on sociality, political order, and education (Ferrara 2017: 2). Particularly The New Heloise (1997 [1761]) was enormously influential, with at least 70 editions in print before 1800 (Darnton 1984: 242). This dispersion of the ideal of authenticity into popular culture was further strengthened by several factors. For instance, a wide array of intellectuals of the nineteenth and the early twentieth century had embraced the idea of authenticity, and even radicalized it by

resisting established codes and publicly defending alternative, "artistic" or "bohemian" modes of life.

The reception of the work of Sartre and Heidegger has surely contributed to the popularization of the idea of authenticity, and the decisive impact of this idea first began to manifest itself after the Second World War (Taylor 2007: 475). Rossinow contends that the politics of the 1960s were centered on questions of authenticity. Following his account, the main driving force towards political and social changes of the New Left movement in the 1960s was "a search for authenticity in industrial American life" (Rossinow 1998: 345). Both J. Farrell (1997) and Rossinow argue that the New Left emerged partly as a reaction to traditional American liberalism and Christian existentialism, replacing the negative concept of "sin" with "alienation" and the positive goal of "salvation" with that of "authenticity". Confronted with what they understood as alienation that "isn't restricted to the poor" (Rossinow 1998: 194), New Left activism reached beyond civil rights to moral rights and attempted to bring about a recovery of a sense of personal wholeness and authenticity by curing the institutions of American society.

The emerging youth culture was characterized by a severe dissatisfaction with the "morass of conformity" of the parental generation (Gray 1965: 57). The critique of the growing conformity of life got more persistent during the 1950s, and a number of social scientists in widely read books criticized what they saw as widespread conformity and inauthenticity. Among these, The Lonely Crowd (1950) by Riesman and The Organization Man (1956) by Whyte received the most attention. Riesman points out that the efficacious functioning of modern organizations requires other-directed individuals who smoothly adjust to their environment. However, he also notes that such people compromised themselves, and a society consisting mostly of other-directed individuals faces substantial deficiencies in leadership and human potential.

On the background of this development, it seems that at a time when relativism appears difficult to surmount, authenticity has become a last

measure of value and a common currency in contemporary cultural life (Jay 2004). So, under the impact of existentialism on Western culture, the ubiquitous desire for authenticity has emerged in modern society as "one of the most politically explosive of human impulses," as Marshall Berman argues (1970: xix).

14.4 CONCEPTIONS OF AUTHENTICITY

14.4.1 Kierkegaard and Heidegger

Kierkegaard's work on authenticity and his suggestion that each of us is to "become what one is" (1992 [1846]: 130), is best seen as linked to his critical stance towards a certain social reality and a certain essentialist trend in philosophical and scientific thought. On the one hand, he (1962 [1846]) condemned aspects of his contemporary social world, claiming that many people have come to function as merely place-holders in a society that constantly levels down possibilities to the lowest common denominator. In more contemporary terms, we can say that Kierkegaard provides a criticism of modern society as causing "inauthenticity". Living in a society characterized by such "massification" lead to what he refers to as widespread "despair" that comes to the fore as spiritlessness, denial, and defiance. On the other hand, he rejected the view that a human being should be regarded as an object, as a substance with certain essential attributes. Rather than being an item among others, Kierkegaard proposes to understand the self in relational terms: "The self is a relation that relates itself to itself…" (Kierkegaard 1980 [1849]:13). This relation consists in the unfolding project of taking what we find ourselves with as beings in the world and imparting some meaning or concrete identity to our own life course. Thus, the self is defined by concrete expressions through which one manifests oneself in the world and thereby constitutes one's identity over time. In Kierkegaard's view, "becoming what one is" and evading despair and hollowness is not a matter of solitary introspection, but rather a matter of passionate commitment to a relation to something outside oneself that bestows one's life with meaning. For Kierkegaard, as a religious thinker, this ultimate commitment was his defining relation to God. The idea is that passionate care about something

outside ourselves gives diachronic coherence in our lives and provides the basis for the narrative unity of the self (Davenport 2012).

Heidegger's conception of human existence (or, as calls it, Dasein, 'being-there') echoes Kierkegaard's conception of the "self". Rather than being an object among others, Dasein is a "relation of being" (Seinsverhältnis; Heidegger 1962 [1927]: 12)—a relation that obtains between what one is at any moment and what one can and will be as the temporally extended unfolding of life into a realm of possibilities. To conceive Dasein as relational means that in living out our lives, we always already care: for each of us, our being is always at issue and this is made concrete in the specific actions we undertake and the roles we enact. Over the course of our lives, our identities are always in question: we are always projections into the future, incessantly taking a stand on who we are.

The most familiar conception of "authenticity" comes to us mainly from Heidegger's Being and Time of 1927. The word we translate as 'authenticity' is actually a neologism invented by Heidegger, the word Eigentlichkeit, which comes from an ordinary term, eigentlich, meaning 'really' or 'truly', but is built on the stem eigen, meaning 'own' or 'proper'. So the word might be more literally translated as 'ownedness', or 'being owned', or even 'being one's own', implying the idea of owning up to and owning what one is and does (for a stimulating recent interpretation, see McManus 2019). Nevertheless, the word 'authenticity' has become closely associated with Heidegger as a result of early translations of Being and Time into English, and was adopted by Sartre and Beauvoir as well as by existentialist therapists and cultural theorists who followed them.[1]

Heidegger's conception of ownedness as the most fully realized human form of life emerges from his view of what it is to be a human being. This conception of human Dasein echoes Kierkegaard's description of a "self". On Heidegger's account, Dasein is not a type of object among others in the totality of what is on hand in the universe. Instead, human being is a

"relation of being", a relation that obtains between what one is at any moment (the immediacy of the concrete present as it has evolved) and what one can and will be as the temporally extended unfolding or happening of life into an open realm of possibilities. To say that human being is a relation is to say that, in living out our lives, we always care about who and what we are. Heidegger expresses this by saying that, for each of us, our being (what our lives will amount to overall) is always at issue. This "being at stake" or "being in question for oneself" is made concrete in the specific stands we take—that is, in the roles we enact—over the course of our lives. It is because our being (our identity) is in question for us that we are always taking a stand on who we are. Since the German word for 'understanding', Verstehen, is etymologically derived from the idea of 'taking a stand', Heidegger can call the projection into the future by which we shape our identity 'understanding'. And because any stand one takes is inescapably "being-in-the-world", understanding carries with it some degree of competence in coping with the world around us. An understanding of being in general is therefore built into human agency.

To the extent that all our actions contribute to realizing an overarching project or set of projects, our active lives can be seen as embodying a life-project of some sort. On Heidegger's view, we exist for the sake of ourselves: enacting roles and expressing character traits contribute to realizing some image of what it is to be human in our own cases. Existence has a directedness or purposiveness that imparts a degree of connection to our life stories. For the most part, having such a life-plan requires very little conscious formulation of goals or deliberation about means. It results from our competence in being members of a historical culture that we have mastered to a great extent in growing up into a shared world. This tacit "pre-understanding" makes possible our familiar dwelling with things and others in the familiar, everyday world.

Heidegger holds that all possibilities of concrete understanding and action are made possible by a background of shared practices opened up by the social context in which we find ourselves, by what he calls the 'They'

(das Man). Far from it being the case that social existence is something alien to and opposed to our humanity, Heidegger holds that we are always essentially and inescapably social beings. As he says,

They itself prescribes that way of interpreting the world that lies closest. Dasein is for the sake of the They in an everyday manner… In terms of the They, and as the They, I am 'given' proximally to 'myself'…. (1962 [1927]: 167, translation modified)

To be a teacher, for instance, I must adopt (and perhaps blend) some set of the ready-made styles of classroom presentation and of dealing with students laid out in advance by existing norms and conventions of professional conduct.

To say that we are always the They is not to say we are automata, however. Heidegger suggests that even in the bland conformism of "average everydayness" we are constantly making choices that reflect our understanding of who we are. Nevertheless, in average everydayness, we are as a rule adrift, acting as one of the "herd" or "crowd"—a form of life Heidegger calls "falling" (Verfallen). Heidegger (1962 [1927]: 220) emphasizes that calling this way of living "falling" does not imply that it is "a bad or deplorable ontical property of which, perhaps, more advanced stages of human culture might be able to rid themselves" (1962 [1927]: 220). On the contrary, since there is no exit from the social world—since it is the "only game in town"—it plays a positive role in creating the background of shared intelligibility that lets us be fully human in the first place. Nevertheless, Heidegger is aware that there is something deeply problematic about this falling mode of existence. In "doing what one does", he suggests, we fail to own up to who we are. We do not take over our own choices as our own and, as a result, we are not really the authors of our own lives. To the extent that our lives are unowned or disowned, existence is inauthentic (uneigentlich), not our own (eigen).

Our condition as They-selves is one of dispersal, distraction and forgetfulness. But this "downward plunge" captures only one aspect of

Dasein, Heidegger says. In order to be able to realize the capacity for authenticity, one must undergo a personal transformation, one that tears us away from falling. This is possible only given certain fundamental insights arising in a life. The first major shift can occur when one experiences an intense bout of anxiety. In anxiety, the familiar world that seemed to assure one's security suddenly breaks down, and in this world-collapse one finds that the significance of things is "completely lacking" (1962 [1927]: 186). One finds oneself alone, with no worldly supports for one's existence. In anxiety, Dasein encounters itself as an individual, ultimately alone. In Heidegger's words, "Anxiety individualizes Dasein and thus discloses it as 'solus ipse'" (1962 [1927]: 188). The second transformative event is the encounter with one's "ownmost" possibility, the possibility of death as the possible loss of all possibilities. In facing our own finitude, we find that we are always future-directed happenings or projects, where what is crucial to that ongoing forward movement is not its actualization of possibilities, but the "How" with which one undertakes one's life. Heidegger tries to envision a way of life he calls anticipatory running-forward (Vorlaufen) as a life that clear-sightedly and intensely carries out its projects, no matter what they may be. The third transformative event is hearing the call of conscience. What conscience calls out to us is the fact that we are "guilty" in the German sense of that word, which means that we have a debt (Schuld) and are responsible for ourselves. Conscience tells us that we are falling short of what we can be, and that we are obliged to take up the task of living with resoluteness and full engagement. Such resoluteness is seen clearly in the case of vocational commitments, where one has heard a calling and feels pulled toward pursuing that calling.[2]

The three "existentialia" that structure Dasein's Being-in-the-world make up the "formal existential totality of Dasein's structural whole", what Heidegger calls care. To be Dasein, an entity must have some sense of what it is "coming toward" (Zu-kunft, the German for "future"), what has "come before" (what is "passed", Vorbei), and what one is dealing with in one's current situation ("making present"). The defining characteristics of Dasein's potentiality-for-Being are displayed in the transformative events

that lead to the possibility of being authentic (eigentlich, as we saw, from the stem meaning "proper" or "own"). When Dasein confronts and grasps its authentic possibility of being, it becomes possible to see the whole of Dasein, including both its being as a They-self and as authentic being-one's-self. "Dasein is authentically itself in [its] primordial individualization", where the "constancy [Ständigkeit] of the Self … gets clarified" (1962 [1927]: 322). What defines the wholeness and unity of Dasein is determined not by an underlying substance (e.g., the sub-ject, that which underlies), but by the "steadiness and steadfastness" (beständigen Standfestigkeit, ibid) of authenticity.

The key to understanding authenticity lies, as we have seen, in the characterization of Dasein's being as a relation between two aspects or dimensions making up human existence. On the one hand, we find ourselves thrown into a world and a situation not of our own making, already disposed by moods and particular commitments, with a past behind us that constrains our choices. With respect to this dimension of human life, we are generally absorbed in practical affairs, taking care of business, striving to get things done as they crop up from time to time. This "being-in-a-situation" naturally inclines us to everyday falling as Heidegger describes it.

At the same time, however, to be human is to be underway toward achieving ends that are understood as integral to one's overarching life-project. My actions at any moment, though typically aimed at accomplishing tasks laid out by the demands of circumstances, are also cumulatively creating me as a person of a particular sort. In this sense, my futural projection as "understanding" has the structure of being a projection onto one's ownmost possibility of being. So, for example, when I attend a boring parent/teacher conference, I do so as part of handling my current duties. But this act is also part of being a parent insofar as it contributes to determining "that for the sake of which" I understand myself as existing. Given this distinction between current means/ends strategic actions and long-range life-defining undertakings, it is possible to see that there are two senses of freedom in play in

Heidegger's account of human existence. There is freedom in the humdrum sense of doing what I choose to do under ordinary conditions, a freedom Heidegger presumably interprets in an agent-libertarian way. But there is also freedom in an ethically more robust sense. In addition to choosing courses of action among options, Dasein is capable of "choosing to choose a kind of being-one's-self" (1962 [1927]: 314) through its ongoing constitution of that identity for the sake of which it exists. Thus, I attend the parent/teacher conference and behave in a particular way because I care about being a parent and a citizen of a particular sort. I understand this stance as having repercussions for my life as a whole, and I grasp the need for resoluteness in holding steady to undertakings of this sort if I am to shape my identity in the way I can care about. For Heidegger, the resolute commitment that is made concrete and defined in one's day-to-day actions is what imparts steadiness and steadfastness to a life. It is also the condition for being responsible for one's own existence: "Only so can [one] be responsible [verantwortlich]", Heidegger says (1962 [1927]: 334, translation modified). Authenticity, defined as standing up for and standing behind what one does—as owning and owning up to one's deeds as an agent in the world—becomes possible in this sort of resolute commitment to the "for the sake of which" of one's existence.

It should be obvious that this conception of authenticity has very little to do with the older idea of being true to one's own pregiven feelings and desires. But there is still a clear respect in which the idea of "being true to oneself" has a role to play here. What distinguishes this conception from the conceptions of pop psychology and romantic views of authenticity is the fact that the "true self" to which we are to be true is not some pre-given set of substantive feelings, opinions and desires to be consulted through inward-turning or introspection. On the contrary, the "true self" alluded to here is an on-going narrative construction: the composition of one's own autobiography through one's concrete ways of acting over the course of a life as a whole. Feelings and desires are, of course, profoundly important, as are the features of one's situation and one's concrete connections to others. Heidegger wants to recover a firm sense of the

wholeness of the existing individual. But this wholeness is found in the connectedness of what Heidegger calls the "happening" or "movement" of a life—that is, in the unfolding and constantly "in-progress" storyizing that continues until death. What is at stake in the ideal of authenticity is not being true to some antecedently given nature, then, but being a person of a particular sort. Heidegger emphasizes that being authentic presupposes that one instantiate such virtues as perseverance, integrity, clear-sightedness, flexibility, openness, and so forth. It should be obvious that such a life is not necessarily opposed to an ethical and socially engaged existence. On the contrary, authenticity seems to be regarded as a "executive virtue" that provides the condition for the possibility of being a moral agent in any meaningful sense whatsoever.

Others argue that Heidegger uses authenticity in both evaluative-normative and purely descriptive senses. In the descriptive use of the term, inauthenticity is simply the default condition of everyday life, in which our self-relations are mediated by others. In this sense, authenticity involves no judgment about which mode of being is superior for Dasein. But sometimes Heidegger's language turns normative (Carman 2003), and the seemingly neutral inauthentic form of relating transforms into something negative. Inauthentic Dasein is now "not itself", loses itself (Selbstverlorenheit), and becomes self-alienated. At this point, it is argued that when introducing the normative-evaluative sense, Heidegger presents three modes of life: authentic—average(ness)—inauthentic, where the authentic and inauthentic modes are existential modifications of average everydayness (Blattner 2006: 130; Dreyfus 1991). In this picture, an authentic way of life is owned, an inauthentic disowned, and the middle one—which is how we live much of the time—is simply one that is unowned. Dasein and authenticity emerge in contrast to this background and out of this background, so that the primordially indifferent mode is the condition of possibility for authenticity or inauthenticity. In addition, Carman (2003: 295) argues that Heidegger's notion of conscience can help us further illustrate his account of authenticity and shows how the "call of conscience" may be interpreted as expressive responsiveness to one's own particularity.

14.4.2 Sartre and de Beauvoir

Published in 1943, Sartre's opus magnum, Being and Nothingness: A Phenomenological Essay on Ontology, had a significant influence on philosophical thought and intellectual life in the second half of the twentieth century. His principal goal in this book is to "repudiate the spirit of seriousness" of traditional philosophy as well as of bourgeois culture (Sartre 1992a [1943]: 796). The spirit of seriousness assumes (1) that there are transcendent values that exist antecedently to humans, and (2) that the value of a thing is part of the actual being of the valued thing. Sartre's view, in contrast, is that all values are generated by human interactions in situations, so that value is a human construct with no extra-human existence in things.

To address the question of human existence, Sartre scrutinizes our everyday lives, focusing on two particular aspects. He notes that human beings, like other entities in the world, have certain concrete characteristics that make up what he calls their "facticity" or what they are "in themselves" (en soi). Facticity makes up the element of "givenness" we must work with: I find myself with a past, a body and a social situation that constrains me in what I can do. This "just being there" is above all contingent: there is no prior justification or reason for the existence of my being. On Sartre's view, the "in itself" does not even have any determinate characteristics, since every determination (every "this, not that") is first introduced into the totality of being by our specific interpretations of things.

While human beings share their "facticity" with other entities in the world, they are unique among the totality of entities insofar as they are capable of distancing themselves from what is "in itself" through reflection and self-awareness. Rather than being an item in the world with relatively fixed attributes, what is distinctive about me as a human being is that I am capable of putting my own being in question by asking myself, for example, whether I want to be a person of a particular sort. This capacity for gaining distance inserts a "not" or a "nonbeing" into the

totality of what is, which allows me to organize what surrounds me into a meaningfully differentiated whole. In addition, human consciousness is the source of the "not" because it is itself a "nothingness". In other words, a human being is not just an "in itself" but also a "for itself (pour soi)", thus characterized by what Sartre calls "transcendence". As transcendence, I am always more than I am as facticity because, as surpassing my brute being, I stand before an open range of possibilities for self-definition in the future.

Sartre's notion of transcendence is closely linked with the idea of freedom. Humans are free in the sense that they have the ability to choose how they are going to interpret things, and in these interpretations they are deciding how things are to count or matter. We constitute the world through our freedom to the extent that our ways of taking things determine how reality will be sorted out and matter to us. At the same time, we constitute ourselves through our own choices: though the facticity of my situation creates some constraints on my possible self-interpretations, it is always up to me to decide the meaning of those constraints, and this means that what I take to be limitations are in fact produced by my own interpretations or meaning-giving activities. Such limitations are grasped in light of antecedent commitments, on the background of which situations becomes intelligible, as affording certain actions and/or modes of evaluation. It is our antecedent commitments that shape our world, making situations and objects intelligible as threatening or favorable, easy or full of obstacles, or more generally, as affording certain actions (Sartre 1992a [1943]: 489). Our engagements provide a hermeneutic structure within which our situations and motives become comprehensible and reveal themselves in the way situations appear to us—as significant, requiring our attention, etc. (1992a [1943]: 485).

It is important to note that Sartre's notion of freedom is radical. Freedom is absolute to the extent that each person decides the significance of the constraints in his or her facticity: "I find an absolute responsibility for the fact that my facticity … is directly inapprehensible", because supposed "facts" about me are never brute facts, "but always appear across a

projective reconstruction of my for-itself" (Sartre 1992a [1943]: 710). For Sartre, only our choices and their projected ends define our situations as meaningful, as threatening or favorable, as affording certain actions etc. The resistances and obstacles that one encounters in a situation acquire meaning only in and through the free choice. Thus, individuals are responsible not only for their identities, but for the way the world presents itself in their experiences. Even others are just "opportunities and chances" for my free creative activity. According to this early formulation, it is up to us to interpret how other people are to matter to us relative to situations in which we find ourselves engaged (Sartre 1992a [1943]: 711).

But human beings are not merely characterized by facticity and transcendence; they are also seen as embodying a deep and irreconcilable tension between facticity and transcendence. This tension comes to the fore in Sartre's account of "bad faith". Bad faith, a kind of self-deception, involves believing or taking oneself to be an X while all along one is (and knows oneself to be) actually a Y. The most familiar form of bad faith is acting as if one were a mere thing—solely facticity—and thereby denying one's own freedom to make oneself into something very different. Thus, the person who thinks she is a coward "just as a matter of fact" is excluding from view the ability to transform her existence through changed ways of behaving. Such bad faith is a denial of transcendence or freedom.

At first, it might seem that one could escape bad faith by making a sincere, deep commitment to something and abiding by that commitment—for example, a total, resolute engagement of the self comparable to Kierkegaard's notion of an "infinite passion". In this regard, Sartre considers a person who tries to wholeheartedly believe that his friend really likes him. "I believe it", he says, "I decide to believe in it and maintain myself in this decision…" (Sartre 1992a [1943]: 114). My belief will be steady and solid, like something "in itself" that informs my being and cuts through all the tenuousness and unsteadiness of my subjective life. I know I believe it, I will say. If I could make myself

believe something in this way, then to achieve this might be what we could call "good faith:" to actually be something, without the questionability of the "not" creeping in. However, Sartre doubts that such an absolute, being-determining commitment is possible. In fact, Sartre claims that any such sort of "good faith" would actually amount to little more than another form of self-deception. For if my decision to believe is in fact a decision, it must always be something that to some extent distances me from what is decided. That is why we use the word 'believe' to imply some degree of uncertainty, as when we say, "Is he my friend? Well, I believe he is". Lucid self-awareness shows us that in making a choice, we can never attain the condition of the "in itself", because what we are is always in question for us. This is what Sartre means when he says human being is always "previously corrupted" and that "bad faith [always] reapprehends good faith" (Sartre 1992a [1943]: 116). Thus, the project of being in good faith seems impossible, as we are always necessarily in bad faith.

The inescapable nature of bad faith seems to leave no room for the possibility of authenticity. This might be why the word translated as "authentic" only appears twice in this vast tome. On one occasion, Sartre attacks Heidegger for introducing the idea of authenticity as a way of providing something foundational in an otherwise totally contingent world. The concept of authenticity "shows all too clearly [Heidegger's] anxiety to establish an ontological foundation for an Ethics…" (Sartre 1992a [1943]: 128). A second and more obscure appearance of the word comes at the end of the discussion of bad faith early in the book. Here Sartre acknowledges that his account of bad faith seems to have the consequence that there can be no such thing as good faith, so that "it is indifferent whether one is in good faith or in bad faith", and that in turn seems to imply that "we can never radically escape bad faith". Nevertheless, he goes on, there may be a "self-recovery of being which has been previously corrupted", a recovery "we shall call authenticity, the description of which has no place here" (Sartre 1992a [1943]: 116n).

One might thus conclude that there is no way to be true to what one is, because there is nothing that one is. However, such a negative conclusion would be reached only by someone who embraced from the outset the "spirit of seriousness" Sartre sets out to attack. Seriousness would lead us to think that there is simply a fact of the matter about a person: the person is either a believer or he is not. But, as Linda A. Bell (1989: 45) has noted, there is another possibility. If one rejects the spirit of seriousness, one might lucidly acknowledge that, as transcendence, one's belief is always in question and so not really a secure belief. Yet, at the same time, one might also recognize that, as facticity, one genuinely holds a belief, and that the belief is central to one's being as an engaged agent in this situation. In Sartre's convoluted style of formulation, "he would be right if he recognized himself as a being that is what it is not and is not what it is" (Bell 1989: 45). On this account, I believe, but I also acknowledge my ability to retract the belief, since nothing is ever fixed in stone.

What is suggested here is that a correlate of authenticity can be found in the idea of being true to the inescapable tension at the core of the human self. This would be attained if one clear-sightedly acknowledged the fundamental ambiguity of the human condition. Authenticity would then be what Sartre calls a "self-recovery of being which was previously corrupted" (1992a [1943]: 116). In a sense, humans can never really be anything in the way brute objects can be things with determinate attributes. In Bell's words, authenticity would be "the awareness and acceptance of—this basic ambiguity" (1989: 46). This conclusion is supported by Sartre's later work, Anti-Semite and Jew where he writes,

Authenticity, it is almost needless to say, consists in having a true and lucid consciousness of the situation, in assuming the responsibilities and risks it involves, in accepting it … sometimes in horror and hate. (1948: 90)

Lucid recognition of the ambiguity of the human condition is the leading idea behind Beauvoir's The Ethics of Ambiguity. Beauvoir takes over Sartre's characterization of the human condition and expands on ideas

only hinted at in Sartre's famous lecture, "The Humanism of Existentialism" (1946), in developing a conception of authenticity. According to Beauvoir, Sartre's conception of the human being as "engaged freedom" implies not just that each individual finds his or her "reason for being" in concrete realizations of freedom, but that willing one's own freedom necessarily involves willing the freedom of all humans. In achieving one's own freedom, she writes, freedom must also will "an open future, by seeking to extend itself by means of the freedom of others" (1948: 60). The point here is that a dedication to freedom, when clearly grasped in its full implications, will be seen to call for a future in which an unrestricted range of possibilities is open to all.

Beauvoir also builds on Sartre's notion of engagement to extend the idea of authenticity. Following Sartre, we are always already engaged in the affairs of the world, whether we realize it or not. To be human is to be already caught up in the midst of social and concrete situations that call for commitments of certain sorts on our part. Sartre takes this ground-level fact of engagement as the basis for exhorting us to be engaged in a deeper sense, where this implies that we decisively and wholeheartedly involve ourselves in what the current situation demands. Of course, once we have abandoned the spirit of seriousness, we will recognize that there are no antecedently given principles or values that dictate the proper course for our existential engagement, so that any commitment will be tenuous and groundless. But the authentic individual will be the one who takes up the terrifying freedom of being the ultimate source of values, embraces it, and acts with a clarity and firmness suitable to his or her best understanding of what is right in this context. In this way, the conception of authenticity is continuous with the ideal of being true to ourselves: we are called upon to become, in our concrete lives, what we already are in the ontological structure of our being.

This is in agreement with the manner in which Sartre describes the consequences of acting against one's deepest commitments.

There is no doubt that I could have done otherwise, but that is not the problem. It ought to be formulated like this: could I have done otherwise without perceptibly modifying the organic totality of the projects that make up who I am?

Sartre goes on to say that the character of the act may be such that

instead of remaining a purely local and accidental modification of my behavior, it could be effected only by means of a radical transformation of my being-in-the-world… In other words: I could have acted otherwise. Agreed. But at what price? (Sartre 1992a [1943]: 454)

Thus, acting otherwise or, more precisely, failing to act on one's fundamental commitments, comes at the price of transforming who one is. This change effectively precludes one from carrying on with an unchanged self-conception.

14.5 RECENT ACCOUNTS OF AUTHENTICITY

In the last three decades, authors like Taylor (1989, 1991, 1995, 2007), Ferrara (1993; 1998), Jacob Golomb (1995), Guignon (2004, 2008) and Varga (2011a) have attempted to reconstruct authenticity by maintaining that the justified criticism of self-indulgent forms of the idea does not justify the total condemnation of the idea itself (see Taylor 1991: 56). Instead of abandoning the notion of authenticity, they attempt to reconstruct it in a manner that leads neither to aestheticism nor to atomistic self-indulgence.

In The Ethics of Authenticity, and the more fully articulated Sources of the Self, Taylor makes a case for retaining the concept of authenticity (and the practices associated with it) on the grounds that the original and undistorted idea of authenticity contains an important element of self-transcendence (Taylor 1991: 15; Anderson 1995). Unsatisfied with the widespread criticism of authenticity as an adequate ethical orientation, Taylor sets out to prove that authenticity does not necessarily lead to

aestheticism or self-indulgence: the justified criticism of self-indulgent forms of the ideal does not justify the complete condemnation of the ideal itself (Taylor 1991: 56). This would mean extricating aestheticism, subjectivism, individualism, and self-indulgent interpretations of this ideal from what Taylor (Ibid.: 15) holds to be an original understanding of that concept as achieving self-transcendence (Anderson 1995). Restoring an undistorted version, Taylor says, could guard against meaninglessness, which is one of the "malaises of modernity" that Taylor regards as tied to trivialized forms of the culture of authenticity. Self-transcendence, which once was a crucial element in the ideal of authenticity, is practically lost from the contemporary version, giving rise to cultures of self-absorption, which ultimately deteriorate into the malaise of absurdity.

Already in Sources of the Self, Taylor draws attention to how modernism gives birth to a new kind of inward turn that not only attempts to overcome the mechanistic conception of the self linked to disengaged reason but also the Romantic ideal of a faultless alignment of inner nature and reason. Instead, for the modernists, a turn inward did not mean a turn towards a self that needs articulation.

On the contrary, the turn inward may take us beyond the self as usually understood, to a fragmentation of experience which calls our ordinary notions of identity into question. (Taylor 1989: 462)

While in modernism, the turn inward still contained a self-transcending moment, the critical point where the ideal of authenticity becomes flattened is when it becomes 'contaminated' by a certain form of 'self-determining freedom' that also contains elements of inwardness and unconventionality (Taylor 1991: 38). Self-determining freedom

is the idea that I am free when I decide for myself what concerns me, rather than being shaped by external influences. It is a standard of freedom that obviously goes beyond what has been called negative liberty (being free to do what I want without interference by others) because that is compatible with one's being shaped and influenced by society and its

laws of conformity. Instead, self-determining freedom demands that one break free of all such external impositions and decide for oneself alone. (Taylor 1991: 27)

Not only is self-determining freedom not a necessary part of authenticity, it is also counterproductive because its self-centeredness flattens the meanings of lives and fragments identities. For Taylor, the process of articulating an identity involves adopting a relationship to the good or to what is important, which is connected to one's membership in a language community (Taylor 1989: 34–35). As he clearly states, "authenticity is not the enemy of demands that emanate from beyond the self; it presupposes such demands" (Taylor 1991: 41). It cannot be up to me to decide what is important, since this would be self-defeating. Instead, whatever is important for me must connect to an inter-subjective notion of the good, wherefrom a good part of its normative force lastly emanates. In this sense, authenticity simply requires maintaining bonds to collective questions of worth that point beyond one's own preferences. Taylor wants to show that modes of contemporary culture that opt for self-fulfillment without regard

(a) to the demands of our ties with others, or (b) to demands of any kind emanating from something more or other than human desires or aspirations are self-defeating, that they destroy the conditions for realizing authenticity itself. (Taylor 1991: 35)

Thus, not only do we need the recognition of concrete others in order to form our identities, but we must also (critically) engage with a common vocabulary of shared value orientations. In other words, Taylor points out that authenticity needs the appropriation of values that make up our collective horizons.

In his Reflective Authenticity, Alessandro Ferrara also sets out to defend authenticity as an ideal, but in contrast to Taylor he is interested in the social and philosophical issues of the relation between authenticity and validity. According to Ferrara's diagnosis, we are currently witnessing a

profound transition that, besides affecting cultures, values and norms, also touch on the "foundations of validity," thereby affecting the "bedrock of the symbolic network through which we relate to reality and reproduce our life-forms" (Ferrara 1998: 1). At the core of this transformation is the reformulation of "well-being" (eudaimonia) as the normative ideal of authenticity, which can be of help in reconstructing a contemporary understanding of normativity. For Ferrara, it can ground a new ideal of universal validity "ultimately linked with the model of exemplary uniqueness or enlightening singularity thus far associated with 'aesthetics'" (Ferrara 1998: 10). Authenticity is then characterized by the "self-congruency" of an individual, collective or symbolic identity (Ferrara 1998: 70), and is thought of as providing a new universal validity that does not build on the generalizable but rather on the exemplary. Ferrara views Simmel's idea of an individual law as an instructive example of such an anti-generalizing universalism, and it is exactly this characteristic that makes it better suited to the pluralist contexts faced by modern Western societies. More recently, Ferrara (2019) has argued that authenticity currently faces a "dual paradox" and is misconstrued by many critics advocating its deconstructionist dismissal.

Golomb (1995) provides an informative historical overview of the genesis and development of the concept of authenticity, paying attention to both literary and philosophical sources. While continuously reminding us of the inherently social dimension of authenticity, one of the achievements here is the focus on boundary situations where authenticity "is best forged and revealed" (Ibid.: 201). Golomb takes a neutral position on the ethical value of authenticity, maintaining that "there is no reason to suppose that it is any better or any more valuable to be authentic than to act inauthentically" (Ibid.: 202).

Guignon (2004) explores both the philosophical roots of authenticity and its contemporary manifestations in popular culture. He thoughtfully criticizes pop-psychological literature that deals with the authentic life by making recourse to the subdued 'inner child'. Since Rousseau, the dichotomy between authentic and inauthentic has often been interpreted

akin to the distinction between child and adult (Guignon 2004: 43). Like the inner child, the authentic self is depicted as not yet corrupted by the pressures, competitiveness, and conformity of modern public life. Guignon draws on the psychoanalytic theories of Freud and Jung to remind us of less romanticized visions of the inner child. Additionally, Guignon (2004: 151) aims to identify the manner in which authenticity can be understood as being at the same time a personal and a "fundamentally and irreducibly" social virtue. Authenticity then involves reflectively discerning what is really worth pursuing in the social context in which the agent is situated (Ibid.: 155). If the ideal of authenticity is possible only in a free society with a solid foundation of established social virtues, it would seem that trying to be authentic, if it is to be coherent, must involve a commitment to sustaining and nurturing the type of society in which such an ideal is possible. A reflection on the social embodiment of virtues therefore suggests that authenticity, like many other character ideals, carries with it an obligation to contribute to the maintenance and well-being of a particular type of social organization and way of life (Guignon 2008: 288; 2004: 161). On the other hand, Guignon (2004, 2008) argues that in a democratic society, in which the authority of government—in setting the political course—stems from the consent of the governed, there is good reason to promote virtues like authenticity that sustain such an organization of government. To be authentic is to be clear about one's own most basic feelings, desires and convictions, and to openly express one's stance in the public arena. But that capacity is precisely the character trait that is needed in order to be an effective member of a democratic society (Guignon 2008: 288).

Varga (2011a) shares the fundamental assumption that authenticity has a certain potential (and therefore deserves to be reformulated), but he also thinks that it could be used for a critical inquiry into the practices of the self in contemporary life. By way of an analysis of self-help and self-management literature, Varga detects a "paradoxical transformation:" the ideal of authenticity that once provided an antidote to hierarchical institutions and requirements of capitalism, now seems to function both as an institutionalized demand towards subjects to match the systemic

demands of contemporary capitalism and as a factor in the economic utilization of subjective capacities. Varga argues that it is in "existential" choices that we express who we are, and that these have a complex phenomenology characterized by a sense of necessity. In such choices, described as "alternativeless choices", we articulate who we are, bringing into reality some tacit intuitions that often only take on a gestalt-like formation. In these cases, we both discover who we are "on the inside", and actively constitute ourselves. Varga's examination of the structure of our commitments culminates in the claim that the internal structure of our commitments commits us to more than what we happen to care about. In many cases it may actually commit us to publicly intelligible values that we take our commitments to embody—an aspect that may constrain the manner of our practical deliberation and the way in which we can pursue our commitments (Varga 2011a,b).

Along similar lines, Bauer (2017) defends authenticity as an ethical ideal, arguing that the ideal should understood as the combination of the ideal of expressing one's individual personality and the ideal of being an autonomous and morally responsible person. Others have argued that authenticity might require more than living in accord with commitments that one wholeheartedly endorses. For example, Rings (2017) highlights an epistemic criterion. The point is that the commitments in question have to be chosen in light of an acknowledgment of facts concerning one's personal history and present context. Thus, self-knowledge might matter more than hitherto recognized, even if the self-relation most pertinent to authenticity is not primarily of an epistemic nature.

Limitation of Authenticity

We are told: "To thine own self be true!" But what do we mean if we say that somebody is an authentic person, or a very genuine person? Personal authenticity is often defined as being true and honest with oneself and others, having a credibility in one's words and behavior, and an absence of pretence. Its meaning is then often clarified by contrasting it to inauthenticity, like comparing light to darkness. But in the absence of any

clear criteria for judging authenticity, the boundaries between being authentic and being inauthentic are amorphous and uncertain, and often porous.

The quest for authenticity is in part related to achieving some measure of autonomy and freedom – to the desire to be the architect of one's own life. Striving for personal authenticity provides an antidote to outside conditioning, and to some extent is a reaction to the inauthenticity prevalent in culture, religion, politics, and everyday life. A desire for authenticity is also essential for the discovery of the truth, and for finding fulfillment in life, making it more meaningful and comprehensible. In general, a state of inauthenticity can be a source of profound dissonance, prompting people to try to become more authentic, in harmony in their inner and outer lives.

Becoming authentic is an individual mission, since each person has their own way of being human, and consequently what is authentic will be different for each individual. Furthermore, personal authenticity is highly contextual, and depends on various social, political, religious and cultural characteristics. But the unique nature of each individual is best seen not in who he is, but in who he becomes, and becoming authentic is a continuous process, not an event. It involves not just knowing oneself, but also recognizing others and the mutual influence between individuals. If the quest for personal authenticity is just for self-fulfillment, then it is individualistic and ego-based; but if it is accompanied with the awareness of others and the wider world, then it can be a worthwhile goal.

Philosophies of Authenticity

The concept of authenticity has been explored throughout history by many writers, from ancient Greek philosophers to Enlightenment authors, to existentialists and contemporary social theorists. The social barrier to achieving authenticity (or self-realization) was emphasized by Jean-Jacques Rousseau (1712-78), who argued that personal authenticity is diminished by the need for the esteem of others in societies characterized

by hierarchy, inequality, and interdependence. According to Rousseau, authenticity is derived from the natural self, whereas inauthenticity is a result of external influences.

The existential philosopher Martin Heidegger (1889-1976) said that authenticity is choosing the nature of one's existence and identity. He also linked authenticity to an awareness of mortality, since only by keeping in mind one's inevitable death can one lead a truly authentic life. His project of realizing one's identity in the context of an external world with its influences, implies a complex relationship between authenticity and inauthenticity which means that they should be viewed not as mutually exclusive concepts, but as complementary and interdependent. Heidegger argued that both authenticity and inauthenticity are basic forms of being in the world, and they cannot be separated.

Another existentialist, Jean-Paul Sartre (1905-80), argued that there is no unchanging essence to the self, but we have a free will that allows us complete freedom to determine our lives from the choices available. According to Sartre, existence precedes essence: in other words, the human being first comes into existence and then continually defines oneself, rather than coming into being with an already given nature. So for Sartre, authenticity requires taking full responsibility for our life, choices and actions. Therefore the anxiety or 'angst' which results from our realisation of our own inescapable freedom is an integral part of authentic living. However, it should be emphasized that the individual's freedom is constrained by nature and society, as well as by their own limitations – what Sartre called their 'facticity'.

Albert Camus (1913-60) claimed that the awareness that we inhabit a universe which doesn't care about us and offers us no salvation compels the individual to recognize that the only path to freedom is authentic self-realization. To be authentic, one must be aware of the absurdity of a world with no objective morality and purpose, and create meaning in life through rebellion against the absurdity. Such personal authenticity emerges from a disregard for any (non-existent) external consolation, and

implies that the individual exists in a permanent exile, alienated from their own life, society and the universe.

Nevertheless the world has no specific inclination for either good or evil: it is what it is. No value judgements can be attached to it, even if life does not make sense from a human perspective.

These philosophical views on personal authenticity vary, but there is a common theme of personal authenticity as a dynamic process of endless becoming in a changing society and world, rather than a fixed state of being. And authenticity and inauthenticity should not be considered as mutually exclusive states, but rather as mutually dependent concepts.

Some Basic Qualifications

The concept of 'authenticity' is a human construct, and as such it has no reality independent of minds. But is authenticity possible, or even desirable? The question is possibly misleading as it implies an absolute yes or no answer, and does not allude to any possibility of 'partial authenticity'. This steers us toward an interpretation of the concept of authenticity as an absolute, but in general the search for absolutes is fruitless. So let's consider some things that can limit absolute authenticity.

Some argue that authenticity is impossible to achieve as an ongoing state of being, since any real authenticity is transient and impossible to maintain indefinitely. And like identity, authenticity cannot be adequately defined or measured, since many characteristics of an individual are in constant change, with no fixed reference points. Individuals undergo changes throughout life with the deluge of observations and interpretations, so human identity is multidimensional and dynamic; it is a work in progress rather than a fixed state. Therefore, attaining some measure of personal authenticity is a lifelong project that may never be fully accomplished. Personal authenticity involves principles and ideals which are continually revaluated through self-examination and social interaction, so who is to judge if someone else is being authentic or not?

The key question is, how can we distinguish between true authenticity and a mere display of authenticity? If one's 'authenticity' is being promoted, highlighted, or exhibited, then it is not true authenticity. Authenticity cannot be declared, publicized, instructed, marketed, or exchanged as some sort of commodity. It must be understated and unpretentious.

Being true and honest to oneself and others is relational, and connected to the outside world as well as to one's inner life. However, to avoid aggravating others, one must observe the need to limit the expression of one's authenticity in specific situations. One may thus distinguish between 'internal authenticity' and 'external authenticity'. To avoid burdening others with our personal issues, we may often be inclined to hide our true feelings. True authenticity isn't about expressing one's inner self with its full range of shifting emotions in all situations. Unbiased self-awareness in the present moment is of great importance, as it can enhance the clarity of one's inner dialogue and diminish the reach of the ego.

But being true and honest is not enough. There are certain attributes without which the concept of 'personal authenticity' would remain an empty shell, ambiguously defined and poorly understood, and without which the quest for authenticity may in fact become detrimental for interpersonal relationships and for society. These characteristics necessary for authenticity include capacities for unbiased self-examination and accurate self-knowledge; reflective judgment; personal responsibility; humility; empathy for and understanding of the other [person], as well as a willingness to listen to feedback from others. Achieving personal authenticity is complicated by the presence of illusions and biases, including self-deception, wishful thinking, and the tendency to behave differently while under observation.

Paradoxes of Authenticity

It's a paradox that one can discover some measure of personal authenticity not by avoiding the outside world, only by immersing oneself in it; and yet authenticity is achieved by resisting outside influences in

one's self-realization. Furthermore, since human lives operate with uncertainties, authenticity can only be discovered in uncertainty. Thus, another paradox is that the authentic can only be attained through an immersion in uncertainty, but uncertainty hinders the discovery of the true self, without which knowledge authenticity cannot be achieved. In addition, any objective discovery of the self is only possible without preconceptions and biases – but we all have preconceptions and biases. Therefore no self-examination, however long and detailed, can ever fully reveal one's true identity, and thus what being authentic would truly involve. And difficult circumstances can also lead to self-doubt and insecurity; true self-knowledge must make allowance for this.

The question is, how do you really know whether you are being authentic or not? One does not consciously consider whether one is being authentic throughout daily life. But on the other hand, complete self-knowledge is impossible: one cannot possibly explore the entire labyrinth of human consciousness. And to a large extent, cognitive processes, such as perception and reasoning and much of the content of memory, are inaccessible to conscious awareness. The tendency is to fill the gaps between the known and the unknown with the known facts and thoughts about oneself, in order to provide a coherent portrayal. As a result, self-examination may lead to an inaccurate self-depiction. The ability of the human mind to examine great amounts of information or multiple aspects of a given topic is limited, further leading to an incomplete understanding or an erroneous representation of what is observed or experienced. Human knowledge always remains incomplete and provisional, yet without full awareness, no complete authenticity is attainable, so at any given moment, authenticity can be only partial.

In addition, there will coexist in one individual multiple identities dependant on the roles the individual holds in society, including personal, occupational, cultural, ethnic, national, political, and religious identities. The dynamics of identity changes can be as complex and unpredictable as changes in society, economics and politics. These factors make any unambiguous discovery of personal identity a difficult challenge,

especially as the analysis is usually simplified, and the intricate interdependence of the various elements is typically overlooked. This can result in the illusion of understanding personal identity, and thus an illusory ideal of authenticity.

Human judgments and attitudes are based on the interpretation of perceptions of reality rather than on the interpretation of reality itself. The limits of human perception, thought and self-knowledge, are some of the main hurdles to personal authenticity. One may never arrive at full self-knowledge, which is constantly being defined and refined on the basis of new understandings against the background of the world and its demands.

Another limitation in the quest for authenticity is related to the language used, which is open to misinterpretation, and words and language are inadequate for expressing the full spectrum of one's thoughts and feelings. Allegories, connotations, and metaphors are the major sources of misunderstandings. In addition, words and sentences are often ambiguous, having more than one possible meaning. A completely clear language with a direct and evident correspondence between experiences and words does not exist. And the individual's shifting thoughts and perceptions about themselves may not always be comprehensible, so the expression of them through language may not be consistent. Also the language to describe authenticity can itself be arbitrary and unclear, often using ambiguous words such as 'true', 'genuine', 'original', 'real', 'self', or 'natural'.

Authentic communication depends on the capacity of individuals to recognize what is true for themselves, and on the adequacy of language to express their thoughts, so the limits to language, interpretation, and expression impede their authentic relationships with each other. In such communications it is not always apparent whether the authenticity or inauthenticity of interactions is due to the circumstances, or the language, or the subject matter, or the participants and their perceptions and interpretations.

Another dilemma with personal authenticity is related to the fact that most personal attributes change with time, yet personal authenticity is expected to demonstrate some measure of consistency. This apparent contradiction involves a requirement for both change and constancy. In other words, if an individual's identity is continually evolving, how can one recognize or discover the meaning of personal authenticity over a lifetime? Perhaps the value of authenticity is not in its constancy, but rather in its consistent evolution throughout the lifetime of the individual.

Further Limitations on Authenticity

Other factors that may hinder the development of personal authenticity include a lack of understanding of authenticity, one's prior programming, the fear of rejection and failure, and social pressures to conform (and thus live inauthentically). In the latter cases, individuals typically try to show their best faces and express what is expected of them so that they will be perceived in a good light. In many situations, the need for collaboration with others may demand some adaptation, that is, some inauthentic compromise.

Being under constant observation and scrutiny at best inhibits authenticity, and at worst makes it nearly impossible. Thus, politicians cannot be authentic, since they always have to appear confident and nearly flawless rather than show any honest doubt or vulnerability. The politicians are on the stage of the theatre of life, where they must perform their art of convincing and pleasing groups of people with different values, aspirations, beliefs, dreams, and needs. So in politics, authenticity is at least difficult to sustain, although a politician may have a well-developed capacity of self-reflection and the desire to ensure consistency between their actions and values, and so may be able to achieve authenticity in some situations. Nonetheless, the public expects politicians to deliver on their promises, and the bottom line is always prosperity, jobs, and security, so in this context, authenticity takes second place to other concerns.

The attempt to achieve personal authenticity is also exacerbated by ever-advancing technologies that inundate an individual's perception of reality with illusions, such as television. Another such technology is virtual reality. The pervasive use of virtual reality may eventually result in difficulties distinguishing between virtual and real experiences, exacerbating the endless human propensity for self-deception and self-delusion. Generally, the future of human experience is related to emerging enhancement technologies, including memory and cognitive enhancement techniques. The merging of human and machine may necessitate new definitions of what a human being is, and generate new problems related to human nature and identity, the nature of society, and the meanings of existence and human authenticity.

If life is an art, as in any art form, one can approach perfection, but one can never arrive. As for personal authenticity, some never bother with it, some discover it in certain actions, some strive to approach it in both life and art, but very few ever arrive.

Check Your Progress 1
Note: a) Use the space provided for your answer
b) Check your answers with those provided at the end of the unit

1. What do you know the Origins and Meaning of the Concept of Authenticity?

   _____

   _____

   _____

2. How do you criticise of Authenticity?

   _____

   _____

   _____

3. How do you find out the Conceptions of Authenticity?

_____

_____

_____

4.   What do you understand the Recent Accounts of Authenticity?

_____

_____

_____

14.6 LET US SUM UP

Authenticity is a concept in psychology (in particular existential psychiatry) as well as existentialist philosophy and aesthetics (in regard to various arts and musical genres). In existentialism, authenticity is the degree to which an individual's actions are congruent with their beliefs and desires, despite external pressures; the conscious self is seen as coming to terms with being in a material world and with encountering external forces, pressures, and influences which are very different from, and other than, itself. A lack of authenticity is considered in existentialism to be bad faith. The call of authenticity resonates with the famous instruction by the Oracle of Delphi, "Know thyself." But authenticity extends this message: "Don't merely know thyself – be thyself."

Views of authenticity in cultural activities vary widely. For instance, the philosophers Jean Paul Sartre and Theodor Adorno had opposing views regarding jazz, with Sartre considering it authentic and Adorno inauthentic. The concept of authenticity is often aired in musical subcultures, such as punk rock and heavy metal, where a purported lack of authenticity is commonly labeled with the epithet "poseur". There is also a focus on authenticity in music genres such as "...house, grunge, garage, hip-hop, techno, and showtunes".

14.7 KEY WORDS

Existentialism: One of the greatest problems facing such abstract approaches is that the drives people call the "needs of one's inner being" are diffuse, subjective, and often culture-bound. For this reason among others, authenticity is often "at the limits" of language; it is described as the negative space around inauthenticity, with reference to examples of inauthentic living.

14.8 QUESTIONS FOR REVIEW

1. What are the limitations to authenticity in Philosophy?

2. Discuss the relationship of Authenticity and non-authenticity.

14.9 SUGGESTED READINGS AND REFERENCES

- Erich Fromm. Escape from Freedom; Routledge & Kegan Paul 1942

- Lionel Trilling. Sincerity and Authenticity; ISBN 0-19-281166-5; Harvard UP 1974

- Charles Taylor. The Ethics of Authenticity; ISBN 0-674-26863-6; Harvard UP 1992

- Alessandro Ferrara. Reflective Authenticity; ISBN 0-415-13062-X; Routledge 1998

- James Leonard Park. Becoming More Authentic: The Positive Side of Existentialism; ISBN 978-0-89231-105-7; Existential Books 2007—5th edition

- Achim Saupe. Authenticity, Version: 3, in: Docupedia Zeitgeschichte, 12. April 2016; DOI

14.10 ANSWERS TO CHECK YOUR PROGRESS

Check Your Progress 1

1. See Section 14.2
2. See Section 14.3
3. See Section 14.4

Notes

4. See Section 14.5