**Chapter (5): Methodological dualism (MD) and the multi-explanation framework (MEF) (II): Mechanistic and mentalistic explanations and theories of understanding**

This chapter is divided into two parts. In the first part, I will develop the MEF approach, which explains how complicated behavior can be explained with the help of two explanatory models: mechanistic and mentalistic. In the second part, I will discuss the influence of the MEF approach on the way we understand behavior, drawing attention to two procedures for understanding a given phenomenon, one based on science and the other on folk psychology (see Chapter 4).

**Part I: The multi-explanation framework (MEF): core ideas**

The ground was prepared to develop the MEF in the previous chapter on methodological dualism. First, it was clarified that a good understanding of behavior requires two explanatory models: mechanistic and mentalistic. Secondly, it was shown that [Motivation–Belief] as a mentalistic explanatory model meets most of the requirements of the scientific method. Therefore, this model may be seen as a procedure for creating specific mentalistic explanations for behavioral phenomena. The present chapter is a natural continuation of what has been achieved by the development of MD. Hence, assuming that the mentalistic explanation model [Motivation–Belief] is methodologically acceptable, the following question arises: How can these different explanatory procedures, mechanistic and mentalistic, be used to explain behavior that is more complicated than, for example, a simple reflex? The MEF offers a solution to this question. To clarify this, I will use an example from daily life, which I shall refer to as the Morning Episode: Ronny woke up in the morning feeling sleepy, got up, went to the kitchen, and made himself a cup of black coffee. How can this behavior be explained?

According to the mentalistic model of explanation, Ronny *felt* thirsty and sleepy and *believed* that making himself a cup of black coffee would satisfy his thirst and reduce his drowsiness. He therefore made himself a cup of black coffee. This explanation answers the question: *Why* did Ronny make a cup of coffee? However, it does not answer the question: *How* did the action of making the coffee and all the other actions related to it take place? If we expand the explanation to answer the question of how, we will have to appeal to a mentalistic description of the Morning Episode that encompasses many actions, some of which have to be explained mentalistically and some mechanistically. I shall first examine a behavior that can be explained by the mentalistic approach. For example, when Ronny opened the fridge, he had to decide between two types of coffee: instant or fresh. After some thought, he chose fresh coffee. Why? The mentalistic explanation is this: Ronny remembered that fresh coffee wakes him up more effectively than instant coffee, and he believed that this time, too, fresh coffee would dispel his drowsiness, so he chose fresh coffee. In this case, conscious recall of the fresh coffee experience motivates Ronny’s choice.

We will now move on to a mechanistic explanation of behavior. Automatic processes operate all the motor actions (for example, standing up, walking, opening one’s eyes, and so on). Ronny wakes up and automatically navigates his familiar environment (he does not ask himself: Where am I? Could it be that I am still dreaming?). He immediately finds his bearings by comparing what he sees when he opens his eyes and the knowledge stored in his mind about his bedroom, a comparison that leads him to routine behavior in his apartment. As can be seen, several conclusions emerge from the Morning Episode.

(A) *Level of explanation*: It is possible to offer several explanations at different levels of expansion. In the present case, with a first degree of expansion, it is possible to answer the question of *why*: Why did Ronny make himself a cup of black coffee in the morning? A satisfactory explanation is offered at the mentalistic level: [Motivation–Belief]. However, if one attempts to answer the question of *how* (how do we explain the chain of actions that led Ronny to prepare a cup of black coffee?), we have to expand the explanation. Now, we must explain several behaviors that are all organized within the general description of the Morning Episode. We must offer a mentalistic explanation for Ronny’s choice of coffee and appeal to a number of mechanistic explanations for the chain of actions involved in getting up and moving around his living space.

(B) *The explanation–behavior match*: It is clear from what was described above that there is a need to match the type of explanation to the type of behavior in question. After all, it would be unthinkable to suggest that chemical processes in the synapses of brain cells in the nervous system are explained by mentalistic concepts such as desire and belief. In this case, the appropriate explanation is mechanistic. Conversely, it is difficult to offer a mechanistic explanation for mentalistic behavior (saturated with CΨ) because, as explained above, no theory has been found that satisfactorily reduces CΨ to neurophysiological processes in the brain. Therefore, it is hard to see how a chemical or neurophysiological process would explain Ronny’s desire to see an action movie. On what basis can the explanation–behavior match be made? The answer is not simple (as far as I know, there is no simple formula), and it depends on two fundamental components: (B1) and (B2). I will call this approach the MEF-matching theory (MMT), a theoretical approach that helps to match the type of explanation (mechanistic or mentalistic) to the behavior and its components.

(B1) *Theoretical and empirical knowledge*. The decision whether to use a mechanistic or mentalistic explanatory model given a certain behavior is based on theoretical and empirical knowledge related to this behavior. To reiterate, I do not believe anyone would try to explain synaptic processes in the brain by using concepts like desire or belief. Moreover, I am not sure that anyone would give a mechanistic explanation for Ronny’s choice of black coffee based on the blood circulatory system. An explanation based on his experience related to black coffee is satisfactory. However, there is another option for determining the explanation–behavior match.

Pylyshyn (1984) proposed the empirical criterion of ‘impenetrability’, which can also be applied to the present subject to decide whether or not the behavior in question needs a mechanistic explanation: If changes in the goals, desires, beliefs, intentions, and knowledge of the individual do not affect the behavior being studied, then this behavior needs a mechanistic explanation. For example, even if we know everything there is to know about the famous Müller-Lyer geometric illusion, this knowledge does not eliminate the illusion. It is therefore reasonable to suggest that the illusion should be explained by a mechanistic process whose output is induced with CΨ.

(B2) *The matching principle*. Rakover (2018) proposed this principle (criterion) to solve the following problem. Consider behavior A, which is based on two different actions: a1 and a2. What is the relation between the type of explanation for behavior A as a whole and the type of explanation for each of the two actions, the components of A, a1 and a2? The answer is this:

(B2.1) *Mechanistic explanation of A*: If the appropriate explanation for behavior A as a whole is mechanistic (the specific type of mechanistic explanation is not important here), then both actions, a1 and a2, must be explained mechanistically. In other words, if A is explained mechanistically, neither a1 nor a2 can have a mentalistic explanation.

(B2.2) *Mentalistic explanation of A*: If the appropriate explanation for behavior A as a whole is mentalistic (the specific type of mentalistic explanation is not important here), then actions a1 and a2 may each separately have a mentalistic or mechanistic explanation. In other words, if A is explained mentalistically, then the actions that make up Amay have mentalistic or mechanistic explanations.

Several interesting conclusions arise from the matching principle:

1. *Explanation in natural sciences*. The matching principle is methodologically consistent with scientific explanations. It is clear from this principle that if the flashlight’s illumination has a mechanistic explanation (due to an electric current passing through the bulb wire), then the workings of the components of the flashlight, such as the battery and the electric current, also require mechanistic explanations. This principle does not contradict how explanations are proposed in the natural sciences.
2. *Organization and explanation of the components*. The illumination of a flashlight and the functioning of its parts can be explained by energy transformations. For example, a chemical reaction in the battery produces electricity that passes through the conductor when the electric circuit is closed, and the electricity passing through the filament in the incandescent bulb creates light when it heats the filament. By comparison, the [Motivation–Belief] model of explanation organizes the elements of a given behavior. For example, the Morning Episode is organized by a mentalistic explanation that encompasses other mentalistic and mechanistic explanations for the various actions that constitute the Morning Episode.
3. *The primacy of the mentalistic explanation*. It seems that any explanation of complicated behavior (such as the Morning Episode) must begin with a mentalistic rather than a mechanistic explanation. If the behavior being studied is not a simple reflex, for which the appropriate explanation is mechanistic, it will be difficult to explain mechanistically. Why? Because a mechanistic explanation for the behavior as a whole requires mechanistic explanations for all its components, but this will not be possible for some of them—for example, Ronny’s decision to choose fresh black coffee over instant coffee. Therefore, a mentalistic explanation must come first, and then it will be possible to decompose the behavior into its components and offer the appropriate explanation: mentalistic or mechanistic.

**MEF-matching theory (MMT): explanation and testing**

In this section, I examine first the following question. What is the process for explaining the natural sciences compared to the explanatory process of the MEF? Then, I will examine the procedure for the empirical testing of a scientific theory.

*Proposing an explanation.* The fundamental idea is that there is a difference between an explanation proposed in the sciences and one proposed by the MEF. In science, the explanatory model uses a law, theory, or mechanism to explain the phenomenon being studied. In the MEF, two fundamental steps are required. In the first stage, the MEF-matching theory (MMT) is used to determine which type of explanation should be used: mechanistic or mentalistic. If it is determined that a mechanistic explanation must be used for a certain behavior, then the process of explanation is the same as that employed in the sciences. For example, Galileo’s law of falling bodies can be used to explain the vertical distance the body will fall at different times. However, if it is determined that a mentalistic model should be used, the [Motivation–Belief] model is activated in the following way. Specific values corresponding to the behavior under study are inserted into the general mentalistic scheme. Consider the following episode. Ronny entered the yard to put some mail in the mailbox, but instead, he quickly ran off and hid behind a tree. How do we explain this behavior? The appropriate mentalistic explanation [Motivation–Belief] is that David *wanted* to put the mail in the box, but suddenly, a large barking dog appeared in the yard, and the *frightened* David ran away and hid behind the tree. (For the sake of brevity, I am omitting in this explanation the parts related to David’s *belief* that entering the yard will fulfill his desire to post the mail and the *belief* that hiding behind the tree will help him to escape from the dog). The important point here is that it is easy to produce a specific mentalistic explanation for David’s behavior by placing the appropriate specific values in [Motivation–Belief].

Now, I will expand a little on the above ideas and illustrate them with simple examples. I will start the discussion with explanations in the sciences and then move to explanations according to the MEF approach. I will examine two mechanical devices and answer the basic question of how to propose a mechanistic explanation for the action of an electric kettle and the operation of a flashlight. Then, I will clarify whether applying the mechanistic explanatory methodology to behavior is possible.

*Sciences.* I distinguish simple phenomena, such as the free fall motion of bodies, which may be explained by appealing to a natural law, from complex phenomena, such as the eruption of a volcano or the functioning of a car, airplane, or even an electric kettle or flashlight. Complex phenomena are explained by appealing to several natural laws or theories. How do we explain why water boils? To do so, we must divide the working of the kettle into three main sub-systems connected with electricity, heat, and water. Using the laws of electricity, we explain how electricity passing through a resistance creates heat, and by employing the appropriate laws of thermodynamics and chemistry, we explain how the heat rises from the bottom of the kettle and makes the water boil. Hence, the kettle’s overall action is accounted for by explanations for each of its components and by the proper combination and organization of the operations of these components. The appropriate organization of the kettle’s components is based on the process of energy transformation (e.g., electricity to heat). This explanation is based on a chain of different processes (acting one after the other and simultaneously), all explained by mechanistic explanatory models that use various natural laws and theories relevant to electricity, heat, and water.

Fundamentally, explaining how a flashlight works is similar to explaining the functioning of the electric kettle. How does this instrument work? To answer the question, one performs the following theoretical and empirical analysis. We dismantle the flashlight into its parts: switch, battery, electric lead, and bulb. We explain every part and their interactions: When the switch is operated, the electric circuit is closed, and the current passes through the filament in the bulb (e.g., a tungsten wire), heats it, and, as a result, the bulb emits light. Here, too, one observes energy transformations, which are accounted for by mechanistic models of explanation: The action of the battery is based on chemical and physical processes, and the action of the filament is based on electrical resistance. We explain the connections across the different parts as processes of energy conversion: chemical energy, which converts into electrical energy, which converts into heat, which emits light. This energy conversion can be calculated precisely and explained by the laws and theories appropriate to the mechanistic models.

As these two examples show, the main difference between the explanations of simple and complex mechanistic phenomena is this: In the latter case, one needs to know (a) how many mechanistic theories (laws) to use and (b) how these explanatory theories interact. In the case of simple phenomena (e.g., free-falling bodies), we apply a natural law appropriate to the simple phenomenon under study within the framework of the explanatory model (e.g., the D-N model); in the case of complex phenomena (e.g., electric kettle, flashlight), we apply different natural laws, theories, or mechanisms, which match the components of the studied complex phenomenon within the framework of the mechanistic explanatory model. The explanatory connections between these components (and their mechanistic explanations) are made through other mechanistic explanations based on the transformation of energies across the components, which unify them into one functioning system.

In view of the above, one may suggest the following relationship between models of explanation, theories, and complex phenomena in the sciences:

*EXPLANATION in the sciences: One or several explanatory models encompass various theories. These theories are connected or interact in a certain order through energy transformations to generate a satisfactory unified explanation of the studied phenomenon.*

*Behavior*. In a similar way, I distinguish simple behavior, which can be explained mechanistically, from complex behavior (pertaining to humans and supreme animals), which, according to the MEF, must be explained mechanistically *and* mentalistically. Thus, one may propose that the above methodological approach (developed for the sciences) cannot be applied straightforwardly to explanations of complex behavior.

The chief reason for this is that we do not have a theory of transformation from the mind to the body and the reverse—we do not possess a mind–body theory, a theory of CΨ (TC). We cannot explain how mechanistic components connect to mentalistic ones and vice versa. Therefore, we cannot propose a unified mechanistic explanation for complex behavior, as in the case of the electric kettle and the flashlight. For this reason, the procedure for explaining behavior (which is more complex than a simple reflex) is complicated. It includes two stages. In the first stage, a theory is applied that matches (a) the type of explanation (mechanistic, mentalistic) to (b) the behavior and its components. In the second stage, the explanation is established. If the explanation for the behavior should be mechanistic, then the procedure for providing a mechanistic explanation is activated, as described above. However, if the explanation should be mentalistic, then the appropriate values (motivation, belief) for the behavior in question are placed in the model of the mentalistic explanation [Motivation–Belief].

For example, a complete explanation of David’s behavior in saying goodbye to his daughter Ruth at the airport differs from the unified explanation for the workings of a flashlight. As mentioned above, the explanation for the latter is provided by decomposing the flashlight into parts (battery, bulb, etc.), and explaining the operation of each part using the appropriate mechanistic theory, and by explaining the interaction between these parts based on the transformation of energy. This kind of explanation cannot work with an MEF explanation of David’s behavior on saying farewell because we still do not know how mental processes (e.g., his desire and belief) interact with neurophysiological processes (e.g., hand waving) or how CΨ emerges from the brain. In short, a successful TC has not been found; therefore, we cannot provide a complete and unified explanation for behavior as we can for the flashlight. Any explanation for David’s behavior must include a mentalistic explanation for his inner world and a mechanistic explanation for his public behavior.

In view of the above, one may suggest the following relationship between models of explanation, theories, and complex behavior in psychology:

*EXPLANATION in psychology: An explanation of complex behavior employs several explanatory models (mechanistic and mentalistic), which cannot be connected since no TC exists. As a result, although the MEF approach provides improved explanations in comparison to mechanistic explanations alone (as observed in behaviorism, cognitive psychology, and neurophysiological psychology), a completely unified account for the studied complex behavior cannot be provided.*

Note that this situation in psychology is probably an important reason why this field is characterized by such a diverse range of specializations and has not succeeded in developing in any area (e.g., perception, learning, cognition) a great theory like the Newtonian theory of gravity, the theory of relativity, or quantum theory in physics (see Rakover, 2020. This subject will be developed in Chapter 9).

**The MEF approach and empirical testing**

In this section, I examine the following question. What is the process for empirically testing theories in the sciences compared to testing theories generated by the MEF? The answer is: There is no difference. The empirical testing process is the same in both cases.

While the natural sciences and the present approach differ methodologically in providing explanations, they do not differ in their methodology of empirically testing theories. It is not problematic to test each theory explaining the functioning of each component in the electric kettle or flashlight: the intensity of the heat produced by the electrical resistance in both the kettle and the flashlight; the time taken for the water to boil; the intensity of the light produced by the filament in the bulb, etc. In each case, one only has to do the appropriate calculations and compare the prediction with the corresponding empirical measurements. Similarly, it is straightforward to test an explanatory mentalistic theory: It can be tested in a similar way to testing a theory in the natural sciences. The main reasons for this are as follows.

The testing method (the hypothetico-deductive [H-D] method) is indifferent to the kind of explanatory theory (T) that the researcher uses to account for the experimental results. A prediction is derived from T and auxiliary hypotheses. If the prediction matches the observation, T is confirmed; if not, T is disconfirmed. The H-D method can be used as long as a prediction can be derived from the theory under study (by employing certain rational means), and the prediction can be compared with the observation. For example, it is not difficult to empirically test the desire–belief explanation that David waved his hand as a sign of his wish to take leave of his daughter, Ruth. Since David is acquainted with Ruth, one can predict that he will recognize (correctly select) her photo out of ten different images, and Ruth, who will recognize David’s photo as well, will confirm that she saw David waving goodbye and waved back.

**Part II: Two-factor theory of understanding (TFTU)**

The present theory is based on two fundamental factors. The first posits that CΨ is a necessary condition for understanding: It is impossible to understand an explanation if it is not, or has not been, represented in an individual’s CΨ. The second factor posits that understanding stems from responding to questions in a particular field, posed according to procedures and knowledge relevant to that field (as detailed below). These procedures are intended to produce explanation and lead to an understanding of the phenomena studied in the natural and social sciences and in daily life according to the MEF approach.

## Consciousness is a necessary condition for understanding.

The fundamental statement is that without CΨ, one is unable to understand what is happening in the world, including one’s own actions; that is, CΨ is a necessary condition for understanding. The emphasis here is on *necessary condition* and not sufficient condition. Several other factors come into play, and if they are not functioning normally, one may not understand the content of one’sCΨ. A brain injury, for example, may interfere with information retrieval processes crucial for understanding a particular behavior (e.g., see Rakover, 2018, 2021).

The assumption that CΨ is necessary to produce understanding is very important since it proposes a distinction between providing an explanation and understanding the explanation. For example, a sophisticated robot—let us call him Robbie—can be programmed to answer any question in classical Newtonian physics to help students learn. For example, Robbie can answer, with limitless patience, any question concerning the free fall motion of bodies, until every single student has a perfect grasp of the answer to each question. However, although the students may grasp Robbie’s explanations completely, Robbie himself does not understand a single element of his explanations because he lacks CΨ. For these reasons, I suggest that an explanation is understood only when it is represented in an individual’s CΨ. Thus, explanations offered by both *scientific procedures* and *everyday procedures* are understood only when represented in an individual’s CΨ.

## Understanding emerges from answers to questions asked according to knowledge and procedures relevant to a specific field.

In addition to the assumption that CΨ is a necessary condition for understanding, I propose that further conditions are required to delineate this concept. I suggest two basic ideas related to understanding: (A) demarcation of understanding and (B) assessment of understanding.

(A) *Demarcation of understanding*. The basic idea is that, in addition to CΨ, which is a necessary condition for understanding, understanding is limited to cases in which a question arises in a particular field of knowledge and which is answered in accordance with certain procedures widespread in the period in which the question arose. The demarcation of understanding also depends on two factors: the relevant knowledge and the procedures with which one attempts to achieve understanding of the phenomenon in question. For understanding to occur, an answer must be given to a question about the phenomenon, such as why or how the phenomenon occurred. The answer suggests new information that sheds light on the phenomenon. However, understanding cannot occur when new information (an explanation) does not reconcile with prior knowledge rooted in an individual’s cognitive system. For example, a person who believes in the geocentric approach will have great difficulty understanding Kepler’s laws based on the heliocentric approach.

What are the specific procedures for imparting understanding? These can be characterized as processes in a particular field of knowledge that form connections between (a) new, relevant information and (b) information about the phenomenon in question; that is, processes that present satisfactory answers to questions concerning the phenomenon in question.

(B) *Assessing understanding*.Assessment of understanding is based on the following assumptions:

1. Every observable behavioral phenomenon (whether external or internal to a person) has an unknown real process (URP) responsible for its occurrence and for the occurrence of other related phenomena (“responsible” is a general term encompassing causes, reasons, mechanisms, functions, etc.).
2. Theories (models, hypotheses) that provide answers to questions of why and how can be rated on a scale of distance from the URP, a distance that can be termed the “understanding distance (UD).”

UD = F (Theory – URP).

The ‘F’ (a function of) is linked to the procedure of empirically testing a theory. That is, it is possible to roughly estimate the UD through an empirical test of a theory that provides answers to these questions. The greater the theoretical success of the explanation, the smaller the UD (for further developments, see Chapter 7).

To clarify the two ideas discussed above, I will briefly analyze four examples from various spheres of knowledge.

*Mathematics*: Consider the question, “What is the value of *x* in the equation 2*x*+1=5?” The answer, *x*=2, is obtained by applying a mathematical procedure relevant to the field of first-degree equations. We may assume that a respondent who gives this answer has understood the question, the answer, and the process that led to it.

*Physics*: Consider the question “What is the vertical distance traveled by a free-falling body in the first second?” The answer, 4.91 meters, is obtained by applying Galileo’s law of falling bodies based on the concepts of gravity and acceleration. We may assume that a respondent who gives this answer has understood the question, the answer, and the theoretical process required to reach it.

*Transportation:*Consider the question “Why did David stop his car when he reached the junction and the traffic lights changed from green to red?” The answer is that David must stop his car in front of the white line on the road when the traffic light at the junction turns red. Since this answer is correct, we consider the respondent to have understood the question, the answer, and the traffic laws according to which motor vehicles must be driven.

*Medicine*: Consider the question “Why, in ancient times, did the village shaman sing and dance around a person complaining of severe abdominal pain?” According to the medical practices and beliefs of that time, the shaman had the power to banish evil spirits that made people sick by singing and dancing. Since we consider this answer to be correct, we may assume that the respondent has understood the question, the answer, and the medical practices and beliefs held at that time.

These examples raise some interesting questions related to the subject under discussion.

(1) *Can one propose a general procedure for reaching understanding?*It has emerged from the discussion so far that understanding depends on two main factors (in addition to the necessary condition of CΨ): *a procedure of understanding, and knowledge relevant* *to a specific domain*. Every field has its explanatory procedure. Thus, for example, the knowledge and procedure for understanding the solution to a first-degree equation cannot be used for the field of transportation. Therefore, one cannot propose that there is a general procedure for reaching understanding.

Moreover, explanatory models in the sciences—*scientific procedures—*do not converge into a single general explanatory procedure or model (see Rakover, 2018). Different scientific fields have different procedures. For example, in classical physics it is convenient to use explanatory procedures consistent with Hempel’s (1965) approach, which provides explanations only when predictions deduced logically from a theory under certain conditions match the observed phenomenon. (If a prediction does not fit the phenomenon, the theory is refuted.) However, in biological, neurophysiological, and cognitive research, it is convenient to use procedures that propose mechanisms comprised of numerous elements in such a way that the particular interaction between these elements ultimately produces the phenomenon being studied (see discussions in Bechtel, 2008; Rakover, 2018).

(2) *Are there different degrees of understanding?* Some scholars have argued that there are various degrees of understanding (see, for example, Gordon, 2020; Grimm, 2011, 2019; Khalifa, 2013, 2017). The two-factor theory of understanding is consistent with this view. Accordingly, for understanding to be acquired, a question must be asked regarding a particular aspect of the phenomenon under study. The answer to an initial question opens the door to a second, deeper question, the answer to which is closer to the fundamental correct answer to the empirical problem. This process ranks the level of understanding according to the depth of the question and how close the answer gets to the fundamental explanation.

(3) *Are there procedures of understanding that are not compatible with the scientific method?*The answer is affirmative. I assume that the class of procedures for creating explanations and understanding includes not only *scientific procedures* but also *everyday procedures*; even though the latter category does not meet allthe requirements of the scientific method. Below are some examples of procedures that provide answers to important questions, but which do not meet all the requirements of the scientific method. I will concentrate on the following two interesting cases.

1. *Behavioral rules.*Much human behavior is explained by the fact that people have learned to obey behavioral rules of society. For example, the following answer to why Gideon stopped his car when the traffic light changed from green to red would be accepted as satisfactory: “Gideon obeyed traffic rules stating that drivers must stop their cars when the traffic light is red.” In these cases, several questions arise concerning the requirements of the scientific method. It is unclear how these rules, which become part of an individual’s inner world, act on that person’s nervous, muscular, and skeletal systems to enable publicly observable behavior, such as stopping a car at a red light. Furthermore, the execution of these rules ultimately depends on individual free will.
2. *Private mental states*.As described in the current and previous chapters, people use the mentalistic explanation model [Motivation–Belief] to explain behavior (however, see below for other kinds of explanations based on one’s inner world). That is, individuals use mental states to generate explanations for the behaviors of living creatures (humans and animals). For example, the answer to the question “Why did the postman throw the mail on the doorstep and sprint away from the yard?” might be: “Because the postman was afraid of the dog in the yard that had started to bark and snarl at him angrily.” Our understanding of the postman’s behavior, then, addresses the feelings of fear that the barking dog evoked in him. In other words, we explain public behavior by appealing to an individual’s inner mental states.

Here, I will briefly list accepted popular explanations belonging to folk psychology that are based on the inner world of the individual: [Motivation–Behavior] and other inner subjective explanations.

*(a) Explanation through [desire–belief]:* If Ruth *wishes* to watch the *Wonder Woman* movie and *believes* going to the Peer Cinema will fulfill that wish, Ruth will go to that cinema.

*(b)Explanation through emotion:* Extraordinary behavior is usually explained by strong emotions that overwhelm an individual. For example, David spent hours sitting in his car surveilling the home of his girlfriend, Ruth, out of jealousy, because he thought that Ruth was dating another man.

(c) *Explanation through inner schemas:* A lifelong New Yorker looked up at the sky and saw a cloud shaped like a car. In contrast, an Inuit man, who had lived all his life in the Arctic and had traveled to New York for a short visit, saw the same cloud and perceived the shape of a polar bear similar to those often encountered by men of his tribe. How do we explain this difference? The schema of a car had been established in the New Yorker’s mind because the city is full of cars. In the Inuit’s mind, the schema of a polar bear had formed because this animal is familiar to him.

(d) *Explanations from imagination, analogies, abstract ideas, and so forth*: Many children have imaginary friends with whom they talk and play. Imaginary friends alleviate loneliness and can help a child cope with conflicts and problems. Analogies are an important tool for understanding an individual’s behavior. For example, the question “Why was Dan behaving irrationally?” might receive the answer “Dan is not the brightest bulb in the box.” In other words, Dan is rather unintelligent and therefore behaves irrationally. Some social movements can be explained through ideologies (e.g., socialism, communism, democracy, and religious castes).

(e) *Explanation from the point of view of another person*: Consider the following example. The cloakroom attendant at a New Year’s Eve party hung Oren’s heavy coat on a small hanger on the right-hand side of the closet. However, after a while, she realized it would be better to hang the heavy coats on large hangers on the left-hand side. She moved the coat to the left and went home at around midnight. At the end of the party, in the early hours, where will Oren look for his coat—on the left- or right-hand side of the closet? The correct answer, of course, is based on Oren’s perspective and knowledge, not the reader’s (i.e., the right-hand side).

*Justifications for the MEF and TFTU*

Briefly, the main aim of this chapter is to propose novel empirical approaches to the following questions: (1) How are explanations for behavior provided by two explanatory procedures: mechanistic and mentalistic? (2) What theoretical approach explicates a person’s understanding of the explained issue? and (3) How does the answer to the second question depend on the answer to the first? The answer to the first question is given with the help of the MEF approach. It describes how mechanistic and mentalistic explanations are matched to the studied behavior and its components. Regarding the second and third questions, two general conditions must be fulfilled if an explanation is to be understood (a robot can give an explanation without understanding it). First, one has to be in a state of CΨ; without CΨ, there is no understanding. Second, one must use certain procedures to provide answers to particular questions. Understanding questions and answers connected to the sciences depends on the individual being in a state of CΨ when they perceive a mechanistic or mentalistic explanation (the latter being connected to everyday psychology, or folk psychology). To put it differently, understanding is provided by answers/explanations to questions that arise in a particular field in accordance with relevant explanatory procedures. This is a broad definition that includes two classes of procedures: class (a) includes answers, explanations, and understandings obtained through *scientific procedures* that meet the requirements of the scientific method; and class (b) includes answers, explanations, and understandings obtained through *everyday procedures* that do not fulfill all the requirements of the scientific method. As mentioned above, these two classes provide understanding when the answers and explanations to questions appear in CΨ. Consciousness is a necessary condition for understanding.

I will now discuss the main justifications for providing explanations by explanatory models, i.e., procedures. I will start with *scientific explanations*. Following the publication of Hempel and Oppenheim’s (1948) paper on the logic of scientific explanation and Hempel’s (1965, 1966) volumes on the philosophy of science and explanation, all models (procedures) are required to provide scientific explanations that meet the requirements of *rationality* and *empiricism*. Meeting these requirements is the main justification for creating new models (new procedures) to provide scientific explanations that offer answers to the questions of *why* and *how* (see discussion in Rakover, 2018). Rationality requires that the structure of an explanatory procedure does not present any internal logical contradictions. Empiricism requires that an explanatory procedure enables a clear and direct link to empirical observations.

Now for the justification of *everyday procedures* for forming explanations. These procedures do not meet every requirement of the scientific method. Behavioral rules and private mental processes do not fulfill the two requirements of rationality and empiricism in the way that *scientific procedures* do. Behavioral rules are not judged as right or wrong (they are not laws of nature) but are evaluated according to how they achieve their purpose. They are judged according to how effectively they regulate human behavior in a given society. Consider the following example: At road intersections, it is best practice to convert the traffic light system to a roundabout. Although roundabouts slow down the flow of traffic, they do not result in the long traffic jams that develop when traffic lights break down. Moreover, roundabouts reduce the number of road accidents that occur as a result of brake failure or not braking in time when the lights change from green to red.

Justifications for *everyday procedures* based on a person’s internal (mental) processes are particularly complex. While it is difficult to base these procedures on rationality and compliance with the requirements of scientific observation (public accessibility, objectivity, and repeatability; see Rakover, 1990), they may be grounded in actual everyday conscious experiences. However, these experiences raise questions that remain unanswered (see Rakover, 2018, 2021). From one perspective, each individual is aware of having an inner world that is unique in an ordinary, everyday way. A person is conscious of their desires, beliefs, thoughts, emotions, and feelings and is well aware that, in many cases, their behavior is merely the realization of this inner world. From another perspective, since each individual is aware only of their particular inner world, a person may doubt that an inner world similar to theirs might exist within another person—this is the well-known problem of other minds, which does not yet have an accepted solution (see Avramides, 2020). Despite this doubt, in everyday life, people conduct themselves out of a clear belief that others do indeed have an inner world that is very similar to their own. It is in these respects—that a person is aware that their inner world is the cause of their actions—that we can find justification for understanding human behavior.

To illustrate the tremendous problems inherent in justifying *everyday procedures*, consider the example of an explanation based on [Motivation–Belief]. On the one hand, it is difficult to suggest that such an explanation completely meets the requirement for rationality simply because an explanation based, for example, on desire and belief depends on an individual’s unique subjective perspective, that is, on their *desire* and their *belief* that a particular behavior will satisfy this desire. Moreover, as already mentioned, the reliance on an individual’s inner world creates problems related to making observations that meet scientific requirements: public accessibility, objectivity, and repeatability (see Rakover, 1990). On the other hand, referring to an individual’s desire and belief as a way of explaining their behavior is so prevalent in everyday life that it has become the most natural method for understanding behavior. This kind of explanation finds expression both in legal judgments (e.g., the motivation for the crime was *jealousy* or *hatred*) and in literature (e.g., the reason for the Trojan War was *jealousy* among the goddesses and a *passionate* *love affair* between Paris and Helen of Troy).

In addition, as discussed in Chapter 4 (and in Rakover, 2011/2012, 2018), it has been shown that an explanation based on the [Motivation–Belief] model fulfills several requirements of a scientifically acceptable procedure of explanation. These include, for example, generality and practical rationality, as well as the fact that the explanatory procedure itself is indifferent to the results of the empirical test.

Nevertheless, one should consider the following criticism. The above justifications can be interpreted, according to Hempel (1965), as reliant on a subjective psychological feeling that accompanies understanding (the so-called “aha moment”), while what is required is the evaluation of a model of explanation from an objective, logical, and methodological point of view. My response to this is as follows.

First, it would be hard to overlook that several scholars have raised important arguments against the idea that understanding is merely a psychological side effect of explanation (see de Regt et al., 2009). For example, it has been suggested that the number of explanatory components should be increased from two—the explanandum and the explanans—to three since the scientist should be added as an important component in explanation/understanding.

Second, as mentioned above, it is difficult to ignore that the [Motivation–Belief] explanatory procedure satisfies some of the methodological requirements of science (see Chapter 4 and Rakover, 2011/2012, 2018).

Finally, TFTU contradicts the conception of understanding as a psychological response, as merely a psychological side effect of explanation that accompanies explanation. According to TFTU, the understanding of an explanation itself requires consciousness. Without CΨ, there is no acquisition of understanding. If we do not accept the assumption about the necessity of CΨ, it follows that we have to accept the bizarre conclusion that Robbie the robot, understands both the questions put to him and the answers he gives, just as a human being understands them.

*Other approaches to understanding*

The accepted account of understanding in the philosophy of science was for many years based on psychological aspects of scientific explanation (inspired by Hempel and Oppenheim’s groundbreaking essay on the logic of explanation in 1948). The concept of explanation, not understanding, was the focus of research until the early 2000s. Since then, the notion of understanding has received increasing attention as a central concept in philosophy and psychology (see reviews by de Reget et al., 2009; Gordon, 2020; Horne et al., 2019; Khalifa, 2017; Pritchard et al., 2018; Rakover, 2018.)

Lipton (2009) proposed a very interesting idea that separates explanation from understanding: understanding without explanation. He proposed that understanding should be identified not with explanation but with cognitive benefits that are created by explanation, such as the apprehension that a particular event E is the cause of the studied phenomenon P or that E is a necessary condition for P. Based on this, Lipton argued that understanding can be reached in various ways, which differ from scientific explanations, such as via visual demonstrations of the occurrence of the studied phenomenon. (Lipton’s approach has given rise to much criticism that I will not dwell on here; see, for example, Khalifa, 2013, 2017; Strevens, 2013.)

As an example, Lipton pointed out that a physical model can be used to demonstrate the astronomical phenomenon known as apparent retrograde motion, according to which Saturn undergoes a very peculiar phenomenon: At first, it seems to move forward, then stops, moves backward, stops, and again moves forward. This perceptual illusion occurs because of the difference in speed between the Earth’s and Saturn’s orbits. As a result of the Earth’s greater speed, at certain angles of observation from our planet, Saturn appears to move forward, while at other angles, it appears to move backward.

The TFTU, which is rooted in Lipton’s main idea (separation between explanation and understanding), is based on the following: It is possible to achieve understanding through the use of appropriate methodologically accepted explanatory procedures; however, explanations can also be given through other means, e.g., using certain procedures that do not meet all of the accepted scientific methodologies. For simplicity’s sake (and see above), I will refer to the first type of procedures as *scientific procedures* (mechanistic explanations) and the second category as *everyday procedures* (mentalistic explanations). (On the distinction between the types of explanation used in the sciences and those used in the humanities, see, e.g., Grimm, 2016, 2019; Rakover, 1990, 2018.) The present theory, TFTU, differs in several aspects from Lipton’s (2009) approach to understanding without explanation. Lipton argued that “it is more natural to identify understanding with the cognitive benefits that an explanation provides rather than with the explanation itself” (p. 43). In this view, cognitive benefits include causes and necessary conditions. He continued: “[f]or by distinguishing explanations from the understanding they provide, we make room for the possibility that understanding may also arise in other ways” (p. 44).

One important difference between Lipton’s approach and TFTU is as follows. While Lipton limits himself to scientific explanatory procedures that answer the question of *why*, TFTU offers a wider approach to the concept of understanding, grounded in the differences between two types of procedures: *scientific procedures* and *everyday procedures*. The proposed theory, TFTU, shows that, in certain fields, there are justified procedures for reaching understanding that are not consistent with all scientific methodologies. These procedures are accepted within certain cultural traditions as explanatory. For example, as mentioned above, almost all explanations in literature use procedures that refer to an individual’s inner world as the cause of behavior. In Leo Tolstoy’s novel *Anna Karenina*, all of the troubles and disasters that befall Anna (and which, ultimately, lead her to take her own life) are ascribed to the powerful feelings of love that have gripped her and her lover, Vronsky.Lipton attempted to ground the possibility of providing understanding not in a particular scientific explanation but in identifying understanding with cognitive benefits. TFTU has expanded this approach. It shows that we may refer to two justified types of procedures for acquiring understanding: *scientific procedures* and *everyday procedures*.

Another important difference between Lipton’s approach and TFTU concerns the concept of CΨ. While in TFTU, CΨ is a necessary condition for understanding, it is not an important factor in Lipton’s approach. In contrast to Lipton, the concept of CΨ is a central axis in Bourget’s (2017) view. Bourget attempted to elucidate the idea of understanding via the concept of *grasping*, which is bounded by CΨ. According to Bourget, *grasping* a sentence with a certain content (a proposition) happens when it is experienced in a conscious phenomenological way. He analyzed several cases using this approach, including Jackson’s (1982) famous thought experiment about Mary, a scientist who is an expert in the neurophysiology of vision (see also Chapter 2). Jackson’s Mary knows everything there is to know on the subject of color. However, she has lived all her life in a black-and-white room. One day, Mary leaves her black-and-white environment and experiences the color red for the first time, and thus learns something new. Bourget argued that, until Mary leaves her room, she has no experience of redness and therefore has no grasp/understanding of its nature. However, when she experiences the color red, she grasps/understands it.

Bourget’s approach is similar to TFTU in considering CΨ to be a necessary condition for understanding. According to the former, it appears that Mary does not grasp/understand the color red until the stimulus appears in her CΨ. However, there is an important difference between these two approaches. While TFTU emphasizes the significance of scientific and everyday procedures for understanding, Bourget’s phenomenological approach does not address the need for an appropriate and justified procedure for arriving at explanation and understanding.

I will end this chapter with the following hypothetical situation: Let us assume that scientists have succeeded in fully explaining the phenomenon of CΨ. In this case, it could be suggested that there is no need for mentalistic explanations and understanding; we will not have to use *everyday procedures*—everything will be well explained by the mechanistic approach. Given this hypothetical situation, the question arises as to whether mentalistic explanations and *everyday procedures* would disappear entirely from scientific and everyday use.

This question is complex and beyond the scope of this current discussion since it belongs to the ongoing, unresolved debate in the philosophy of mind concerning eliminative materialism. This approach posits that the everyday theoretical conception of mental states, such as desire and belief, as well as explanations based on these concepts, are incorrect and have no place within the modern scientific approach; they will ultimately be superseded by the development of a suitable cognitive theory (see discussion in Ramsey, 2020). However, this is not the status quo. The scientific method has not yet succeeded in addressing the problem of CΨ. Eliminative materialism remains a subject of fierce debate in the professional literature and has not yet become the accepted, dominant approach. Therefore, I suggest that mentalistic explanations and *everyday procedures* continue to have an important role in explaining and understanding human behavior by other means, which do not satisfy all the requirements of the scientific method.

**References**

Avramides, A. (2020). Other minds. In E. N. Zalta (Ed.), *Stanford encyclopedia*

*of philosophy*. <https://plato.stanford.edu/archives/spr2020/entries/other-minds/>

Bechtel, W. (2008). *Mental mechanisms: Philosophical perspectives on cognitive*

*neuroscience*. London: Routledge.

Bourget, D. (2017). The role of consciousness in grasping and understanding.

*Philosophy and Phenomenological Research*, *95*, 285–318.

de Regt, H. W., Leonelli, S., & Eigner, K. (Eds.). (2009). *Scientific*

*understanding: Philosophical perspectives*. Pittsburgh: University of Pittsburgh Press.

Gordon, E. C. (2020). Understanding in epistemology. In M. A. Peters (Ed.),

*Encyclopedia of educational philosophy and theory*. https://doi.org/10.1007/978-981-287-532-7\_698-1

Grimm, S. R. (2011). Understanding. In S. Bernecker and D. Pritchard (Eds.),

*Routledge companion to epistemology* (pp. 84–94). New York: Routledge.

Grimm, S. R. (2016). How understanding people differs from understanding the

natural world. *Philosophical Issues*, *26*, 209–225.

Grimm, S. R. (2019). Understanding as an intellectual virtue. In H. Battaly

(Ed.), *The Routledge handbook of virtue epistemology* (pp. 340–351). New York: Routledge.

Hempel, C. G. (1965). *Aspects of scientific explanation and other essays in the*

*philosophy of science*. New York: The Free Press.

Hempel, C. G. (1966). *Philosophy of natural science*. Englewood Cliffs:

Prentice–Hall.

Hempel, C. G., & Oppenheim, P. (1948). Studies in the logic of explanation.

*Philosophy of Science*, *15*, 135–175.

Horne, Z., Muradoglu, M., & Cimpian, A. (2019). Explanation as a cognitive

process. *Trends in Cognitive Sciences*, *23*, 187–199.

Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly*, *32*, 127–

36.

Khalifa, K. (2013). The role of explanation in understanding. *British Journal for*

*the Philosophy of Science*, *64*, 161–187.

Khalifa, K. (2017). *Understanding, explanation, and scientific knowledge*.

Cambridge: Cambridge University Press.

Lipton, P. (2009). Understanding without explanation. In H. W. de Regt, S.

Leonelli, & K. Eigner (Eds.), *Scientific understanding: Philosophical perspectives* (pp. 43–63). Pittsburgh: University of Pittsburgh Press.

Pritchard, D., Turri, J., & Carter, J. (2018). Knowledge. In E. N. Zalta (Ed.),

*Stanford encyclopedia of philosophy*. https://plato.stanford.edu/archives/spr2018/entries/knowledge/

Pylyshyn, Z. W. (1984). *Computation and Cognition: Toward a Foundation for*

*Cognitive Science*. Cambridge, Mass.: MIT Press.

Rakover, S. S. (1990). *Metapsychology: Missing links in behavior, mind and*

*science*. New York: Paragon/Solomon.

Rakover, S. S. (2011/2012). A plea for methodological dualism and multi

-explanation framework in psychology. *Behavior and Philosophy*, *39/40*, 17–43.

Rakover, S. S. (2018). *How to explain behavior: A critical review and new*

*approach*. Lanham, Maryland: Lexington Books.

Rakover, S. S. (2020). Why has the field of psychology not developed like the

natural sciences? *The Journal of Mind and Behavior*, *41*, 247–266.

Rakover, S. S. (2021). *Understanding human conduct: The innate and acquired*

*meaning of life*. Lanham, Maryland: Lexington Books.

Ramsey, W. (2020). Eliminative materialism. In E. N. Zalta (Ed.), *Stanford*

*Encyclopedia of philosophy*. <https://plato.stanford.edu/archives/sum2020/entries/materialism-eliminative>

Strevens, M. (2013). No understanding without explanation. *Studies in History*

*and Philosophy of Science*, *44*, 510–515.