

CAUSALITY IN ISLAMIC PHILOSOPHY:
THE ARGUMENTS OF IBN SĪNĀ

Syamsuddin Arif

This article is intended to provide insight into aspects of Ibn Sīnā's natural philosophy. It will summarize his interpretation of the Aristotelian four causes, explicate his theory of efficient and necessary causal linkage, and analyze his arguments for causal efficacy. Finally, it will discuss Ibn Sīnā's views on chance happenings in nature.

Keywords: Ibn Sīnā; natural philosophy; cause-effect relation; causality; causal efficacy; chance and luck; determinism.

Few ideas in the history of philosophy in Islam have been so much debated, attacked, and defended as has the thesis that a necessary connection exists between cause and effect, that cause and effect are so inextricably linked that the existence of one necessitates and implies that of the other, and that if the cause has occurred, the effect cannot fail to occur. For instance, it is this conception of causality, with its hidden assumptions and far-reaching ramifications (e.g. the fettering of God's Will, the eternity of the world, the denial of the possibility of miracles, etc.), that prompted al-Ghazālī (d. 505/1111) to charge Ibn Sīnā with heresy and infidelity. This article reexamines Ibn Sīnā's views on this crucial issue by looking first into his exposition of the four causes and his account of causal efficacy and necessity, before finally discussing his theory of chance.

Syamsuddin Arif is Assistant Professor, Department of General Studies, KIRKHS, International Islamic University Malaysia; Email: tagesauge@yahoo.com.

Islam & Science, Vol. 7 (Summer 2009) No. 1

© 2009 by the Center for Islam and Science
ISSN 1703-7603 (Print); ISSN 1703-7602X (Online)

1. The Four Causes

Ibn Sīnā, like Aristotle,¹ recognizes four kinds of causes, or rather four modes of explanation in the study of nature: (1) the efficient cause (*'illah fā'iliyyah*), namely, that which causes a thing to move or change and hence is dubbed the source or principle of motion and change (*mabda' al-ḥarakah*); (2) the material cause (*'illah 'unṣuriyyah*), defined as that out of which a thing is generated, that in which change or motion is produced, and that which persists before, during, and after the process of change; (3) the formal cause (*'illah ṣūriyyah*), that is, that into which something is changed or moved; this is the essence or 'what it was for so-and-so to be' of a thing; and (4) the final cause (*'illah ghā'iyyah*), or that for the sake of which a change or motion is produced.

Ibn Sīnā takes these four causes as explanatory principles which are indispensable as a theoretical framework for the successful investigation of the natural world. This four-cause method of analysis is applied by him to natural bodies: everything that is generated and destructible, everything subject to change (*li-kulli wāqi' fī al-ḥarakah*), and everything composed of matter and form. For example, he says, a bed is produced by a carpenter (its efficient cause) by his imposing changes upon a block of wood (its material cause) for the purpose of possessing a piece of furniture on which to sleep (its final cause), the wood thereby acquiring the form, or certain distinctive properties, of a bed (its formal cause).²

While he admits that the ancient philosophers utilized all these four explanatory principles, Ibn Sīnā criticizes them for emphasizing some of the causes to the neglect of other equally, perhaps even more important factors in explanation. Thus he laments that some of the early thinkers were concerned too much with material causes, believing they could explain the natural world by discovering the basic matter out of which all things are made, such as water (for Thales), air (according to Anaximenes), or fire (in Heraclitus' view). On the other hand, others, like Plato, often spoke as if the explanation of all things would be achieved simply by discovering their Forms, or formal causes. The concept of efficient cause did not emerge until the time of Empedocles, who postulated two motive forces, Love (*maḥabbah*) and Strife (*ghalabah*), to account for the various changes observed everywhere in the universe.³

According to Ibn Sīnā, the material cause is important inasmuch as it is the potential bearer (*ḥāmil*) of form and insofar as it is the locus of change from potentiality to actuality. Every natural thing has a material cause that is subject to generation, be it through the change of attribute or

state (*istiḥālah*), by way of aggregation and combination (*bi-ḥasab ijtīmā' wa tarkīb*), or both by way of aggregation as well as change of state.⁴

The formal cause is understood by Ibn Sīnā as having a variety of meanings. This is because he says the term 'form' (*ṣūrah*) can stand for: (1) essence (*māhiyyah*), which, when present in matter, transforms it into a species (*naw'*); (2) the species itself; (3) figure or shape (*shakl*); and (4) the reality (*ḥaqīqah*) of a thing. As for the final cause, Ibn Sīnā defines it as the purpose for which the form is produced in matter, the purpose being either the Real Good (*khayr ḥaqīqī*) or a supposed good.⁵

Ibn Sīnā points out that the formal, efficient, and final causes may coincide in one thing. Consider, he says, the case of a human being: one is at the same time the efficient, formal and final cause of his child. In this case, though conceptually separable, the efficient cause (a man), the formal cause (manhood), and the final cause (to become a man) are collectively embodied in the parent alone.⁶ It may also happen, Ibn Sīnā adds without giving any example, that the four causes coalesce in one thing (*wa qad yattafiqu an yajtami'a hādhihi al-asbāb kulluhā li-shay' wāhid bi al-dhāt*), or that a thing has only two causal factors, the efficient and the final, as in the case of the separated Intelligences.⁷

Explaining the relation between the causes, Ibn Sīnā asserts that in a sense the efficient cause or agent is a cause for the purpose (*sabab li al-ghāyah*), since it makes the purpose exist (*yuḥaṣṣil al-ghāyah mawjūdah*). In yet another sense, the purpose is a cause for the agent, since it is precisely for its sake that the agent does what it does, and without which the agent would not act. This is why if someone asks: "Why do you exercise?", the correct reply would be: "To be healthy." Similarly, if someone asks: "Why is it that you always look healthy?", it would not be incorrect to say in reply: "Well, because I always exercise." In that case, one is therefore the cause for the other. Ibn Sīnā remarks, however, that this by no means implies that an agent is the cause for the being-purpose of a purpose (*al-fā'il laysa 'illatan li-ṣayrūrat al-ghāyah ghāyah*), nor does it mean that an agent is a cause for the essence (*māhiyyah*) of the purpose in itself. Rather, an agent is but a cause for the actual existence of the purpose's essence in concrete things, since for Ibn Sīnā the essence of a thing is distinct from its existence. By contrast, the purpose is a cause for the being-agent of an agent (*al-ghāyah 'illah li-kawn al-fā'il fā'ilan*), a cause for its being a cause.⁸

Each of these causes can be either essential (*bi'l-dhāt*) or accidental (*bi'l-ʿarad*), proximate (*qarīb*) or remote (*ba'īd*), specific (*khāṣṣ*) or general (*ʿāmm*), particular (*juz'ī*) or universal (*kullī*), simple (*basīl*) or composite

(*murakkab*), potential (*bi al-quwwah*) or actual (*bi al-fi'l*). An efficient cause is said to be essential if it acts as a principle for the essence of its act (*mabda' li-dhāti dhālika al-fi'l*), as in the case of a physician when he gives treatment, or fire when it heats water. The accidental efficient cause may be any of the following: (1) that whose act eliminates the contrary which prevents its contrary and strengthens the other contrary, e.g. scammony,⁹ which, when cooled, can ease yellow bile; (2) that whose act removes the obstacle that prevents a thing from acting naturally, such as anything that destroys the end or purpose of something by eliminating its supporting condition; (3) that whose act does not destroy its individuality, e.g. our saying that the doctor builds, by which we mean that being-a-builder is a subject for him, so that he builds not insofar as he is a doctor, but insofar as he is a builder; (4) that whose natural or voluntary act is directed toward a certain end, whether achieved or not, like a falling stone that breaks upon hitting a reptile; or (5) that which in fact does not act at all, as in the case of chance happenings.¹⁰ The efficient and final causes are usually remote, while the formal and material causes are proximate. The proximate causes act without an intermediary, as in the case of motion of a part of the body by the muscles. By contrast, remote causes act by means of intermediaries, such as the motion of the parts of the body by the soul.¹¹

In natural philosophy, says Ibn Sīnā, by 'efficient cause' or 'agent' (*al-fā'il*) is meant "the principle of change (*mabda' al-ḥarakah*) in something other than itself," the term "change" here referring to "every passage in matter from potentiality to actuality" (*kull khurūj min quwwah ilā fi'l fī māddah*). It is the principle which acts as a cause (*sabab*) or agent of change in natural things from one state to another. To anticipate a possible objection to this definition, Ibn Sīnā hastens to emphasize the difference of cause and effect, adding the phrase "by way of its [i.e., the effect's] being other [than the cause]" (*min ḥaythu huwa ākhar*). Thus, for example, even the physician curing himself is also the principle of change 'in another by being another' (*fī ākhar bi-annahū ākhar*)—for he is changing himself (by undergoing treatment) not insofar as he is a physician, but insofar as he is a patient.¹² Ibn Sīnā further distinguishes two aspects of efficient cause as the agent or principle of change. First, the principle or agent of change may be regarded as 'that which prepares' (*muhayyi'*) and adjusts or modifies the matter of a thing so as to make it suitable for receiving a particular form—for example, that which causes different changes in sperm. Second, an agent of change or efficient cause may be seen as 'that which brings to completion' (*mutammim*) the whole change by giving to

each thing the form it deserves.¹³

2. Efficient Causality

Unlike Aristotle, however, Ibn Sīnā classifies the four causes into the “essential” and the “ontological”. By ‘essential causes’ (*‘illatā al-māhiyyah*) he means those which constitute the essence of a thing—namely, its material and formal causes, whereas by ‘ontological causes’ (*‘illatā al-wujūd*) he means those which bring the thing into actual existence—namely, its efficient and final causes.¹⁴ This holds true of natural things whose essences are distinct from their existences, in contrast to the Necessary Being (i.e., God) whose essence is identical with His existence. Ibn Sīnā also goes beyond the ‘First Teacher’ by expanding the concept of efficient cause in order to fit his own metaphysical system. Whereas for Aristotle the efficient cause refers to that which produces motion or change in something already existing (since for him coming to be and passing away is always from and into something, and never *ex nihilo*), Ibn Sīnā gives a deeper and far more fundamental meaning of efficient cause, taking it as the cause of existence. As he put it in the *Physics* of his *Shifā’*:

If, however, the efficient cause is taken, not in relation to natural things, but in relation to existence itself (*bi-ḥasab al-wujūd nafsihi*), it would then constitute a meaning more general than this [i.e. as used in natural philosophy]. Consequently, everything which is a cause of an existence different from itself becomes—inasmuch as it is different and inasmuch as the new existence is not for the sake of the former—an efficient cause.¹⁵

He points out that whereas in natural philosophy the efficient cause or principle common to natural things is identified as “nature” (*al-mabda’ al-fā’ilī al-mushtarak li al-ṭabī‘iyyāt huwa al-ṭabī‘ah*),¹⁶ which in the case of living beings or animals is called ‘soul’ (*nafs*), in metaphysics the efficient cause is posited as the principal source of existence (*mabda’ al-wujūd*). And whereas the role of the efficient cause in natural philosophy is confined to that of producing change in things and to being the principle or cause of motion and rest, for the metaphysicians like Ibn Sīnā the efficient cause is the one responsible for producing existence and maintaining it, apart from causing change and motion. As he explains, by the term ‘agent’ (*fā’il*) he means the cause which bestows existence on things, as something ‘super-added’, as it were, to the thing’s essence (*allatī tufīdu wujūdan mubāyinan li-dhātihā*): “The metaphysicians (*al-falāsifah al-ilāhiyyūn*) do not mean by ‘agent’ only ‘the principle of motion or change’ (*mabda’ al-taḥrīk*), as do the

natural philosophers, but rather they mean ‘the principle and bestower of existence’ (*mabda’ al-wujūd wa mufīduhu*) as in the case of God with respect to the universe.”¹⁷

3. Two Principles of Causality and Their Corollaries

At the heart of Ibn Sīnā’s theory of causality therefore lies the notion of bringing into existence (*ijād*). For him, the efficient cause does not merely produce change and motion in its effect; it brings the latter into existence, renders it actual, putting an end to the latter’s potentiality. Indeed, Ibn Sīnā weaves his physical theories into his metaphysics, in order to lay a solid rational foundation for religion. It is on the basis of his assertions about the necessary connection between cause and effect that he develops his proof for God’s existence and the eternity of the world. We shall now proceed to extract the main principles of causality from Ibn Sīnā’s philosophy.

As we know, Ibn Sīnā classifies things “considered in terms of what they are in themselves (*idhā u’tubira bi-dhātihi*)” into three, namely: (1) that whose existence is necessary; (2) that whose existence is merely possible or contingent; and (3) that whose existence is totally impossible. The last he rules out because the impossible cannot exist without no longer being impossible. Left with the other two categories, he advances that if anything *is* (literally: if anything enters existence (*tadkhul fī al-wujūd*) then it would have to be either something whose existence is in itself necessary or in itself is only possible. Now, (1) if its existence is necessary in itself, then such a being cannot be caused, for otherwise it would be necessary through something else and not by itself. According to Ibn Sīnā, such a being (whose existence is necessary by itself and independent of any cause) would have to be numerically one, the only such being, simple and changeless. (2) On the other hand, if the thing’s existence is in itself only possible, then, in terms of its own essence, it may and may not exist. For such a being is in need of some cause, whether to be or not to be (*fa-wujūduhu wa ‘adamuhu kilāhumā bi-‘illah*), and hence is contingent. If such a contingent entity exists, then it must have been specified with existence rather than non-existence. Similarly, if it does not exist, this is because it has been specified with non-existence rather than existence—for the contingent’s own essence alone cannot account for this specification (*takhṣīs*). Thus, if the contingent exists, its existence is due to something extraneous which is its cause. By the same token, if it does not exist, then this is also due to something extraneous—in this case, the absence of the cause.¹⁸ With this

ontological analysis Ibn Sīnā establishes the *first principle* of his causal theory, namely that everything contingent, if it ever exists, must have a cause and must be caused to exist by something other than itself.

The *second principle* of causality enunciated by Ibn Sīnā states that everything contingent that is caused to exist is caused necessarily—that is, its existence is necessitated. His argument for this may be paraphrased as follows. Any contingent being that exists, though in itself is possible, must become necessary through its cause and with respect to it (*yaṣīru wājiban bi al-‘illah wa bi al-qiyās ilayhā*). For otherwise—that is, if it remained possible with respect to its cause—then the fact of its existence rather than its non-existence cannot be explained. We would then have to suppose another cause and if again, in relation to it, it remains possible, then we would have to suppose yet another cause and so on *ad infinitum*. Even if we allow the possibility of the infinity of such causes, this will not explain the fact of the contingent’s existence; for it would remain possible, not only in terms of its own essence, but also in relation to all these infinite causes. However, given the assumption that the contingent exists, its existence can only be explained in terms of a cause that renders it necessary. Thus it becomes necessary through the cause and with respect to it.¹⁹

According to the second principle, therefore, if the cause exists, all other conditions being fulfilled, the effect must exist necessarily. This means that, on Ibn Sīnā’s view, the connection which exists between (an efficient) cause and its effect is a necessary one. To say that the cause entails (*lazima ‘anhu*) and necessitates (*wajaba ‘anhu*) its effect is just another way of saying that the effect necessarily proceeds from the cause; that the effect must happen in case the cause exists; that once the cause exists—all other things being equal, that is, in normal circumstances—its characteristic effect must occur; and that the effect cannot fail to occur simply because it is compelled to occur by the very nature and power of the cause. If somehow the anticipated effect does not in fact occur, this does not mean that there is no such thing as necessary causal connection between the cause and its effect. Rather, it merely indicates that not all things were equal as had been supposed—that the condition *ceteris paribus* was not satisfied—as, for example, the cause may have been absent, misdescribed, or impeded by the efficacy of other causes.

It must be noted that when Ibn Sīnā speaks of efficient cause as necessitating the existence or occurrence of its effect, he is referring to the ‘essential’ and ‘proximate’ cause (*‘ilal dhātīyyah wa qaribah*), which alone are the ‘true cause’ (*‘ilal bi al-ḥaqīqah* or *‘ilal ḥaqīqīyyah*), in contrast to the

accidental cause (*ʿilal bi al-ʿarad*) or the auxiliary and preparatory cause (*ʿilal muʿīnah wa muʿiddah*). He is insistent that the essential, true cause, and it alone, necessitates the occurrence of its effect and hence coexists with it, whereas the other two kinds of cause do not.²⁰ This is implied in the phrase *ʿbi al-dhāt* from which the adjective *ʿdhātīyyah* is derived. When applied to the term ‘cause,’ the phrase has two senses: first, the general one referring to the very thing which becomes the cause, and second, the specific sense that refers to the essence or nature of the agent thing, where the action is a consequence or effect of that essence. It is important to note further that, contrary to the widely held view, Ibn Sīnā nonetheless acknowledges that empirical observation of regularities in nature does not prove the existence of necessary causal connection between natural things or events, and that regularities merely show that one is concomitant with the other. However, he argues, there is a hidden syllogism to the effect that if this regularity were accidental or coincidental, it would not have continued to exist or occur ‘always or for the most part’ (*dāʿīman aw akthariyyan*). From this he concludes that the regularity is indeed essential and due to the natural power inherent in the cause.²¹

Several corollaries follow from the two principles mentioned above. The first of these is that the (true, essential efficient) cause and its effect are coexistent (*al-ʿilal al-ḥaqīqīyyah mawjūdah maʿa al-maʿlūl*), which means that the relationship between the cause and its effect is not only coextensive in the sense that what holds true of one is equally true of the other, but also reciprocal in the sense that the existence of one can be inferred from and implies the existence of the other (*al-ʿillah musāwīyah li al-maʿlūl munʿakisah ʿalayhi*). To put it in general terms, when the proper causal conditions obtain and no impediment intervenes, an agent or efficient cause not only produces its proper effect, but logically coexists with it. This holds true, however, only in cases where the cause is unique in relation to its effect (*idhā kāna al-maʿlūl ʿillatuhu wāḥidah*)—that is, when it alone produces that particular effect.²² For when the same effect can be produced by any of several efficient causes or agents, the relation is no longer reciprocal, since no single cause can be inferred from the effect. The reciprocity relation is not restored until one element, common to the variety of possible causes, is identified and treated as the proximate cause or agent.²³ Of course, this same reciprocity relationship between cause and effect may also be severed or interrupted by the intervention of various impediments. Yet according to Ibn Sīnā, such exceptions to the unique and invariable character of causal necessity are mostly restricted

to a single domain, namely the world of generation and corruption (*‘ālam al-kawn wa al-fasād*). For in this realm the different natures of things result in frequent intersections between lines of causal influence so that the efficacy of some natural causes may be impeded by that of others, while non-natural causes, such as the human deliberative faculty, do not invariably produce one kind of effect in any case.²⁴ Each of these counter-instances, therefore, represents for Ibn Sīnā the exception that proves the rule, for even the impediment impedes in keeping with its essential nature under appropriate causal conditions, while the deliberations and choices of human beings are likewise consistent with their own natures.

Ibn Sīnā does not say that the cause necessitates the occurrence of its effect without qualification. As he explains, the necessary relation between efficient causes and their effects may not obtain if the conditions peculiar to their natures are not met. In the sublunar, terrestrial world of generation and corruption, for instance, no less than six requirements must be fulfilled in order for effects to occur: (1) the (efficient) cause must be both proximate and specific; (2) the cause must represent the only kind of cause productive of a specific effect; (3) the cause must be essential as well as actual, while the effect must be in itself merely potential;²⁵ (4) the cause must relate to the essence or nature of the (natural, involuntary) agent; (5) the recipient of the action must exist and must be properly disposed for the reception of the act;²⁶ and finally (6) there must be no impediments.²⁷ Of course, all these conditions are not always present. One should also consider the fact that the conditions of actuality ('being-actual') and necessity ('being-necessary') do not always combine. For example, the efficient cause may be a necessary condition (*mimmā yūjīdu dhātan*), but may not yet be actual (*wa laysa ba‘du ‘illatan bi al-ḥi‘l*), such as the wall being a necessary condition for the existence of the ceiling, but may not yet be actual.²⁸ Also, a cause may be a sufficient (*tāmm*) but not necessary (*mūjīb*) condition for the existence or occurrence of the effect; for the same effect may be produced by any one of a number of causes.²⁹

Another corollary is that the true efficient causes are always prior to their effects, ontologically if not temporally (since the two things or events may be contemporaneous or simultaneous, coexisting at the same time). That a cause must precede its effect 'in existence' is explicitly stated by Ibn Sīnā in the *Metaphysics* of his *Shifā'*:

Thus, the existence of every effect is necessary in conjunction with the existence of its cause, while the existence of its cause necessitates the existence of the effect from it (*wujūdu kullī ma‘lūl wājib ma‘a wujūdi ‘illatihi, wa wujūd ‘illatihi wājib ‘anhu*

wujūd al-mā'lūl). Both occur together in time or in eternity [that is, temporally or eternally] or what have you, but not in relation to the attainment of existence [i.e., ontologically the two are not on a par]. This is because the existence of the one [i.e., the cause] did not come about from the existence of the other [i.e., the effect], for the occurrence of existence (*huṣūl wujūd*) belonging to the one does not derive from the occurrence of existence belonging to the other, whereas the occurrence of existence of this [effect] derives from the occurrence of existence of that [cause]. Therefore, one is prior [to that of the other] in relation to the attainment of existence.³⁰

Ibn Sīnā distinguishes between two kinds of ontological priority. In one case the prior [i.e., the cause] is a necessary condition for the existence of the posterior [i.e., the effect], but does not necessitate the existence of the latter. For example, the number one is necessary for the existence of subsequent whole numbers but hardly necessitates that there are two, three, etc. of whatever we should like to count. In the second case, what is prior is both a necessary condition for the existence of the posterior and a factor which necessitates the latter's existence as well. This kind of priority is represented by essential efficient causes as Ibn Sīnā conceives them.³¹ On this view, an essential efficient cause must be regarded as prior to its effect even though they are temporally coexistent. That is to say, the cause must exist prior to its effect; it must be there if the effect is to occur at all. This also means that while it is true that causes are said to necessitate the occurrence of their effects, the reverse is not true—there is no 'backward causation.' Of course, causes may either precede or coexist with their effects temporally; they may either occur successively (one preceding the other) or simultaneously (neither occurring before the other), in which case they are contemporaneous in the sense that there is never any real temporal succession or sequence such that one could distinguish between two things or events, which being the cause and which the effect, merely by asking which occurs first in time, since neither of them occurs first. Consider, for instance, the relationship between one's hand and the key one is holding. The key will move whenever the hand moves, the movement of the key being caused by the hand's movement, and not vice versa. There seems to be no temporal gap between the two movements—they move simultaneously, the movement of neither being followed by the movement of the other, such that it would be clearly false to say that the hand moves and then, after it stops moving, the key begins to move. Describing this second type of ontological priority, Ibn Sīnā says:

The mind is not repelled at all by our saying, “when Zayd moves his hand, the key moves,” or when we say, “Zayd moves his hand, then the key moves.” But it is repelled by our saying, “when the key moves, Zayd moves his hand,” even though the mind rightly says, “when the key moves, we know that Zayd moves his hand.” Thus, with respect to the temporal coexistence of the two movements, the mind assigns a priority to one and a posteriority to the other. For it is not the existence of the second movement that causes the existence of the first, but it is the first movement that causes the existence of the second.³²

That the cause is ontologically prior to its effect further means that the latter depends upon it for existence and is necessitated by it. Against his own assertion, however, Ibn Sīnā raises a possible objection: It might be argued that *if* each one of two things is such that if one exists, the other exists, and if one is removed, the other is removed, *then* one is not the cause and the other is not the effect, since neither one of them has a greater claim (*awlā*) than the other to be the cause with respect to existence.³³ It is obvious that the critic here is assumed as having taken Ibn Sīnā at his word on the criteria of simultaneity and reciprocity in the implication of existence, which Ibn Sīnā has asserted to exist between essential efficient causes and effects.³⁴ But the critic uses them to deny that the so-called cause is ontologically prior to the effect. Neither one can have a greater claim on existence than the other, if they both arise together and disappear together. Indeed, it is so argued, their simultaneity connotes temporal coexistence, rather than priority of one to the other. Similarly, their mutual reciprocity connotes only an ontological association, so that wherever they are found, they are found together. But this does not justify any concept of ontological priority. While the opponent here is not identified, the character of his objection strongly suggests that Ibn Sīnā is dealing with an occasionalist. For in attacking the notion of ontological priority in the second sense (which is both necessary and necessitating), Ibn Sīnā clearly indicates his rejection of essential efficient causation itself.

Ibn Sīnā responds with a conceptual analysis of the various ways in which the “if” mentioned in the objection may be understood. The obvious aim behind this procedure is to answer the objection itself by clarifying the precise sense in which the coexistent cause is prior to the effect. To this end Ibn Sīnā distinguishes between four senses in which the coexistence of cause and effect can be understood. The statement “if one exists, then the other must exist,” he argues, can mean either of the following: (1) if one (either cause or effect) has existed in external reali-

ty (*idhā ḥaṣala fī al-wujūd*), then the other must consequently come into existence (*yajibu ‘anhu an yaḥṣula*) in external reality; (2) if one has existed in external reality, then the other must have already existed (*qad ḥaṣala*) in external reality; (3) if the existence of one has occurred in the mind, then the existence of the other must consequently occur to the mind; or (4) if the existence of one has occurred, then the existence of the other must have already occurred either in external reality or in the mind.³⁵

The first alternative is rejected, since it is true with regard to the cause, but false as regards the effect. For it is only the cause which is such that if it occurs then the occurrence of the other—after being possible—follows necessarily from it, whereas the occurrence of the effect does not necessitate the occurrence of the cause. The second alternative is also rejected, since it is false both with regard to the cause and the effect. For as far as the cause is concerned, it is not true that if the cause exists, then the effect would have come into existence by itself or without the cause. Nor is it true as far the effect is concerned, because (i) the actual existence of the cause in external reality is not a necessary consequence of the effect, and (ii) because the actual existence of something which has already occurred cannot be necessitated by the occurrence of something supposed to be occurring—unless the word “occur” is understood otherwise. Thus only the third alternative is entirely correct, since it is possible to say that “if the cause exists in the mind, it is logically necessary that the effect of which the former is its essential cause should exist in the mind,” and likewise “if the effect exists in the mind, then the cause too must exist in the mind.” As for the fourth alternative, only the first part of it (namely, if the existence of one has occurred, then the existence of the other must have already occurred in external reality) is true, while the other (i.e., if the existence of one has occurred, then the existence of the other must have already occurred in the mind) is false. It is true that if the effect exists, the mind testifies (*shahida al-‘aql*) that the existence of the cause has occurred in external reality so that the effect has come to exist, although the existence of the cause might occur in the mind after that of the effect.³⁶ From the correct alternative it follows that if the existence of either cause or effect in reality presents itself to the mind, the mind attests to or makes a necessary inference to the existence of the other, although the other may not be known or represented in the mind until long afterward. Speaking in psychological terms, his point is that in thinking of one we must ultimately think about the other, although he offers no reason why we should do so.

In short, to say that the cause is ontologically prior to the effect simply means that the existence of the cause is necessary for the existence of the effect, and that the former necessarily results in the latter. Much of the same set of arguments applies to the removal of cause and effect from existence. On Ibn Sīnā's account, the removal of the effect is dependent upon and necessitated by the removal of the cause, but not vice versa. At most one could conclude from the removal of the effect that the cause has necessarily ceased to exist first. From these considerations, Ibn Sīnā draws his conclusion as follows:

Let us return to the point where we drew this distinction, for we say in regard to the state of perplexity [which prompted the question] that it is not the conjunction (*ma'yyah*) [of cause and effect] that necessitated causation for one of the two, so that it is [argued that] "one of them does not have a greater claim on causation than the other because, in terms of conjunction, they are on a par." Rather, the two of them differ because we suppose in the case of one of them that its existence does *not* necessarily occur *through* the other, but *with* the other, while in the case of the second we suppose that just as its existence occurs with the existence of the other, it likewise exists through the other. Thus the question should be resolved in this way.³⁷

Two noteworthy points emerge from Ibn Sīnā's analysis. First, his defense of ontological priority as both necessary and necessitating depicts causes and effects as essentially discrete and rather static entities, which are related in various ways depending upon the sense and context assigned to each. There is nothing intrinsically active about the cause which is necessary for a given effect and from which the effect necessarily follows. Like two links in a chain, we may say that they are connected but, to all appearances, the links do not do anything. Second, Ibn Sīnā does not clearly indicate whether he regards causes primarily as substances or as events or equally as both. His examples tend to include both. Although he seems inclined to regard causes as substantive agents, here the clear illustration he gives of a cause and its effect in discussing ontological priority—that of the movement of someone's hand in turning a key—portrays the cause as an event (*ḥādīthah*). This ambiguity cannot be ignored, since it is precisely these elements in Ibn Sīnā's account of causality that the theologians like al-Ghazālī, who regard causes and effects strictly as events, have exploited to destroy and condemn his metaphysical system.³⁸

4. Chance and Luck

We now come to the opposite of causal necessity, namely chance (*ittifāq*)

and luck (*bakht*). Ibn Sinā begins by reviewing and refuting the opinions of ancient thinkers on these issues. First, he tells us, there are some who reject the whole idea of chance and fortune, those who claim that every event has a definable cause and that everything that seems to occur by chance is really due to a determinate cause. According to them, only the fool would think it is by ‘good luck’ (*al-bakht al-sa’id*) that someone found a treasure while digging a hole, much as it would be due to ‘bad luck’ (*al-bakht al-shaqīy*) if he slipped and broke his leg. There is no such thing as luck, they argue. Thus if someone found a treasure it is because he had dug for it. Similarly, if, while digging for it, the person slipped, it is because he did not mind his step. Also, they further argue, when someone goes to the market for some purpose of his own and meets his debtor and then gets back his money, this cannot be due to luck either, but rather because he intentionally went to the place where his debtor was around, spotted him, and so was able to collect his money. Admittedly, they say, his going out to the market was not planned for the purpose of getting the money back, although it eventually had the same result as if so designed. For it is possible that an activity or event have several goals (*yajūz an yakūn li-fi’l wāḥid ghāyāt shattā*), but when either one of them is achieved, the rest are ignored.³⁹

Others have taken the totally opposite view, overestimating the role of chance and luck. Those who hold that luck is a divine cause (*sabab ilāhīy*) and a hidden, mysterious power inscrutable to human intelligence have gone so far as to worship it like a deity, built it a temple and erect a statue dedicated to it. There are others who, like Democritus and his followers, regard luck as the most prevalent of all natural causes and hold that the whole universe is a spontaneous product of chance. They say that the fundamental constituents of everything are unsplitable corpuscles (*ajrām ṣiḡhār lā tatajazza’*). These so-called atoms or minimal parts are said to be numerically infinite, dispersed in an infinite void, having various shape and size, continuously in motion, and colliding with each other as they roam in the void, giving rise as luck would have it to the cosmic vortex out of which the universe came into being in its present order. They also hold that there are infinitely many universes similar to this one we live in. Finally, there are those who despite their denial of chance as the cause of the whole universe nevertheless maintain that natural things which are made of basic elements are produced or come to be by chance (*al-kā’ināt mutakawwimah bi al-iltifāq*). Even continuation and reproduction of species are said to occur by chance.⁴⁰

Ibn Sīnā next gives his own views, stating that things and events may be classified into: (1) ‘those which happen always’ or regularly (*mā hiya dāʾimah* or *al-dāʾim*); (2) ‘those which happen for the most part’ or usually (*mā hiya fī akthar al-amr* or *al-aktharī*); (3) ‘those which seldom happen’ or rarely (*mā yakūnu fī aqall al-amr* or *al-aqallī*); and (4) ‘those which happen as often as not’ or occasionally (*mā yakūnu bi al-tasāwī* or *al-mutasāwī*). An example of the first category is the occurrence of burning when fire comes into contact with firewood, while events of the second category are such as the expected arrival of someone heading for some place at his destination upon leaving the house. Things or events which happen always or usually, Ibn Sīnā tells us, do so either because the nature of the cause acts regularly (*ʿan iṭṭirād fī ṭabīʿat al-sabab*)—the act being necessitated by the very nature of the cause or agent—or because the cause needs no accompanying factor (*qarīn aw sharīk*), that is, when it is sufficient to produce the effect, or because there is no impediment. Events of the first category are necessary (*wājib*) by virtue of the nature of the cause. But those belonging to the second category too, we are told, can become necessary “provided all impediments present are removed and potential obstruction prevented” (*bi sharṭ dafʿ al-mawānīʿ wa imāḥat al-ʿawāriḍ*). Ibn Sīnā then concludes that things or events which always or usually occur cannot be due to chance or luck.⁴¹

The last two categories to be considered relate to the things which happen as often as not or occasionally (lit. ‘equally’ = *biʿl-tasāwī*) and those which seldom occur. It is a problem whether or not what occurs equally is said to be so by chance. Ibn Sīnā disagrees with the later peripatetics who said that the things which occur by chance are found among the things which occur rarely, and that the things which occur as often as not, like walking and not walking, eating and not eating, result from the thing’s principles, that is, from someone’s choice and decision to perform that action or not, and this cannot be called by chance. According to Ibn Sīnā, however, it is possible that something may occur necessarily or usually (*akthariyyan bal wājiban*) from one point of view, and that it may occur as often as not or rarely from another viewpoint. For instance, the generation of a sixth finger seldom occurs, but it is necessary from the standpoint of nature as a whole in which no matter should get lost, so that it is necessary that it receives the form it was prepared for. Therefore it is possible that something occurs as often as not from one point of view and usually from another point of view. If walking or eating is attributed to the will which decides whether or not this activity will take place, then it

usually occurs. But if it is regarded not as such but only insofar as it may or may not occur, then it may be said to occur as often as not and by chance. Thus it is correct to say: “When I entered his house, he just happened to be eating,” eating in that case being considered not from the point of view of the choice of the eating person, but from the point of view of the entering person, for whom it is by chance that the person in the house is eating (*bi al-qiyās ilā al-dukhūl lā ilā al-irādah*).⁴²

Ibn Sīnā then concludes that something is said to happen by chance only if it is something which is unexpected (*ghayr mutawaqqaʿ*) and which neither always nor usually happens in the same way, provided it is chance or luck that leads to the thing’s occurrence (*idhā kāna min shaʿnihi an yuʿaddiya ilayhi*). Therefore, chance and luck are regarded as accidental causes. Consider, for example, the case of someone sitting at the moment of lunar eclipse. We cannot say that his sitting happens to be the cause of the eclipse, although it is correct to say that both events happen to coincide. The activity of sitting was not the (true, essential) cause of the eclipse, but rather was merely an accidental cause leading to the coincidence or ‘happening together’ with the eclipse (*al-kawn maʿa al-kusūf*), for ‘occurring or being together with the eclipse’ is not an eclipse.⁴³ To sum up, what happens by chance must therefore be something which occurs for the sake of something (*min ajli shayʿ*), whether it be according to choice or nature. Chance causes (*asbāb ʿitifaʿiyyah*) are accidental; they can never necessitate the occurrence of something, neither always nor usually (*laysa dāʾim al-ʿijāb wa lā akthariyya al-ʿijāb*).⁴⁴ Thus, chance events are those which happen otherwise—that is, other than always or usually. Chance causes operate in exceptional cases, when the four causes, the principles on which nature habitually works, appear to be absent.

Endnotes

1. See his *Physics* II.3 and *Metaphysics* V.2. For modern accounts, see R.K. Sprague, “The Four Causes: Aristotle’s Exposition and Ours,” *Monist* 52 (1968), 298-300; M. Hocutt, “Aristotle’s Four Because,” *Philosophy* 49 (1974), 385-399; and R.B. Todd, “The Four Causes: Aristotle’s Exposition and the Ancients,” *Journal of the History of Ideas* 37 (1976), 319-322.
2. See Ibn Sīnā, *al-Shifāʿ: al-Ṭabīʿiyyāt: al-Samāʿ al-Ṭabīʿī*, ed. Saʿīd Zayed (Cairo: Organisme Général des Imprimeries Gouvernementales, 1983), 48 lines 8-11 and *al-Shifāʿ: Ilāhiyyāt*, ed. G.C. Anawati, Y. Mūsā, S. Dunyā and S. Zāyed (Cairo: Organisme Général des Imprimeries

- Gouvernementales, 1960), 257 lines 7-17.
3. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿi*, 46-7, 26-8, and *al-Shifāʾ: al-Ṭabīʿiyyāt: al-Kawn wa al-Fasād*, ed. Maḥmūd Qāsim (Cairo: Dār al-Kātib al-ʿArabī, 1969), 86-9.
 4. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿi*, 50 lines 11-17.
 5. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿi*, 52 lines 4-15.
 6. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿi*, 54 lines 10-17.
 7. *Al-Shifāʾ: al-Mantiq: al-Burhān*, ed. A.E. Afifi (Cairo: Organisme Général des Imprimeries Gouvernementales, 1956), 181 lines 1-7.
 8. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿi*, 53 lines 4-12.
 9. Known in Arabic as *saqmūniyā* (apparently from Gr. *skammōniā*), it is a certain plant from the hollows of which is extracted a mucilage, which is dried, and is called by the name of its plant; it is more repugnant to the stomach and the bowels than all the laxatives.
 10. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿi*, 55 lines 4-15; and 56 lines 1-4.
 11. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿi*, 56 lines 7-8.
 12. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿi*, 48 lines 14-17; and 49 lines 1-2.
 13. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿi*, 49 lines 1-3.
 14. *Al-Shifāʾ: Ilāhiyyāt*, 258 lines 1-8.
 15. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿi*, 49 lines 11-12.
 16. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿi*, 35 line 15.
 17. *Al-Shifāʾ: Ilāhiyyāt*, 257 lines 10-14.
 18. *Al-Shifāʾ: Ilāhiyyāt*, 37 lines 6-17; 38 lines 1-18; and 39 lines 1-4.
 19. *Al-Shifāʾ: Ilāhiyyāt*, 39 lines 6-16.
 20. *Al-Shifāʾ: Ilāhiyyāt*, 264 lines 5-7; 265 lines 1-8; and 266 lines 5-8.
 21. *Al-Shifāʾ: Mantīq: Burhān*, 95 and *Al-Shifāʾ: Ilāhiyyāt*, 8 lines 5-9.
 22. *Al-Shifāʾ: Mantīq: Burhān*, 322 lines 13-19.
 23. *Al-Shifāʾ: Mantīq: Burhān*, 323 lines 20-21 and 324 lines 1-7.
 24. *Al-Shifāʾ: Mantīq: Burhān*, 298 lines 4-17.
 25. *Al-Shifāʾ: Mantīq: Burhān*, 299 lines 1-9 and *Al-Shifāʾ: Ilāhiyyāt*, 276-8.
 26. *Al-Shifāʾ: Mantīq: Burhān*, 298 lines 4-13.
 27. *Al-Shifāʾ: Mantīq: Burhān*, 96 lines 4-11 and *Al-Shifāʾ: Ilāhiyyāt*, 180.
 28. *Al-Shifāʾ: Mantīq: Burhān*, 303 lines 9-16.
 29. *Al-Shifāʾ: Mantīq: Burhān*, 322 lines 7-12.

30. *Al-Shifāʾ: Ilāhiyyāt*, 167 lines 1-5.
31. *Al-Shifāʾ: Ilāhiyyāt*, 164 lines 12-18 and 165 lines 1-4.
32. *Al-Shifāʾ: Ilāhiyyāt*, 165 lines 4-9.
33. *Al-Shifāʾ: Ilāhiyyāt*, 167 lines 6-8.
34. Cf. his *al-Shifāʾ: al-Mantiq al-Qiyās*, ed. S. Zayed and I. Madkour (Cairo: Organisme Général des Imprimeries Gouvernementales, 1964), 322-4.
35. *Al-Shifāʾ: Ilāhiyyāt*, 167 lines 9-17.
36. *Al-Shifāʾ: Ilāhiyyāt*, 167 lines 1-17 and 168 lines 1-2.
37. *Al-Shifāʾ: Ilāhiyyāt*, 169 lines 9-14.
38. Al-Ghazālī's critique of Ibn Sīnā's views has been subject to many studies: Michael E. Marmura, *The Conflict over the World's Pre-eternity in the Tahāfuts of al-Ghazālī and Ibn Rushd*, Ph.D diss. University of Michigan, 1959; Muhammad Yasin El-Tahir Uraibi, *Al-Ghazālīs Aporien im Zusammenhang mit dem Kausalproblem*, Ph.D diss. University of Bonn, 1972; Carol L. Barger, *The Concept of Causality in Abū Ḥāmid Muḥammad al-Ghazālī's Tahāfut al-Falāsifah*, Ph.D diss. University of Wisconsin-Madison, 1978; and Ilai Alon, "Al-Ghazālī on Causality," *JAOS* 100 (1980): 397-405.
39. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿī*, 60 lines 5-17.
40. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿī*, 61 lines 5-16.
41. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿī*, 62 lines 5-18.
42. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿī*, 63 lines 4-17 and 64 lines 1-5; cf. Paul Lettinck, *Aristotle's Physics and Its Reception in the Arabic World* (Leiden: E.J. Brill, 1999), 164-5, who observes that this category is not mentioned by Aristotle but is discussed by John Philoponus in his *In Aristotelis physicorum octo libros commentaria*, ed. H. Vitelli (Commentaria in Aristotelem Graeca XVI-XVII; Berlin, 1887-88), 266,18-275,11 (ad 196b10 ff.).
43. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿī*, 64 lines 6-13.
44. *Al-Shifāʾ: Ṭabīʿiyyāt: Samāʿ Ṭabīʿī*, 65 lines 1-3.