gest basins have melted large volumes of the crust, creating LKFM impact melts and bringing up fragments of the deep crust for our inspection. The creation of the visible craters of the highlands has thoroughly broken up and churned the upper crust into a massive deposit of breccia, the megaregolith. The high rates of heat flow within the early Moon caused the recrystallization of some of this early breccia into granulites.

The formation of the megaregolith has mixed together the materials of the crust, with such mixing being in a mostly vertical, not lateral, manner. We see preserved on today's Moon some of the original provinces of different composition, provinces created in the era of crustal formation over 4 billion years ago. Mare basalts were extruded onto this complex, recycled surface, resulting in the creation of the mare-highlands dichotomy we see today. Since then, only the occasional formation of an impact crater has disturbed the silence of this once-violent, turbulent planet, whose earliest events are shrouded by the overprint of an impact record that may hold the key to understanding the early history of all the planets.

Chapter 7

Whence the Moon?

The origin of the Moon has been pondered, modeled, used as a justification for lunar exploration, studied, and endlessly debated. The popular press has often asserted that determining the Moon's origin is the ultimate goal for scientists. As we have seen, much of the story of the Moon really does not directly concern its origin. However, some aspects of the story of the Moon's evolution do merge into the story of its origin. Where did the heat come from to make the magma ocean? Was there a cataclysm, and if so, did it have something to do with accretion? The Moon is locked into synchronous rotation with the Earth (i.e., its "day" of 709 hours approximately equals its "year," the time it takes to revolve once around the Earth, see Chapter 1). Is this fact, as well as the Moon's curious orbital properties, related to its origin? These questions alone are enough to make the origin of the Moon of interest to lunar geologists.

Many people think that we have "solved" the problem of lunar origin. It is claimed that the "giant impact" model (a.k.a., the "Big Whack" or, in a more dignified mode, the "Collisional Ejection Hypothesis") has made earlier models of origin obsolete. All that remains to be done is to mop up the mere details of explaining the facts about the Moon and its history. Perhaps this is a solved problem, but I hope to show in this chapter some of the reasons the Big Whack model has such wide appeal and to suggest that perhaps we are not quite so smart as we tend to think we are.

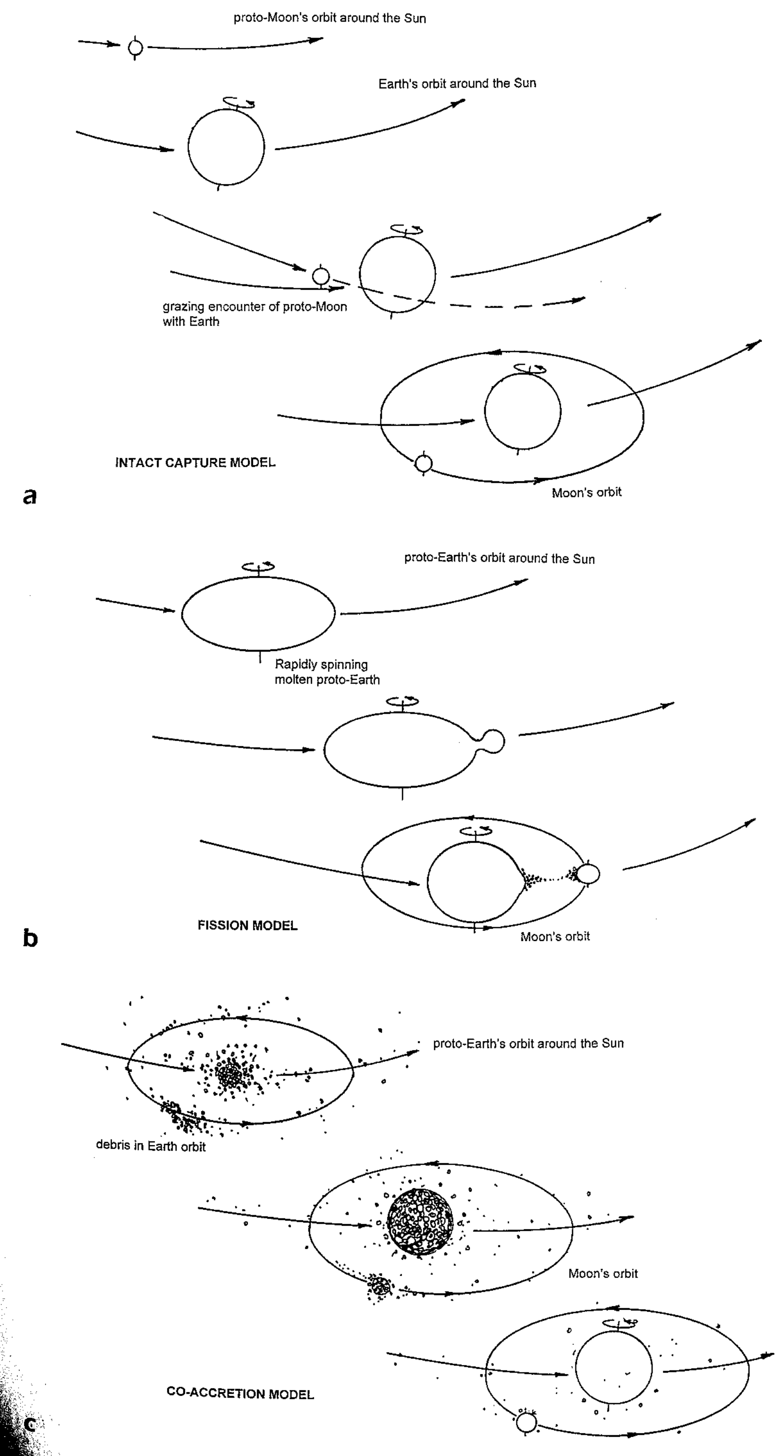
The Moon Cannot Exist!: Early Models for Lunar Origin

Ignoring for the moment the mythological stories of the birth of the Moon, lunar origin was first considered as a scientific problem by the classical theories of the origin of the solar system in general. The dominant model for the origin of the solar system holds that it condensed out of a cloud of hot gas, called the *solar nebula.* In such a model the planets are small pieces of the condensed material left over after star formation. Thus our Moon and the moons of the giant planets are similar to the planets that form around a star, a moon being the leftover bits of planet formation. This model, now called the *binary accretion model* (or the co-accretion model), was strongly defended by the astronomer Edouard Roche (who developed the idea of the tidal breakup of large bodies). In the co-accretion model of origin, the Moon is sometimes referred to as Earth's "sibling" because they both accumulated and grew as separate planetary objects in orbit around the Sun (Fig. 7.1).

As mentioned earlier, the Moon is receding, or moving away from Earth over time. This recession is extremely slow, about 4 cm per year. The converse of a lunar recession is that the Moon must have been much closer to Earth in the geological past. The astronomer Pierre Simon Laplace, inventor of the solar nebula model, attempted to account for the Moon's recession from Earth and concluded that tidal interactions between Earth, the Moon, and the Sun would conspire to cause the Moon's retreat from Earth over time. A corollary of this idea is that while the Moon recedes, Earth's rate of rotation is decreasing. Thus the evolution of the orbital relations in the Earth-Moon system is an important constraint on models of origin.

One of the first scientific models for the origin of the Moon was developed by George Darwin, the son of the great biologist and geologist Charles Darwin. Darwin, following up on Laplace's suggestion that the Moon is gradually receding from Earth, traced this recession back in time and concluded that Earth and the Moon were originally in physical contact with each other. After this idea is combined with the concept that

Figure 7.1 (opposite). Contending models for the origin of the Moon. The three classical origin models are (a) the "spouse" (the capture model), (b) the "daughter" (the fission model), and (c) the "sibling" (the binary accretion, or co-accretion, model). They have now been superseded (or joined) by the "Big Whack" model, or the collisional ejection hypothesis.



Earth's rate of spin was once much greater, it is a short leap to the suggestion that the Moon originated by being flung off Earth very early in Earth's history, when it was spinning very rapidly (Fig. 7.1). We now refer to this concept as the *fission model* for lunar origin. Because the Moon came from Earth, the Moon is considered to be Earth's "offspring." An early version of this idea held that when the early Moon spun off from Earth, the scar left on Earth was the Pacific Ocean basin. We now know that the ocean basins are very young features (younger than 70 million years), but this knowledge, unavailable to 19th-century astronomers, does not negate the fission model—other facts do that.

The third classical model for lunar origin grew out of the recognition that some satellites, particularly a few of the moons of the outer planets, orbit their primaries in a retrograde (opposite to planet rotation) direction. These orbits indicate that the small moons are wanderers in space that have been captured by their planets. The American astronomer Thomas Jefferson Jackson See, notable for claiming to be able to see craters on the surface of Mercury through the telescope (they are there but are too small to be visible from Earth), advocated that our Moon was captured into Earth orbit early in Earth's history. See claimed that the Moon had formed in the outer solar system and had wandered closer to the Sun, as had all of the planets through time. Eventually, some perturbation, probably an encounter with Jupiter, had flung the Moon into the inner solar system, where it was captured by Earth. The *capture model* (Fig. 7.1), modified appropriately to account for modern ideas about varied compositions in different parts of the solar system, was the third model of lunar origin to be tested by the data from the Apollo missions. In the capture model, the Moon is Earth's "spouse" (a forced marriage!).

What was Apollo's contribution to our understanding of the origin of the Moon? Essentially, the lunar samples and the Apollo program provided a variety of indisputable facts about the Moon's composition, the ages of its materials, and the reconstruction of certain past events—facts that any model for origin must explain. When these new facts are combined with certain astronomical data, we have information that can constrain the origin models. If a model for lunar origin makes predictions

contrary to some known property of the Moon, it cannot be correct. If a model cannot explain a certain observation or if it accounts for many but not all of the observations, it may or may not be correct; most certainly it is incomplete. We should look at the properties of the Moon that any successful model of origin must explain and the constraints that such a model must satisfy.

First, let's look at some general facts about the Moon and its orbit. The Earth-Moon system as a whole possesses a great amount of *angular momentum.* This property of rotating systems includes not only the rates of spin but how much mass is spinning and how difficult it would be to stop such spinning. Angular momentum can be neither created nor destroyed—it can only be transferred. The Earth-Moon system has the greatest amount of angular momentum of all the planet-satellite systems in the solar system. The Moon's orbit is neither in the plane of the ecliptic (the plane in which nearly all the planets orbit the Sun) nor in the equatorial plane of Earth (Fig. 1.6). The spin axes of Earth and the Moon are not aligned. Earth's axis is inclined 23°5 from perpendicular to the ecliptic plane (the plane of its orbit around the Sun), whereas the Moon's axis is nearly vertical, being inclined only 1.5° from the ecliptic plane. These physical properties of the Earth-Moon system do not allow the dynamic requirements of some origin models, such as the capture model, to be easily satisfied.

Another constraint on lunar origin is the bulk composition of the Moon. Estimating the composition of an entire planet is tricky. Because we know approximately what materials the planets originally assembled from (asteroids similar in composition to primitive meteorites), the density of a planet tells us something about its bulk composition. The Moon's density is about 3.3 g/cm^3 (grams per cubic centimeter), much less than Earth's density, at 5.5 g/cm^3 (but about the same as the density of the *mantle* of Earth). The most straightforward explanation for this difference is to assume that the Moon *as a whole* is depleted in iron, the most abundant high-density element in the solar system. Note that this bulk depletion in iron has nothing to do with the iron-rich nature of the mare basalts or certain other lunar rocks. The density data mean that the *whole* Moon has less iron than Earth. We think that this is because Earth has a massive core of iron-nickel metal, whereas the Moon has either no core or a very small one. Estimating the amounts of the other elements depends on knowing how and when the Moon melted and how these magma compositions changed during cooling. For that information, we needed data from the Apollo missions.

The legacy of Apollo has added significant details to our constraints on lunar origin. We have already noted that lunar rocks have no detectable water and very small amounts of elements that we consider to be "volatile" (i.e., have low boiling temperatures). The volatile hydrogen and helium that we find in the lunar soil come from the Sun. Note that the lunar samples are *depleted* in volatile elements (relative to Earth), not *devoid* of them. The distinction is important. Whatever process was responsible for creating the Moon, it could not totally eliminate all of the volatile elements. The bulk Moon is probably enriched in more *refractory* elements (those with high boiling points) than typical planetary material. The age of the Moon, as inferred from its content of radiogenic elements and isotopes, is about the same as the age of Earth, 4.5 billion years old.

Most elements have more than one stable isotope. For the Moon, study of the nonradioactive, stable isotopes of oxygen show that the Moon and Earth are closely related (Fig. 7.2). Because meteorites that have formed in different parts of the original solar nebula have different oxygen isotope compositions, the data from samples imply that Earth and the Moon must have formed in about the same place in the solar nebula, at roughly the same distance from the Sun. Additionally, the evidence from the samples indicates that the Moon had an ocean of magma early in its history. Such an episode of global melting requires a heat source, indicating that the initial formation of the Moon must have been a high-energy process.

None of the three classical models of lunar origin are particularly successful at satisfying these constraints or at explaining the properties of the Moon. In fact it has been said that because none of the models of origin are outstandingly successful, perhaps the Moon cannot exist! Although the fission model is supported by the depletion of the Moon in iron and volatile elements and by the oxygen isotope relations of Earth and the Moon, it has difficulty explaining the dynamics of the Moon's

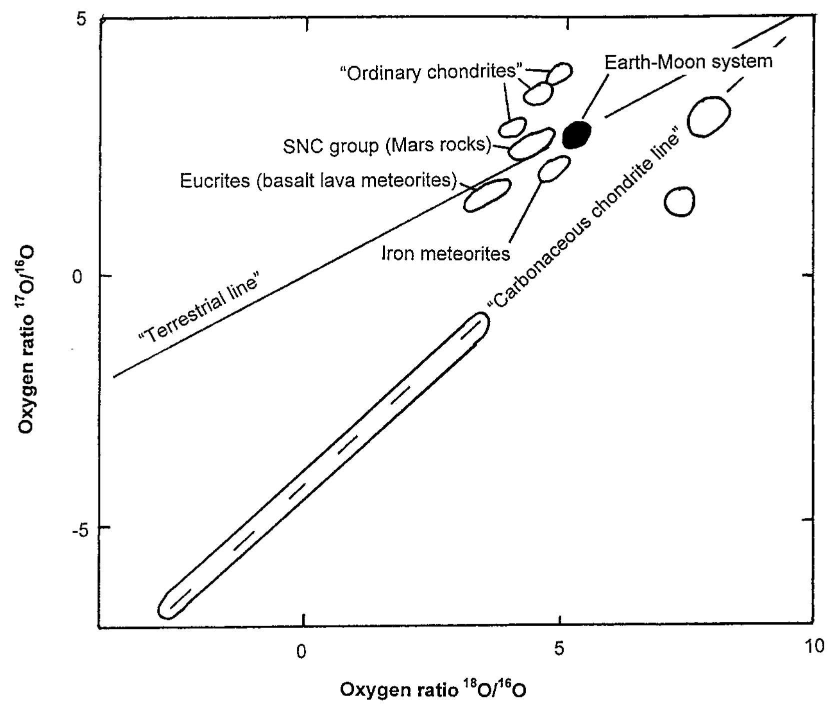


Figure 7.2. A map of the various isotopes of oxygen, a very common element in rocks, showing where various materials in the solar system formed. Whereas all known meteorite samples group into clusters, indicating that they formed at various locations around the solar system, the Earth-Moon samples make up a single cluster, indicating that however the Moon formed, it is closely related to Earth and was created in the same part of the solar system.

orbit and its orbital evolution. The capture model does not adequately account for these same dynamics and additionally has difficulty explaining the similarities in the oxygen isotopes (why are they the same if the Moon came from elsewhere in the solar system?). The co-accretion model perhaps does the best job of accounting for lunar properties, but even it does not really explain the Moon's depletion in volatile elements and in fact may be *too* flexible (a model that can be stretched to fit any data does not really explain anything). Even so, the co-accretion model was the favorite model of origin for most scientists in the years immediately after Apollo, probably because it was the least objectionable one of an admittedly poor lot.

This was the state of affairs when some lunar scientists decided to convene a conference on the origin of the Moon so that quiet contemplation of this thorny problem in the midst of active intellectual stimulation might produce some new insight into the problem of lunar origin. The conference was held in 1984 in Kona, Hawaii, and amazingly enough, it produced the hoped-for synergy (and possible breakthrough) in thought.

Giant Impacts and Rare Events: The Big Whack

The idea that very large impacts occurred early in the history of the planets grew out of two separate lines of thought. Scientists making theoretical models of the dynamics of planetary accretion found that nothing prevented the formation of fairly large planetoids (hundreds of kilometers across) or the subsequent collision of these objects with and incorporation into the existing planets. Some of these early planetoids could have approached sizes comparable to those of the surviving planets. Other scientists mapping the geology of the Moon and planets noted that the largest craters on these bodies were very large indeed; the South Pole–Aitken basin on the Moon is over 2,500 km across (about the distance between Houston and Los Angeles), and another supposed feature, the now discredited Procellarum basin, was alleged to be over 3,200 km in diameter (about the distance between New York and Phoenix). It required no great leap of faith to connect these two disparate observations into the concept that early planetary accretion involved very large impacts or, perhaps a better term, "planetary collisions." Such events might or might not have actually happened. Like the formation of large basins, a planetary collision was literally "hit or miss."

The idea that the Moon formed as a consequence of a giant impact on Earth was originally proposed independently by two research groups in 1976. That idea lay fallow until the Kona conference, when it exploded into the consciousness of the community. In one fell swoop the giant impact model seemed to have solved many problems of lunar origin. An off-center impact

could have simultaneously increased the spin rate of Earth and launched material into Earth orbit. An oblique collision also would have greatly increased the angular momentum of Earth and its newly added partner, accounting for one of the most curious of the dynamical properties of the Earth-Moon system. The ejected matter from the collision would have been superheated vapor, material that would certainly be depleted in, if not devoid of, volatile elements. The Moon's lack of iron could be the result of the fact that the colliding planet had already separated into core and mantle. Thus the material out of which the Moon formed came from the mantles of the two bodies—the proto-Earth and the colliding planetoid. Both of these parts of the planets would already have been depleted in iron, having removed most of their iron to their respective cores.

The Big Whack explains the Moon as debris ejected from a planetary collision. About 4.5 billion years ago, two planets were in the position around the Sun now occupied by the Earth-Moon system. The "proto-Earth" (Terra before the impact) was already planet-sized (only slightly smaller than the present Earth, 12,756 km in diameter), had differentiated into a core and mantle, and rotated at a much slower rate than Earth does currently. The other planet (the "impactor") was also differentiated into core and mantle and was about the size of Mars (6,787 km in diameter). The two planets collided off-center (Fig. 7.3); the impact strike by the Mars-size body was in the same direction as the rotation of the proto-Earth, resulting in a speeding-up of Earth's rate of rotation. But more important, a vaporized cloud of debris was flung off the colliding planets and into orbit around Earth (Fig. 7.3). It is still uncertain how much of this material came from each object, but current models suggest that most of the ejected material came from the impactor, the currently nonexistent, Mars-sized planet.

This superheated cloud of vapor expanded rapidly and cooled in space, forming an orbiting cloud of small, recondensed droplets of vaporized mantle material. One might expect such droplets to be enriched in refractory elements (those with very high boiling temperatures), having formed at extremely high temperatures and being depleted in volatile elements. The Moon is made from this material. A cloud of debris orbiting Earth would

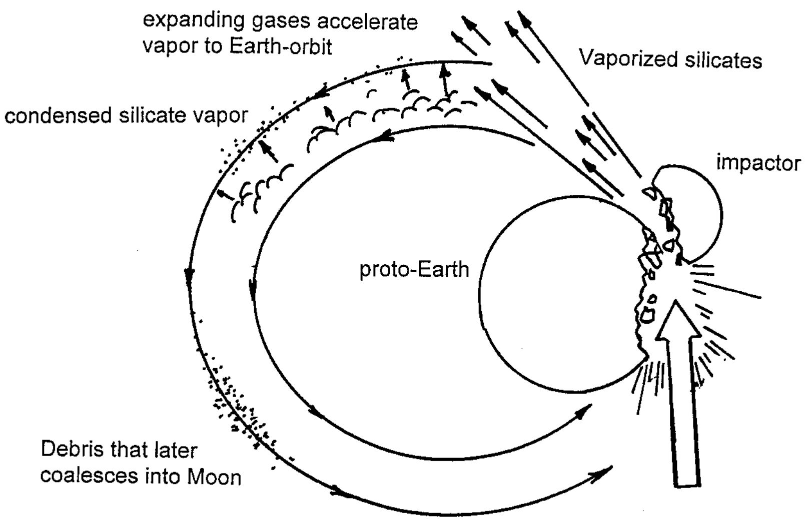


Figure 7.3. Diagram showing how the Big Whack model launches vaporized debris into orbit around Earth. An off-center, oblique impact seems to be required for this model. This vaporized rock later cools into many very tiny droplets that later accumulate very rapidly into a single body. According to advocates of this model, the rapid accumulation is responsible for the heat that created the magma ocean. Solid material would follow the inner ballistic path, reaccreting onto Earth; this material includes most of the impactor. After J. Wood, "Moon over Mauna Loa: A Review of Hypotheses of Formation of Earth's Moon," in Hartmann, Phillips, and Taylor, *Origin of the Moon,* 17–55

not be stable as an extended body, and this debris would have assembled itself into the Moon on very short timescales, tens of thousands of years, according to the dynamical models. Accretion within such short periods released large amounts of thermal energy, and it was this energy that was responsible for the creation of the magma ocean. The bulk of the mass of the impacting planet was incorporated into Earth. Such an incorporation might incidentally explain a curious (and until now apparently unrelated) fact about Earth's upper mantle: it appears to have rather high concentrations of the siderophile ("iron-loving") elements, such as nickel, cobalt, and iridium. In the Big Whack model this elevated content of siderophiles is the result of the incorporation of the core of the impactor planet into the mantle of Earth.

By simultaneously explaining so many properties of the Moon and its motions, this model of origin became an instant hit. Numerical modelers rushed to run new, ever more powerful mathematical simulations of giant impacts on their new supercomputers. Geochemists assembled data on the abundance of ever more obscure elements, like tungsten and molybdenum. Geophysicists reexamined their meager data on the interior to compute new and better models for the bulk composition of the Moon. Geologists assembled new geological maps showing that giant basins and, by implication, giant impacts were even more common on the planets than previously thought. All were eager to jump on the Big Whack bandwagon.

The Big Whack not only appears to solve many problems about the origin of the Moon but also appears to solve other problems of planetary science. The planet Mercury appears to be exceedingly dense, about 5.4 g/cm^3, suggesting that it has a relatively large amount of iron. Reasonable models for the structure of Mercury indicate a very large core, almost 60 percent of the volume of the planet. Venus rotates on its axis very slowly, once every 266 days, and backwards. Put these facts together with the Big Whack model and voilá—problem solved! Mercury and Venus are fragments of a giant impact, but in the opposite sense of the Earth-Moon collision. This impact reversed and slowed the rotation rate of Venus, stripped off much of the silicate mantle of Mercury (leaving the residuum of a planet, accounting for its anomalously large iron core), and sent Mercury careening in toward the Sun to become the innermost planet of the solar system. In another example, the spin axis of Uranus is tilted over 90° from perpendicular to the ecliptic plane. This extreme amount of tilt is proposed as a likely result of a giant impact early in the history of Uranus.

The Big Whack idea is now accepted wisdom in planetary science. It is seldom that a model gains such widespread acceptance so rapidly. It has done so largely because it readily explains so many different aspects of the Moon and its history.

Diverse phenomena from geochemistry to orbital dynamics appear to be accounted for. But is this because they really have been explained, or have they been explained away?

A Solved Problem?

Despite the popularity of the Big Whack model, it is far from certain that we have solved the problem of lunar origin. In science no model is ever "proven" to be true. Instead ideas (hypotheses) are created and then tested repeatedly. If an idea passes these tests, it becomes generally accepted. Any good hypothesis not only will explain existing facts but also will make predictions that are themselves testable and incidentally explain apparently unrelated facts. For a hypothesis to be a valid one, it must be capable of being "falsified," that is, while a hypothesis cannot be proven to be true, it must be capable of being proven untrue. We always have more error than truth in the multitude of hypotheses and models that scientific research produces, and unless this criteria is adhered to, we have no way to discard unusable, wrong concepts in favor of more correct ones.

One of the biggest difficulties with the Big Whack model is its elasticity. It has so many loose ends that it can be stretched to fit many observed facts. A model so unconstrained loses its predictive potency. For example, in terms of geochemistry, the Big Whack model (as outlined above) is actually a variant of the capture hypothesis, a variant that solves the dynamical problems that made the conventional capture unworkable. But it is not clear how geochemical data can constrain the model. The material that makes up the Moon comes from elsewhere in the solar system. If the Mars-sized impactor came from somewhere else, why do the oxygen isotopes of Earth and the Moon match so closely? Advocates of the Big Whack explain this problem away by saying that the isotopic data cannot be overinterpreted and that isotopic affinities merely indicate where materials are created in the solar system in general. If so, then perhaps the oxygen isotope data are not really constraints on origin at all.

Another difficulty with the Big Whack is the nature of the Moon's depletion in volatile elements. It is generally thought that this depletion is one of the lunar facets most readily explained by the Big Whack. However, the patterns of elemental depletion do not seem to make sense. For example, some volatile elements clearly were incorporated into the accreting Moon, as evidenced by the gases driving the eruptions of dark mantle material (ash) and the vesicles in the mare basalt lavas. How were these gases preserved in such a high-temperature environment? Some elements (such as manganese) are found in the samples that are volatile in a geochemical sense, yet these elements are not as depleted as other volatile elements of comparable geochemical behavior. Is this also a consequence of the Big Whack? Clearly, the effects of volatile depletion are not understood well enough to use them as a real constraint on the processes operating during giant impacts.

My discussion of these difficulties does not aim to disparage the Big Whack model but to suggest that we do not yet fully comprehend the origin of the Moon. In particular we need to understand how such an event would affect properties that we can determine and measure to a precision adequate to allow us to really test the model. If we cannot devise rigorous tests of such a hypothesis, then it becomes merely an interesting idea, not a generalized theory of lunar origin. The Big Whack model is the most promising idea yet developed to explain some of the more puzzling aspects of the origin of the Moon. Future work on all of its various components will enable us to better assess its true value.