Impact Origin of the Moon?

Erik Asphaug

School of Earth and Space Exploration, Arizona State University, Tempe, Arizona, 85287; email: easphaug@asu.edu

Annu. Rev. Earth Planet. Sci. 2014. 42:551–78

First published online as a Review in Advance on March 21, 2014

The *Annual Review of Earth and Planetary Sciences* is online at earth.annualreviews.org

This article's doi:
10.1146/annurev-earth-050212-124057

Keywords

planets, collisions, Moon formation, accretion

Abstract

Earth formed in a series of giant impacts, and the last one made the Moon. This idea, an edifice of post-Apollo science, can explain the Moon's globally melted silicate composition, its lack of water and iron, and its anomalously large mass and angular momentum. But the theory is seriously called to question by increasingly detailed geochemical analysis of lunar rocks. Lunar samples should be easily distinguishable from Earth, because the Moon derives mostly from the impacting planet, in standard models of the theory. But lunar rocks are the same as Earth in O, Ti, Cr, W, K, and other species, to measurement precision. Some regard this as a repudiation of the theory; others say it wants a reformation. Ideas put forward to salvage or revise it are evaluated, alongside their relationships to past models and their implications for planet formation and Earth.

1. INTRODUCTION

Giant impact: a solar system collision involving two planet-sized bodies

Coaccretion: the theory that the Moon and Earth accreted together in the protoplanetary nebula, as an orbiting pair

Capture: the theory that the Moon and Earth formed separately, and the Moon was later captured into orbit

Fission: the theory that Earth once spun so rapidly that the Moon was launched into orbit from its upper mantle

Theia: the proposed planet that collided with Earth to make the Moon

In the aftermath of the Apollo landings, a handful of scientists made novel arguments for a Moon-forming giant impact. Hartmann & Davis (1975) advanced the emerging view (Safronov 1972, Wetherill 1976) that planet formation ends with pairwise accretionary collisions, also known as giant impacts, that might eject mantle rock into orbit. Cameron & Ward (1976) argued that a Mars-sized off-axis collision bestowed the anomalously large angular momentum to the Earth-Moon system. These and other papers set the stage for a modern theory that would, within twenty years, all but displace the previous Moon formation concepts of coaccretion, capture, and fission.

Resolute efforts of dynamical modelers (e.g., Benz et al. 1989, Cameron & Benz 1991, Ida et al. 1997, Canup et al. 2001, Canup & Asphaug 2001), supported by increasingly advanced simulations and visualizations, led to the adoption or acceptance of a specific giant impact scenario, that of a Mars-sized terrestrial planet Theia accreted by the proto-Earth toward the end of Earth's formation. This standard model is well aligned with the modern theory of planet formation, and it satisfies the major physical, petrological, and bulk compositional aspects of the Earth-Moon system, including its anomalously large angular momentum and satellite mass.

Lulled into thinking that Moon formation was largely settled, scientists awoke to find the barn door open, perhaps on fire: Concurrently (e.g., Wiechert et al. 2001), the painstaking isotopic geochemistry of lunar rocks was leading inexorably to the conclusion that the Moon is too Earth-like to include substantial fractions of Theia, which it does in the standard model, unless Theia is indistinguishable from Earth. What follows is an account of the so-called isotopic crisis (Melosh 2009) and its responses.

For a comprehensive introduction to the giant impact hypothesis, earlier reviews (e.g., Stevenson 1987, Melosh 1990, Jones & Palme 2000) are of long-standing value. Canup (2004) described the modern theory along with a comprehensive review of simulations. Shearer et al. (2006) provided an excellent introduction to the various concepts of lunar origin in the context of the thermal and magmatic evolution of the Moon. The hypothesis casts a great shadow, too. Zahnle et al. (2007) adopted the giant impact as the basis for thermal, physical, and chemical models that characterize the emergence of a habitable planet. How stable is the edifice of these ideas?

The crisis is that the Moon derives predominantly from the mantle of the impacting planet Theia, in the standard model, for simple reasons of angular momentum. This is in direct conflict with precise isotopic measurements (e.g., Wiechert et al. 2001, Jacobsen 2005, Touboul et al. 2007) that find that lunar rocks are indistinguishable from Earth rocks in O, Ti, Cr, W, K, and other species. There are two classes of response: that we must put a stake in the heart of the concept of Theia, and that we must identify a new kind of giant impact. Certainly, previously discarded theories deserve a revised and careful look.

Foremost is the concept of fission, caused by rotational instability in a fast-spinning proto-Earth (Darwin 1879). This concept has evolved into a novel theory (Ćuk & Stewart 2012) called impact-triggered fission; also, as we shall see, the giant impact itself works through rotational instability. Disintegrative capture (e.g., Matsui & Abe 1986) is relevant to the modern picture, and so are tidal capture (Gerstenkorn 1955, Goldreich 1966) and variations on coaccretion (Morishima & Watanabe 2001) described below. Yet, it attests to the robustness of the theory that novel alternatives such as exploding planets (Terez & Gerasimov 2009) have not gained traction.

We are cautioned by another popular giant impact theory that met a dead end. Motivated by telescopic observations of solar prominences, and by the orbits and densities of planets, Buffon (1749) wrote convincingly about how a grazing collision of a massive comet into the Sun created the Solar System. His idea foundered when comet masses proved to be unmeasurably small. But stars are much larger, and the understanding of stellar proper motions led Moulton (1905), Chamberlin (1916), Jeans (1919), and Jeffreys (1924) to propose that the Solar System was created by a tidal collision with a passing star. Supported by detailed calculations and by images of clumpy spiral nebulae (later shown to be galaxies), this became the textbook explanation (Jeffreys 1924).

Irreconcilable problems led to the theory's fragmentation and demise, as recounted by Brush (1978): The angular momentum distribution of the Solar System (Jupiter and Neptune) could never fit the theory, and parcels of matter extracted from a star's convective zone would be dispersive, not planet forming. It took decades for a modern theory to replace it (e.g., Safronov 1972, Wetherill 1980). Coming full circle, the new theory features its own planet-forming collisions, including the one that is proposed have caused the Moon.

So theories fall apart, and then there are great advances. We are at the cusp of a great advance, and however exciting this may be scientifically, it means that a review is never finished. One builds until there can be no further piling on of timbers. As for exploring all the intellectual avenues, I have been remiss, and biased by a modeler's perspective. When given a choice I have cited earlier works along with modern efforts; regrettably this comes at an expense of acknowledging many of the significant contributions that got us here.

2. MOTIVATION

In six historic voyages to the lunar nearside (Wilhelms 1993), Apollo astronauts returned hundreds of kilograms of igneous rocks, cumulates, and breccias. Their endeavors added to the puzzle of the Moon's bulk composition, which was known since the first determination of the Earth-Moon barycenter to be $\rho_{☾} = 3.3$ g/cm^3, only $\sim$3/5 the density of Earth. Although this is the same bulk density as chondrites (e.g., Urey 1952), lunar samples are igneous, unlike primitive meteorites. Measurements of their water abundance suggested a bone-dry planet.

Apollo seismometers confirmed the Moon to be of rocky composition throughout, with hardly any core. Subsequent analysis of seismic detections (e.g., Weber et al. 2011) and studies of the induced magnetic dipole (Hood et al. 1999) have determined that the lunar core is at most 4% of the mass of the Moon and $\lesssim$400–700 km in diameter. By comparison, Earth's core is $\sim$30% of the planet's mass and half of its diameter, comparable to the free iron mass fraction in unmelted chondrites.

Coaccretion and capture, two of the leading hypotheses before Apollo, were put on the defensive. Already challenged to explain the high angular momentum and mass ratio of the Earth-Moon system, they now had to explain the missing iron and water—this stranded ball of rock. The other leading theory, rotational fission, could satisfy the geochemistry by flinging Earth's mantle into orbit, but it lacked a primary mechanism. The time was right for a modern theory.

2.1. A New Planet-Forming Process

Beginning in the early 1980s, the giant impact hypothesis could be simulated using pioneering computational techniques originally devised to study stellar interactions (Benz et al. 1986) and nuclear explosions (Kipp & Melosh 1986). The earliest calculations showed the effective segregation of mantle silicates from core iron, the production of global shock waves as the impactor plowed through the target, and the initial stages of the ejection of a silicate-dominated plume.

Although not well resolved and prone to artifacts (see Section 5), the first simulations were evocative and their visualizations impressive. The giant impact was immediately understood to be able to explain the loss of iron: In simulations, the cores tended to merge at Earth's center while a fraction of the ejected silicates got stranded into orbit. But 3D codes with self-gravity were able to resolve the colliding bodies with only ~3,000 particles total, putting only a few dozen particles into orbit to make the Moon.

Escape velocity: the ejection speed at which one object escapes the gravitational potential of another; this is also the speed at which two bodies falling from great distance will collide. For the standard model this is ~10 km/s

Protolunar disk: the disk of material (melt and vapor) that was left orbiting Earth following the proposed giant impact

Magma ocean: melted silicates hundreds or thousands of kilometers deep, whose solidification is accompanied by fractional crystallization and floatation of crustal cumulates

Anorthosites: the oldest lunar rocks, comprising much of the highlands crust, believed to have formed when plagioclase feldspar floated to the top of a magma ocean

By the 1990s, the ejected, captured protolunar material could be resolved with hundreds of particles, so its dynamics (e.g., Ida et al. 1997) and composition could be evaluated. Investigations began to converge upon a standard or reference model, in which a Mars-sized terrestrial protoplanet collides into the proto-Earth near the end of Earth's formation (**Figure 1**). That is, $M_2 \sim 0.1$–$0.15\, M_1$, $M_1 + M_2 \approx M_\oplus$, $v_{\rm imp} \sim 1.0$–$1.1\, v_{\rm esc}$, and $\theta \sim 45°$ (Canup & Asphaug 2001). The escape velocity $v_{\rm esc}$ is the speed at contact when two planets of radii R_1 and R_2 and masses M_1 and M_2 fall toward each other from many radii away; $v_{\rm esc} = \sqrt{2G(M_1 + M_2)/(R_1 + R_2)}$. So in a dynamical sense, a giant impact at $v_{\rm imp} \sim v_{\rm esc}$ is slow, in that the impact energy is less than the gravitational potential, resulting in accretion.

Velocity is relative to size, but in absolute terms $v_{\rm esc}$ transitions from subsonic to hypersonic as accreting bodies get larger (**Figure 2**; **Table 1**). Planetesimal-scale accretion is a subsonic process dominated by mechanical and gravitational interactions, whereas Mars-sized and larger accretion features intense global-scale shocks, causing vaporization. This comparison emphasizes the importance of using a good equation of state in modeling giant impacts, because the theory for Moon formation relies upon the captured fraction of the ejected melt-vapor plume.

The idea that the Moon accreted out of a completely melted protolunar disk led to the development of magma ocean models, reviewed by Warren (1985) and Shearer et al. (2006), that showed how the predominantly anorthosite composition of the lunar highlands would be the expected result of the crystallization of plagioclase feldspar at depth and its buoyancy to form a rafted lunar crust. Also, an initial melt-vapor state for the protolunar disk would explain the efficient differentiation of the $\lesssim$4% iron fraction, which would otherwise require high temperatures and pressures to precipitate in situ out of an iron-poor composition.

2.2. Isotopic Constraints

However convincing these arguments or convenient these explanations, critics have long raised the point (e.g., Drake 1986, Ringwood 1986) that the Moon is too Earth-like in oxygen isotopic composition to be substantially contaminated by another planet. This argument came to the forefront when well-resolved calculations (Canup & Asphaug 2001, Canup 2004) honed in on a model that satisfied all bulk aspects of the Moon, except its formation mostly out of Theia. Meanwhile, geochemists (e.g., Wiechert et al. 2001, Jacobsen 2005, Touboul et al. 2007) were finding that not only oxygen but also other species were indistinguishable from Earth.

Consider the three-isotope system ^{16}O, ^{17}O, ^{18}O. Earth is enriched, by some process or processes, in heavier oxygen by about 7% relative to the Sun (McKeegan et al. 2011). Whatever the cause, lunar rocks are enriched the same as Earth's in both heavier isotopes, so that the deviation Δ^{17}O of lunar rocks from the terrestrial fractionation line (TFL) in the three-isotope diagram is zero to measurement precision (see Section 4). Any mechanism involving capture, including the standard model, must explain this.

The original terrestrial planetary embryos, including Mars and presumably Theia, are believed to have accreted in a few million years (e.g., Dauphas & Pourmand 2011) as amalgams of isolated feeding zones. Each would have obtained a distinctive isotopic trace, like an unblended whiskey. Differentiated meteorite parent bodies have strongly identifiable isotopic signatures (see **Figure 3**), presumably because they sampled unique regions of the protoplanetary disk, accreting out of diverse thermodynamical, radiative, and chemical environments. As reviewed by Meier et al. (M.M.M. Meier, A. Reufer, and R. Wieler, manuscript in revision), Δ^{17}O ≈ 0.3 for Mars and -0.2 for Vesta, so Theia would also plot quite distinct from the TFL except by remarkable

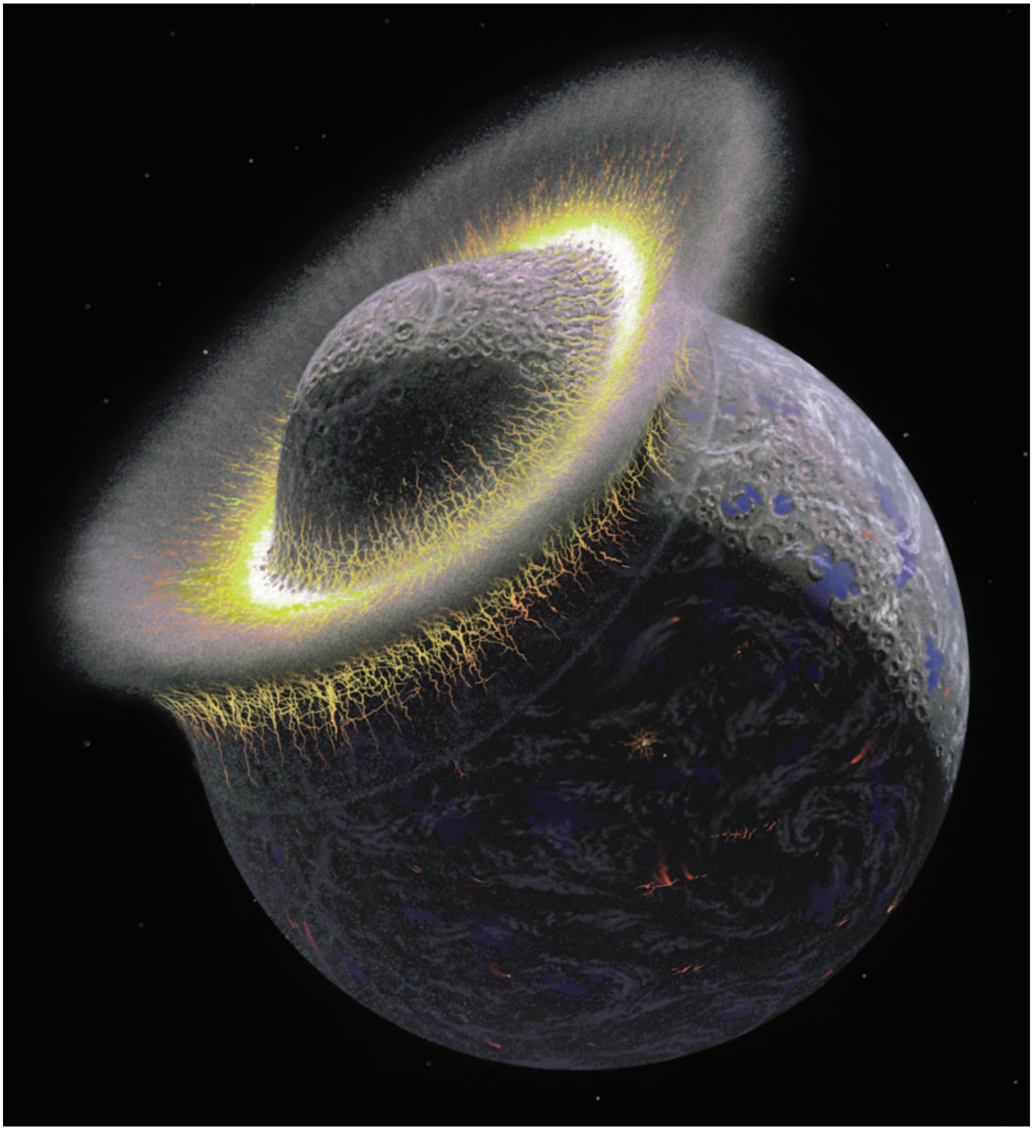

Figure 1

Artist's rendering of the standard model. The impact is from the upper left. The antipode of the colliding planet (Theia) and most of the proto-Earth remain untouched ~10 min after initial contact. The shock front progresses at ~10 km/s. Surface materials in the contact zone are jetted into space while shock heating forms a magma ocean deep inside. Over the next several hours the collision unfolds as an accretion dominated by angular momentum and gravity. In the standard model, the leading component of Theia (farthest to the right) is shredded into a disk about the finished Earth, forming an alien Moon. Illustration by Don Davis, used with permission.

coincidence. Instead, $\Delta^{17}O = 0.003‰ \pm 0.003‰$ (2σ) for the Moon (Wiechert et al. 2001), well inside the $\pm 0.011‰$ range of measurements for the silicate Earth.

Pahlevan & Stevenson (2007) have argued that turbulent diffusion would homogenize the isotopic composition of the protolunar disk with Earth's postimpact silicate atmosphere. The hotter and more expansive the protolunar disk, the more efficient the isotopic homogenization;

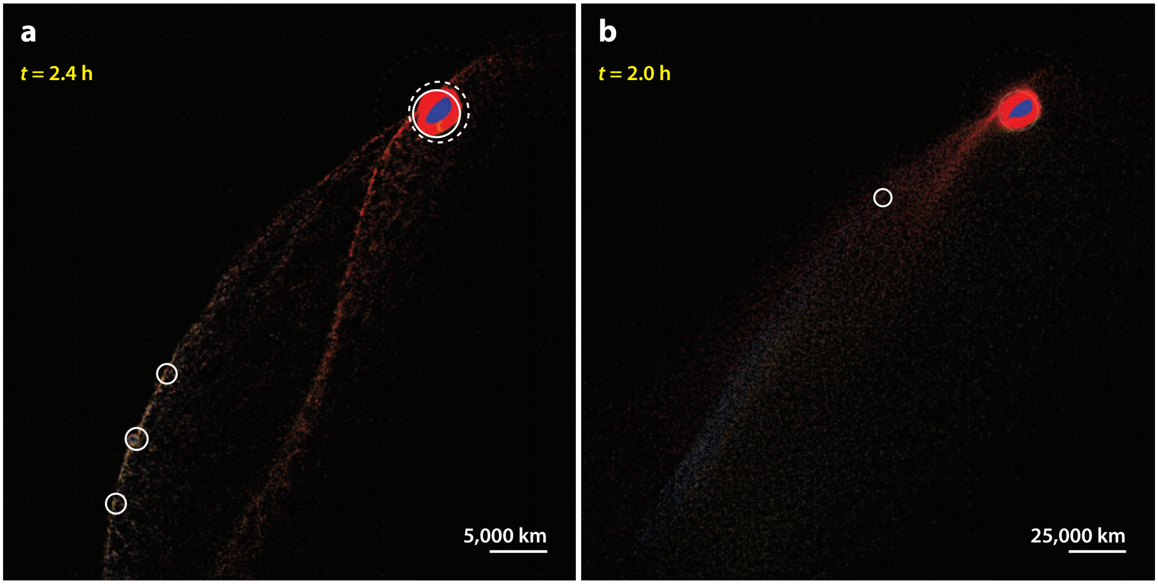

Figure 2

Giant impacts that have identical $M_2/M_1 = 0.2$, $\theta = 30°$, and $v_{imp} = 3v_{esc}$ but vary by 10^3 in colliding mass. Shown are hit-and-run collisions that demonstrate the combined effects of shock and self-gravitation in giant impacts. (*a*) A planetesimal-scale collision ($R_1 = 700$ km, $R_2 = 400$ km) involves no shocks; mantle silicates (*red*) are diverted gravitationally from the dense streams of iron-dominated material (*blue*). Self-gravitating clumps are circled. (*b*) An Earth-scale collision ($M_1 = M_\oplus$) produces intense shocks throughout both bodies ($v_{imp} \approx 30$ km/s), causing complete vaporization that drives plume expansion. Simulation by A. Reufer (University of Bern/Arizona State University).

this favors the most energetic giant impact scenarios. It is a powerful idea that would reconcile the isotopic crisis, but the details are problematic.

For one thing, the process of diffusion is dynamically self-limiting (Desch & Taylor 2013), as isotopic equilibration implies mass exchange and hence angular momentum transport that might collapse the disk. Also, this process cannot be played out to completion because isotopic equilibration implies chemical equilibration. How are elemental systems like oxygen homogenized isotopically when the volatiles and semivolatiles are not? Diffusion equilibration is also challenged to explain the recent analysis of lunar titanium (Zhang et al. 2012), a highly refractory species that

Table 1 Escape velocity v_{esc}, the characteristic velocity of planet-forming collisions, for various bodies of terrestrial-like composition

Planet	Radius (km)	v_{esc} (m/s)
Earth	6,371	11,200
Mars	3,390	5,000
Moon	1,737	2,400
Vesta	263	360
Lutetia	50	70

For uncompressed rocky bodies, v_{esc} (in m/s) is comparable to the radius (in km). The speed of sound c_s is ~3,000 m/s in nonporous water ice and ~5,000 m/s in rock. Collisions faster than the sound speed create shocks, leading to vaporization and melting in planet-scale collisions.

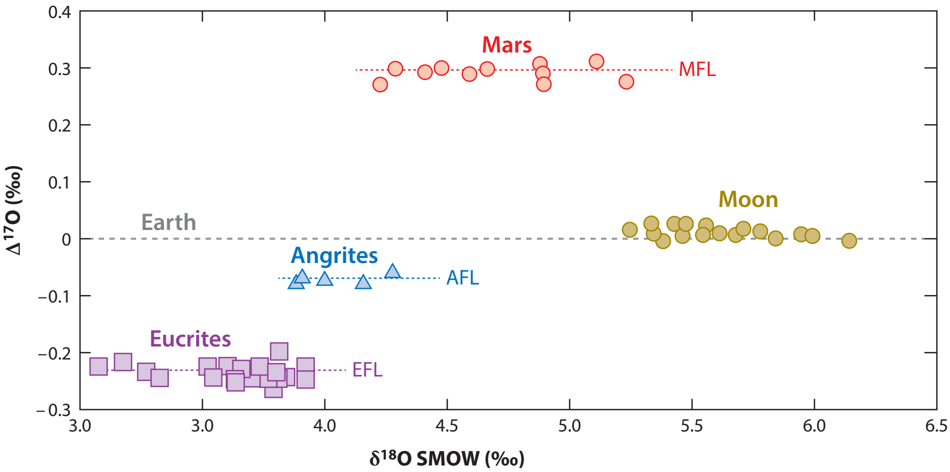

Figure 3

Oxygen isotope data for Earth (*gray dashed line*), the Moon, Mars, Vesta (eucrite meteorites), and angrite meteorites. This diagram plots Δ^{17}O, the departure from the TFL in ‰, versus δ^{18}O, the ratio of ^{18}O to ^{16}O normalized to Standard Mean Ocean Water (SMOW). The spread along the x-axis is due to chemical processing within each body, but all values for each type lie close in Δ^{17}O, representing mass-dependent fractionation during melting and crystallization. Terrestrial, Mars, angrite, and eucrite fractionation lines are labeled (TFL, MFL, AFL, and EFL, respectively). Moon rocks plot on top of the TFL. From Kevin Righter (NASA Johnson Space Center).

would be considerably more difficult to mobilize than oxygen. The Moon is Earth-like in Ti to 4 ppm, compared with the ~600-ppm range typical of meteorites.

2.3. Earth-Moon Dynamics

The Earth-Moon system has an unusually high mass ratio, $M_☾/M_\oplus = 1.2\%$, causing powerful tidal coupling. It also has high specific angular momentum, $\vec{L}_{EM}/(M_\oplus + M_☾)$, where $\vec{L}_{EM}$ is the rotational angular momentum of both bodies plus their orbital angular momentum about their barycenter. If somehow brought together into one place, $\vec{L}_{EM}$ would equal the angular momentum of an Earth-mass planet spinning with a period of ~5 h—an anomalously large value that Cameron & Ward (1976) took to indicate a giant impact.

Tides raised on Earth by the Sun and Moon, and on the Moon by Earth, make the system "by far the most complex" (Goldreich & Soter 1966, p. 381), especially when one traces the dynamics back to the first few million years after lunar formation. Orbital energy is converted to elastic and potential energies that are dissipated by internal friction; this heats the bodies and effects the transfer of angular momentum. Consequently, the lunar semimajor axis $a_s = 60.3\,R_\oplus$ is expanding at $\dot{a}_s = 3.8$ cm/year today. Because $\dot{a}_s \propto a_s^{-11/2}$, tidal expansion might have been tens of kilometers per year when $a_s \sim 5\,R_\oplus$.

Gerstenkorn (1955) studied the tidal evolution of the Moon, assuming it was initially captured by nebula drag into a high-inclination retrograde orbit about Earth. Tidal drag on the retrograde orbit would cause the Moon to spiral in, and then it would flip, becoming a prograde, receding satellite and evolving to its present orbit and inclination. The intensive working of tides that at the closest approach are kilometers high would lead to a globally melted Moon, a consequence of Gerstenkorn's model that turned out to agree with lunar geology. Other aspects—such as creating

$\vec{L}_{EM}$: the combined angular momentum of Earth and the Moon: their orbits about their barycenter plus their individual rotations an iron-free silicate body in the first place and having the nebula disappear just in time—were unexplained.

Corotation: the semimajor axis a_c where the orbital period $T = 2\pi\sqrt{a_c^3/GM_\oplus}$ of a satellite equals the planet's rotation period P_{rot}; for the postimpact Earth, $a_c = 2.3 R_\oplus$ assuming a 5-h period

Roche limit: $R_{roche} = 2.456 R_p (\rho_s/\rho_p)^{1/3}$ is the orbital radius where a liquid satellite of density ρ_s begins to tidally disrupt around a planet of radius R_p and density ρ_p; for the Moon, $R_{roche} = 2.9 R_\oplus$

Evection: a resonance between a satellite's period of orbital precession around a planet, and a planet's orbital period around the Sun

A satellite orbiting inside the corotation radius will spiral in toward the planet due to tides, as Phobos does around Mars. A satellite orbiting outside will spiral out, as the Moon must have done. This sets an additional constraint on where the Moon might have formed, having to be outside of $a_c \sim 2.3\ R_\oplus$.

Whether captured or coaccreted, or formed in a fission or giant impact, the Moon was once much nearer to Earth (Darwin 1879). Starting from the Roche limit $R_{roche} \sim 2.9\ R_\oplus$ and evolving out to $\sim$10 $R_\oplus$ during the first $\sim$100 Ma makes for a fascinating and important calculation (e.g., Goldreich 1966, Kaula 1971, Touma & Wisdom 1994), because on its way the Moon falls into a "devious network of resonances—courtesy of the Earth and Sun—that mercilessly distort its orbit" (Touma 2000, p. 165).

Among the most intensive of these is the evection resonance, where the precession of the lunar periapse is synchronous with Earth's orbit about the Sun. During evection, thought to occur at $a_s \sim 6$–$8\ R_\oplus$, the lunar orbit's periapse stands still in the Sun-Earth frame, raising its eccentricity and causing massive tidal dissipation and possible melting inside the Moon (Peale & Cassen 1978, Touma & Wisdom 1998). Evection can cause the transfer of angular momentum from the planet-satellite pair to the planet's orbit. In the case of Earth and the Moon, Touma & Wisdom (1998) found that evection could remove at most $\sim$10–20% of the original angular momentum, so that $\vec{L}_{EM}$ would be relatively constant since lunar formation. For better or worse, this conclusion anchored the hydrodynamical simulations that led to the standard model and falsified hypotheses requiring too much angular momentum.

More recently, Ćuk & Stewart (2012) have obtained a different result, finding that half of $\vec{L}_{EM}$ could be lost through evection. But this requires slow migration through the resonance (low dissipation inside Earth) and a ratio $A \sim 0.1$ for the tidal dissipation inside the Moon relative to Earth. The detailed calculations are challenging and require some shortcuts. Also, tidal dissipation varies greatly depending on whether a planet is liquid, solid, partially melted, or covered in oceans, so there is no particular reason to expect $A \sim 0.1$. Because angular momentum is a key constraint on giant impact scenarios, evection is a hot topic.

The lunar inclination $i = 5.14°$ relative to Earth's equator must also be reconciled with any formation theory; it motivated Gerstenkorn's theory. Inclination creates an offset tidal bulge; projecting the orbit back to R_{roche} leads to a starting inclination $\gtrsim 10°$. Upon first examination, this is a fundamental obstacle for the giant impact theory, because a postimpact disk containing $\gtrsim 2\ M_☾$ (the minimum amount required to accrete the Moon according to Ida et al. 1997) would damp in months to a low inclination through mutual collisions.

Ad hoc solutions include a second giant impact that knocked Earth off-kilter after Moon formation and a disk dominated by a single high-inclination clump (e.g., disruptive capture). Coupling between the Moon and a massive ($\sim$0.5 $M_☾$) interior disk (Ward & Canup 2000), or with companion satellites (Ćuk & Stewart 2011), would also be able to excite the Moon's inclination following its formation. Lunar inclination is a powerful constraint and contains key information, but is by no means a fatal blow to the giant impact hypothesis.

3. EVIDENCE OF A GIANT IMPACT

The resilience of the standard model is due in no small part to its close relationship to the late-stage hypothesis of planet formation. This well-accepted idea, that terrestrial planet formation concluded with a series of giant pairwise collisions (Wetherill 1985), is strongly supported by geochemical heterogeneity (e.g., Rudge et al. 2010) and is the natural outcome of planet formation according to all modern dynamical simulations (e.g., Raymond et al. 2006). Also, recent observations of circumstellar debris disks (Johnson et al. 2012) show what could be direct signatures of giant impacts.

So the concept of colliding planets is here to stay. Five lines of evidence argue for its relevance to Moon formation: (*a*) the combined Earth-Moon angular momentum $\vec{L}_{\mathrm{EM}}$, (*b*) the globally melted geology of the Moon, (*c*) the lack of a lunar core, (*d*) the lack of lunar volatiles, and (*e*) the relatively late timing of Moon formation. Analogous moons in the outer Solar System that are thought to have formed by giant impact also shed light on the debate.

3.1. Melting, Mixing, and Differentiation

Petrology, geology, remote sensing, and gravity all indicate pervasive and perhaps global melting of the Moon (Warren 1985, Shearer et al. 2006). A planet-scale collision provides the enthalpy for global melting in two ways. The deep mantle of Theia, already statically loaded to tens of gigapascals (Asphaug et al. 2006) and further loaded by shock (e.g., Stevenson 1987), explodes upon decompression into an orbiting torus of melt and vapor.

Production of an immediately melted and vaporized disk is an appealing explanation for the dry Moon, but the details are not yet satisfactory. Water loss requires lunar accretion to be slower than vapor diffusion, which depends upon temperature, pressure, and droplet size. It is not clear how much water can get out before coagulation. Conversely, coagulation must have occurred before volatile loss was complete, to account for the substantial water that is retained by the Moon ($\gtrsim$10ď% the fraction in Earth; Hauri et al. 2011) and for the retention of semivolatile sodium and potassium. Evidently the process shut down before completion (cf. Desch & Taylor 2013).

A giant impact would be entirely consistent with a terrestrial magma ocean (Melosh 1990, Elkins-Tanton 2012). It might furthermore explain the loss of the primordial atmosphere (Melosh & Vickery 1989, Ahrens 1993, Genda & Abe 2005), making Earth unlike Venus. After the impact, a massive silicate atmosphere radiating at ~2,500 K would rain out in a few thousand years, whereas a longer-lived steam atmosphere might insulate Earth's magma ocean for $\gtrsim 10^6$ years according to Zahnle et al. (2007)—a Hadean planet evolving toward a habitable biosphere. A giant impact might even have ejected an original, too-salty ocean (Sharp & Draper 2013), making the planet perfect for life.

The giant impact is becoming deeply embedded in modern geologic thought, so it is important that we get it right. But the slim record of Earth's Hadean geology is found only in zircons, some as old as ~4.4 Ga (Harrison 2009). The constant exchange of impact material between Earth and the Moon, and the early solidification and truly ancient nature of the lunar crust, makes the Moon itself a better place to look for relics of Earth that might provide closure to this puzzle.

3.2. Age of the Moon

The tidal bulge, offset from the Earth-Moon center line by friction, causes a torque that raises the Moon's orbit and spins down Earth. Integrating backward, and making assumptions that he suspected were incorrect, Darwin (1879) obtained $\gtrsim$56 Ma for an age of closest proximity, leading to his fission hypothesis described above. This number was in agreement with Kelvin's estimate (Thomson 1864) for the age of Earth, $\gtrsim$24 Ma (later revised upward), derived from conductive heat flow in a cooling solid (the "turkey in the freezer"; see England et al. 2007).

Subsequent analysis of tidal evolution (Jeffreys 1924, Gerstenkorn 1955) and the radiometric dating of rocks led to the understanding that whatever happened took place before 2 Ga. The oldest inclusions in meteorites have now been dated precisely using Pb-Pb and Al-Mg chronometers, indicating $t_0 \sim 4.568$ Ga for the birth of the Solar System (Bouvier & Wadhwa 2010).

Radiometric ages for the birth of Earth and the Moon will always be much less precise, because the short-lived chronometers are extinct. Moon formation happened between ~30 and 200 Ma after t_0, according to best estimates derived using various chronometers (Jacobsen 2005, Halliday 2008, Yu & Jacobsen 2011); saying anything more definitive requires a clear understanding of what event or events we are dating.

Differentiation of metal from silicate (the Hf/W system) gives the oldest ages, but this is specifically the age when the metallic phases (for which W has affinity) separated from the silicate phases (for which Hf has affinity). Hafnium decays into tungsten $^{182}\mathrm{Hf} \rightarrow {}^{182}\mathrm{W}$ with a half-life of 9 Ma. When a planet melts, Hf is retained by the silicates while W goes to the core, under the assumption of complete differentiation. Excess $^{182}\mathrm{W}$ thus indicates early core formation (by melting), while Hf/W ratios, normalized to $^{183}\mathrm{W}$, give the age of differentiation.

The disadvantage of the Hf/W chronometer is that it might record core formation in the precursor planetesimals that formed Earth, instead of dating the proposed Moon-forming event, if core and mantle do not remix intimately during giant impacts. If and when we understand the process of Moon formation, and acquire more diverse lunar samples, we will be able to construct model-dependent ages that give more precise answers.

The Moon is no younger than its oldest rocks, the anorthosites. The oldest of these would be the flotation cumulates rafted from the original magma ocean (Ohtake et al. 2009) that if retained on the surface and sampled would date the giant impact. Borg et al. (2011) reported 4.36 Ga for the oldest sampled anorthosite; if we add to this a timescale of crustal solidification of tens of millions of years (Elkins-Tanton et al. 2011), this sets a lower limit of 4.4 Ga for the age of the magma ocean. But the Moon is poorly sampled, and these lunar anorthosites could date remelted magma oceans created by later global-scale collisions during a long and complex history.

The concept of a planetary late veneer evolved alongside the concept of a giant impact, to explain the abundance and diversity of siderophiles in Earth's crust. These can be attributed (Bottke et al. 2010) to massive ($\gtrsim$1,000-km diameter) late collisions, ~4.4–4.2 Ga (well after Moon formation), that added their diverse metallic components. Most of these massive planetesimals would also have borne volatiles, perhaps even bringing a late ocean to Earth. The late veneer has grown popular among dynamicists (e.g., Schlichting et al. 2012) because a leftover population of massive bodies would circularize the orbits of the finished planets, whose eccentricities and inclinations otherwise end up too large in simulations. By modeling the dynamics of highly siderophile elements delivered to Earth's mantle since the last giant impact, and their radioisotopic evolution, Jacobson et al. (2014) have obtained an age of 4.47 Ga, in agreement with the late-forming models.

The Moon has been used to argue against a late veneer, because otherwise the Moon would have accreted substantial iron and water and be more geologically heterogeneous. But a late sweep-up of massive planetesimals would be erosive to the Moon, not accretionary. Consider a zero-velocity planetesimal falling toward the Earth-Moon system. By the time it falls inside 22 $R_\oplus$ it is already traveling faster than lunar v_{esc}. It is therefore difficult to add a veneer to the early Moon, other than by objects from inside the Earth-Moon system (Jutzi & Asphaug 2011) that might plaster themselves onto the crust. So if Earth acquired a late veneer, then the Moon might be deeply eroded. Raymond et al. modeled the delivery of a late veneer to Earth and noted "the possibility that the primordial Moon was more massive than the current one, perhaps by up to 25%" (Raymond et al. 2013, p. 680). Instead of a veneer, we might be looking at the "bones" of the Moon.

3.3. Analogous Satellites

Earth's satellite mass ratio is $M_{☾}/M_\oplus = 0.012$, over 50 times those of the giant planets and ~10^6 times those of Mars. Venus and Mercury have no moons. In searching for analogs we have to go to the dwarf planets beyond Neptune, most famously Pluto-Charon, whose mass ratio is 0.12 and whose origin is also thought to be a planet-scale collision (Dobrovolskis et al. 1997, Canup 2005).

The middle-sized moons of Saturn, 0.04 times the mass of Titan, led Asphaug & Reufer (2013) to consider a giant impact analogous to Moon formation, forming them out of the icy mantles of ~2–3 massive progenitors that accreted to form Titan. The Kuiper belt object Haumea also has significant moons, as well as a dynamical family of ejected bodies, and it was modeled as the outcome of a giant impact by Leinhardt et al. (2010). If the Moon is a giant impact relic, why are its only analogs these icy bodies?

4. PLANET-SCALE COLLISIONS

To evaluate this evidence we need a better understanding of planet-scale collisions, a process whose variants range from effective mergers to hit-and-run collisions. The consequences are remarkably diverse (Asphaug 2010), providing multiple scenarios for Moon formation.

4.1. Angular Momentum

It took decades to appreciate that angular momentum is not simply additive in collisional accretion (Agnor et al. 1999). Like clay on a potter's wheel, it becomes harder to add more once a planet is spinning. If Moon formation was simply a giant cratering event, blasting material from Earth, there would be no isotopic crisis. A much better analogy to the standard model is stellar merger (Rasio & Shapiro 1994), in which the combined angular momentum exceeds what can be bound gravitationally by a single body, so material spills over and escapes.

In a giant impact, the planet is not spun up so much as it accretes high-angular-momentum material that is otherwise passing by. Even for $v_{\rm rel} = 0$ there is always some mass loss, and correspondingly, some angular momentum is partitioned into spiral arms and orbiting and escaping bodies (Asphaug & Reufer 2013). At the most common impact angle $\theta \approx 45°$, the standard model is a graze-and-merge collision, in which Theia bounces into a captured orbit and swings by a second time for a high-angular-momentum merger—a cosmic pinwheel. The phenomenon is like unequal figure skaters pulled into a spin: The bigger skater defines the center of mass, and the legs of the smaller skater get flung to the outside. This is why the Moon is made mostly out of Theia in the model.

The giant impact explanation for Earth-Moon angular momentum (Cameron & Ward 1976) begs the question why Venus, Mercury, and Mars rotate so slowly. Surprisingly, the efficiency of angular momentum accretion by giant impact has not been systematically explored in the manner of mass accretion efficiency (described below). From simulations performed to date it appears (Reufer 2011) that simulated giant impacts of all kinds leave behind final planets spinning no faster than $P_{\rm rot} \gtrsim 4$ h, far shy of the ~2.3-h starting condition of the target Earth in the impact-triggered fission model (Ćuk & Stewart 2012).

4.2. A Panoply of Outcomes

The accretion efficiency is the change in mass of the largest body M_1,

$$\xi = \frac{M_{\rm F} - M_1}{M_2}, \tag{1}$$

where $M_{\rm F}$ is the mass of the largest final body and $M_2 < M_1$ is the mass of the impactor. In a perfect merger, $M_{\rm F} = M_1 + M_2$ and $\xi = 1$. In the standard model, $\xi \approx 0.95$. The colliding mass

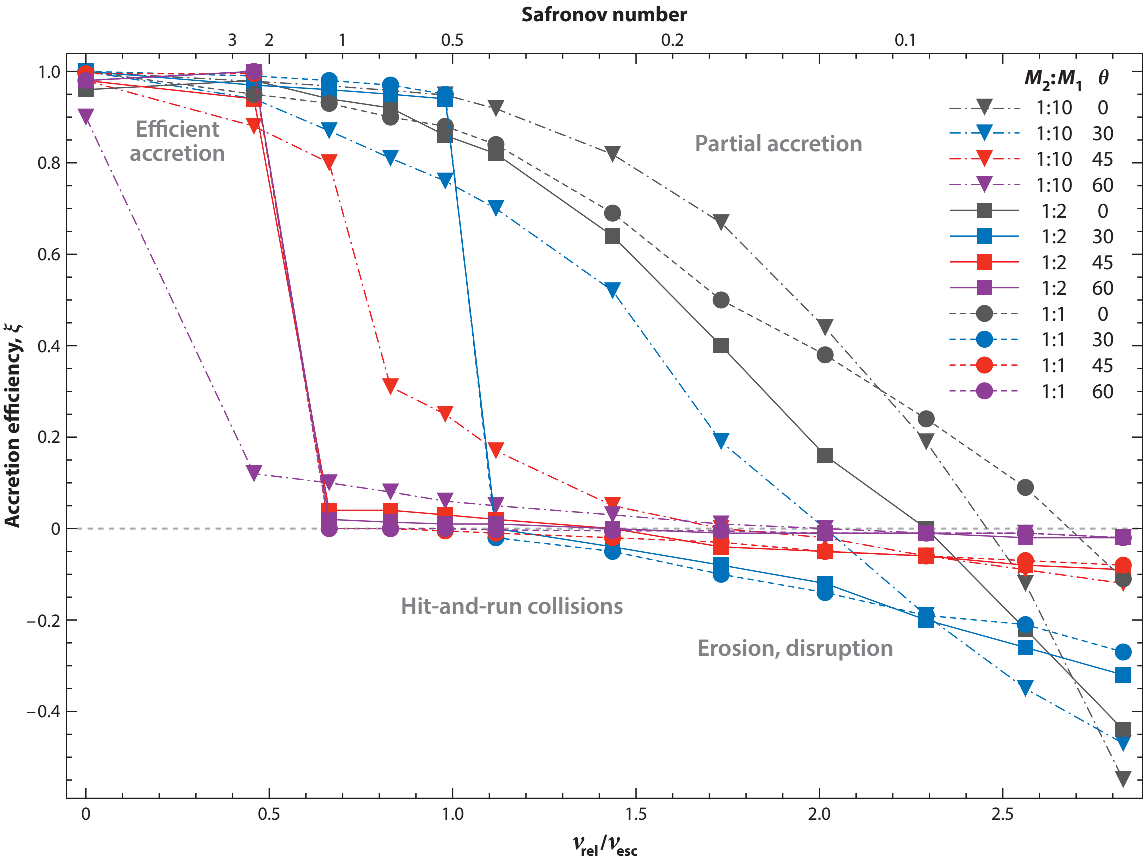

Figure 4

Accretion efficiency $\xi = (M_F - M_1)/M_2$ for giant impacts as a function of v_{rel}, for colliding masses M_2:M_1 and for varying impact angles θ as labeled. Also shown is Safronov number, which is >1 for efficient accretion. Smooth particle hydrodynamics (SPH) simulations were performed by Agnor & Asphaug (2004) and Asphaug (2010) assuming differentiated terrestrial planets with 30 wt% iron and 70 wt% rock. M_1 is the mass of Mars in this plot. Note the abrupt transition from efficient accretion to hit-and-run collision for characteristic velocities $v_{rel} \sim v_{esc}$.

ratio is defined as the ratio of the impactor mass to the total colliding mass,

$$\gamma = \frac{M_2}{M_1 + M_2}. \tag{2}$$

For cratering collisions, not considered here, $\gamma \sim 0$ and $\xi \sim 1$.

Simulations show that for similar-sized collisions (that is, mass ratios $\gamma \gtrsim 0.03$), there are four main categories of outcomes (Asphaug 2010, Stewart & Leinhardt 2012):

- efficient accretion ($\xi \sim 1$) for $v_{imp} \simeq v_{esc}$,
- partial accretion ($0 < \xi < 1$) for intermediate velocity and $\theta \lesssim 45°$,
- hit-and-run ($\xi \sim 0$) collision for intermediate velocity and $\theta \gtrsim 45°$, and
- erosion and disruption ($\xi < 0$) for $v_{imp} \gg v_{esc}$,

where θ is the impact angle. **Figure 4** plots these regimes from simulations; we see that intermediate velocity is $v_{imp} \sim 1.2$–$2.7\ v_{esc}$. Quite significantly, this happens to be the characteristic range

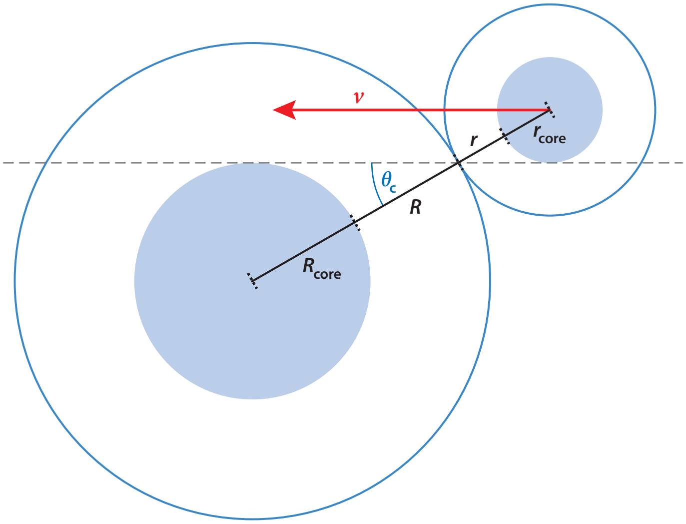

Figure 5

When similar-sized bodies collide, even in relatively head-on collisions, there is a region that "misses" the collision. In high-energy events this leads to a hit and run. In low-energy events a fraction of the impactor can be decelerated from v_{esc} to the orbital velocity, forming a protosatellite. As shown here, the cores of two colliding bodies do not overlap in a collision 30° from vertical.

of impact velocities expected during planet formation (Wetherill & Stewart 1993, O'Brien et al. 2006), requiring us to expect a panoply of collisional outcomes.

Hit and run is especially important, not only because it is common (the clustering near $\xi = 0$ in **Figure 4**) but also because it allows for a new kind of catastrophic disruption not previously considered. In giant impacts, the colliding cores are shielded from impact shocks by thousands of kilometers of rock and can merge gravitationally even while the colliding mantles explode. Once merged, they can be demolished only by the most extreme collisions (Scott et al. 2001). But a more massive body M_1 can dismantle a smaller differentiated planet M_2 (Asphaug et al. 2006), even down to its core, provided larger bodies exist to run into (**Figure 5**). This might explain, by attrition, the iron-rich minor bodies in the Solar System, and perhaps Mercury (Asphaug 2010).

Reufer et al. (2012) propose a hit-and-run scenario for Moon formation by giant impact, carrying Theia away from the scene of the crime and leaving Earth with a hot, massive disk composed mostly of Earth material. It may prove to be the only model capable of satisfying all of the diverse requirements. Although its circumstances seem exotic, hit and run is a typical outcome of planet formation, as shown in **Figure 4**. But the model is not complete until we find out what happens to Theia (e.g., Jackson & Wyatt 2012).

4.3. Energy of Collision

Assuming perfect accretion ($\xi = 1$), and ignoring body and system rotations, a fraction f of the final accumulated gravitational binding energy is available (Asphaug & Reufer 2013),

$$f = 1 - \gamma^{5/3} - (1 - \gamma)^{5/3}, \qquad (3)$$

above the original gravitational binding energy of two well-separated uniform spheres. In the standard model, $\gamma = 0.1$, so $f \sim 14\%$, which is an enormous input of energy (loss of potential energy). In the case of equal-mass planets, $f = 37\%$.

The loss of potential is expressed by postimpact enthalpy that causes a magma ocean and even a globally molten silicate Earth (e.g., Rubie et al. 2007), plus the ejection (in the new frame of reference of M_F) of the protolunar disk and escaping materials. Shock heating is principally responsible for increasing the entropy (temperature) of the colliding materials.

The disk ends up, in simulations of the standard model, as a continuous two-phase fluid, a torus $\sim$2–3 $R_\oplus$ in radial extent and $\sim$1 $R_\oplus$ in vertical extent. In some simulations it is embedded with massive (>1,000 km) clumps, discussed in Section 6. Viscous spreading causes further heating (cf. Stevenson 1987) as part of a longer epoch (years to centuries) when melted and vaporized silicates can accrete only as fast as they radiate to space (Matsui & Abe 1986).

In the standard model, the outer leading hemisphere of Theia "misses" Earth and spawns the disk. Less intensively shocked than the part that collides directly, the disk is relatively cold, $T_{\rm disk} \sim 3{,}000$ K, in characteristic simulations (Canup 2004). By comparison, the hit-and-run scenario described by Reufer et al. (2012) is a higher-velocity event that blasts the Moon from the intensively shock-accelerated material in the proto-Earth; little of Theia is left behind. Their model therefore produces a disk that is (*a*) primarily of Earth composition and (*b*) substantially hotter than in the standard model, at $\gtrsim$10,000 K.

The impact-triggered fission scenario described by Ćuk & Stewart (2012) is also highly energetic, with a much smaller Theia impacting an oblate spinning proto-Earth at $\sim$2–3 $v_{\rm esc}$, colliding retrograde to its presumed $\sim$2.3-h rotation. They also obtain a hot ($\gtrsim$10,000 K) protolunar disk primarily of Earth composition. The twin-impactors hypothesis (Canup 2012) also produces a hot disk, corresponding to the substantial gravitational binding energy $f \sim 0.37$ that is made available by the accretion.

So, because of its geometry, the standard model produces the least-shocked, least-Earth-like initial protolunar disk of the available models. Because diffusion increases strongly with T, hot extensive disks would be an advantage in the context of the diffusion model of Pahlevan & Stevenson (2007), especially as these hotter disks are already better matches isotopically to begin with. From this standpoint they are better models (M.M.M. Meier, A. Reufer, and R. Wieler, manuscript in revision).

It is cautioned that modeled temperatures are uncertain. Furthermore, gravitational energy is released by the final accretion of the protolunar disk; this translates into an additional temperature $T_{\rm acc} \sim \frac{3}{5}\frac{GM_☾}{R_☾ c_{\rm p}} \sim 2{,}000$ K [assuming heat capacity $c_{\rm p} \sim 1$ J/(g K)] that must be radiated away for accretion to commence. So the process is thermally, not gravitationally or collisionally, limited at first. The radiative cooling time of an optically thick protolunar disk is $\sim$10–100 years (Stevenson 1987). Once cooled and coagulated into massive bodies, the disk becomes transparent and radiates effectively; then, the clumps might coagulate in months (Ida et al. 1997).

4.4. Timescales of Collision

A giant impact occurs over hours to days, tens of gravitational timescales $\tau_{\rm grav} \sim (G\rho)^{-1/2}$, and several dynamical self-crossings $\tau_{\rm coll} \sim 2(R_1 + R_2)/v_{\rm imp}$. What happens next depends on the relative energy of the event, the Safronov number $\boldsymbol{\Theta} = \frac{1}{2}(v_{\rm esc}/v_{\rm rel})^2$. Efficient accretion occurs in low-energy collisions ($\boldsymbol{\Theta} > 1$), and these must ultimately be dominant for planets to grow.

In low-energy collisions such as the standard model, the core of the projectile M_2 spirals through the target mantle toward the center, while the target M_1 gets encircled by extensive tidal arms derived primarily from the silicate mantle of M_2. Details depend on the mass ratio, velocity, impact parameter, and initial rotations. Over the next few days, strands and massive clumps captured into orbit are shredded when they pass inside Earth's Roche limit. In days to months, the orbiting melt and vapor become a protolunar disk.

We do not yet know whether to expect massive clumps in the protolunar disk (see Section 6), but we expect it to be composed of melted and vaporized silicates and a percentage of iron. For the endgame of lunar accretion, the key questions raised by Stevenson (1987) are still unanswered. If there were massive clumps, then the Moon would rapidly accrete. If the disk began as a foamy diffusive cloud of microdroplets (Thompson & Stevenson 1988), cooling and condensation might take centuries and be further frustrated by viscous spreading and the associated heating.

The latest detailed modeling (Salmon & Canup 2012) revealed some interesting possibilities. In particular, the authors found that interior disk material could have been hot enough, and in small enough sizes, to equilibrate isotopically with Earth's silicate vapor atmosphere, thereafter migrating out to be accreted by the mostly finished Moon. They argued that this could mask Theia's isotopic signature beneath Earth-equilibrated materials, hiding the alien material.

Jutzi & Asphaug (2011) also considered burying the Moon beneath low-velocity materials, in this case a ~1,000-km companion satellite that becomes unstable (Ćuk & Gladman 2009) and accretes onto one side. But their model involves a Trojan moon, formed alongside the Moon out of the same protolunar disk, so it would not affect the isotopic chemistry, or perhaps even the bulk chemistry. This proposed event would play out tens of millions of years after the giant impact, putting icing on the cake.

Trojan points: dynamically stable points that lead and trail a body in its orbit by ~60°

5. EVOLUTION OF IDEAS

The giant impact hypothesis is so intimately connected to the earlier hypotheses for Moon formation that it is useful to recap its variants in the context of each scenario: coaccretion, rotational disruption, and capture. For example, the standard model is a special instance of collisional capture. I then consider how geochemical models are able to constrain these variants, emphasizing the importance of obtaining new suites of lunar samples.

5.1. Coaccretion

The idea that Earth and the Moon coaccreted as a binary pair was refined by Morishima & Watanabe (2001) in the context of the waning solar nebula. But W isotopes and other chronometers (Touboul et al. 2007) have shown that the Moon must have formed long after the disappearance of the gas, at ~4.5 Ga or later. Without the nebula, coaccretion appears impossible to support dynamically.

Another idea—another kind of coaccretion—is that Theia formed at one of Earth's Trojan points (Belbruno & Gott 2005) or otherwise in the same feeding zone near 1 AU. Coformation of some kind is consistent with the low-velocity collision that is required of the standard model. But even if the physical conditions could be satisfied to make a Mars-mass Trojan, it might not be sufficient to make Theia in the same feeding zone as Earth, or in a Trojan point of Earth, to give it indistinguishable isotopes. A Trojan point exists only after Earth is substantially accreted, so it would be a depleted and much less massive region of the disk. It remains to be demonstrated how a Trojan Theia or a nearby Theia, forming at lower pressures, fugacities, and temperatures and with different boundary conditions, would have the same isotopes.

5.2. Rotational Disruption

Darwin (1879) supposed that Earth and the Moon might once have been a single planet spinning with a period of $\lesssim$5 h. Although not nearly fast enough to result in direct fission (this requires over twice the Earth-Moon angular momentum; Chandrasekhar 1969), tides raised on Earth by the Sun every ~3 h might resonate, Darwin supposed, with Earth's free oscillations. The resulting instability might break the Moon away, perhaps forming the Pacific Ocean. This prescient idea proved dynamically impossible.

O'Keefe & Sullivan (1978) postulated that rapid differentiation changed Earth's moment of inertia, so that an already rapidly spinning planet superrotated, launching a debris ring of silicates. This also requires rapid initial rotation and an undifferentiated initial planet. One possible way to achieve both of these things is via a giant impact, if the impact energy is sufficient to emulsify the core into the mantle and give the postimpact planet a maximal moment of inertia. An emulsified, homogenized planet could in principle accrete substantially more angular momentum than could a completely differentiated planet, and thereafter the abrupt settling out of iron could trigger mass loss and Moon formation.

Although it solves the isotopic crisis by homogenizing the reservoirs, this model conflicts with the evidence that iron segregation occurred much earlier than Moon formation (Rudge et al. 2010). The silicate reservoirs need to be homogenized, but the whole planets cannot be. But the standard model is not that far removed from this scenario: two accreting planets going from high to low moment of inertia as their cores merge, transferring angular momentum to the spiral arms that form the Moon, although evidently without much iron-silicate mixing. For now, the timing and pace of postimpact iron-silicate segregation, and the influence of this segregation on accretionary dynamics, are the subject of ongoing study (e.g., Dahl & Stevenson 2010).

Impact-induced fission (Ćuk & Stewart 2012) is another variation of rotational disruption. It is actually a sequence of models: (*a*) spin up of Earth to near disruption, (*b*) fission of Earth's mantle, triggered by a smaller and more energetic Theia, and (*c*) orbital evolution of the Earth-Moon system to lose half its angular momentum to evection. The model's key advantage is explaining a Moon that forms, as in Darwin's theory, mostly from Earth. Theia is small, and the impact velocity is great, so most of Theia's mantle escapes.

One criticism of the theory concerns its final step, discussed above, because the transfer of $\vec{L}_{\mathrm{EM}}$ to the Sun requires special conditions. Another concern is that Earth must spin at $P_{\mathrm{rot}} \sim 2.3$–2.7 h in this model. Rotation periods reported by Agnor et al. (1999) do range faster than ~1–2 h, but the emphasis of that study was that perfect accretion is an invalid assumption (Agnor & Asphaug 2004). Although systematic explorations have not been concluded, no reported giant impact aftermath has a spin period faster than 4 h.

5.3. Capture

Another venerable idea (Urey 1952) is that the Moon was a wayward planet, captured by Earth and circularized into stable orbit. Dynamical theories of capture include those of Gerstenkorn (1955), described above, and Goldreich (1966). These theories require a nebula that aids the capture and then goes away. Even if plausible, capture does not explain why the Moon is such a dramatically unique planetary body—how it got to be that way, and why Earth did not instead capture a common protoplanet.

Because the nebula is likely to have vanished within ~1–10 Ma, it would not be available to do the work of capture, given the recognized age of the Moon. But a collision with Earth can decelerate a planet, to the point of capturing all or part of it into orbit. Matsui & Abe (1986) proposed that a collision led to the production of a swarm of captured planetesimals eventually accreted onto Earth. They proposed that this might have delivered the ocean and atmosphere. A capturing collision is not far from the standard model, and we now recognize that a giant impact is the only way to capture another planet, or parts of it, into orbit. This brings us full circle to the quandary of how lunar isotopic composition can be indistinguishable from that of Earth, which is equivalent to asking, Where is Theia?

5.4. Where Is Theia?

The deep mantle and core of Theia are added to the proto-Earth in the standard model, coming to rest along with a half-dozen comparable planets, dozens of Moon-sized embryos, and thousands of Vestas. Whether or not the standard model explains the Moon, these kinds of low-v_{rel} events are well represented among the collisions that are thought to characterize late-stage accretion (e.g., O'Brien et al. 2006). So we can certainly consider whether accreted volumes of Earth survive as discrete deep reservoirs (Mukhopadhyay 2012) and whether they might be related to the formation of the Moon.

Hydrocode: time-stepped integrations of the partial differential equations of fluid mechanics, including material equations of state, shocks, and gravity

In hydrocode simulations, the iron cores of the two planets agglomerate without much mixing of silicate with iron, although immiscibility may be numerically exaggerated. Nimmo & Agnor (2006) modeled Hf/W evolution and found evidence in the isotopic heterogeneity data for core-mantle mixing in most collisions, whereas Dahl & Stevenson (2010) found that only a fraction of Earth's core would equilibrate with silicate during a giant impact. Mixing might also occur thousands to millions of years after the giant impact, by intense geodynamical evolution (Golabek et al. 2011) aided by irregular heating and rapid rotation. Gravitational and possibly magnetic interactions between the rapidly rotating Earth and a very nearby Moon would further stir things up (Garrick-Bethell et al. 2009).

Considering the more energetic, higher-angular-momentum solutions [impact-triggered fission (Ćuk & Stewart 2012) and hit-and-run collision (Reufer et al. 2012)], Theia's mantle either escapes or becomes part of Earth's postimpact silicate atmosphere. In several of the models of Reufer et al. (2012), most of Theia continues downrange, in which case fragments of Theia might be expected in the asteroid belt, although Earth would be expected to accrete the majority. Cameron (2000) considered equal-mass proto-Earths colliding, in a low-velocity accretion, but the scenario was tabled because it produced a system with double the allowable angular momentum. The idea was revisited by Canup (2012), after the dynamical constraint was lifted by Ćuk & Stewart (2012), to solve the isotope problem in a different way: making the Moon and Earth out of equal fractions of each colliding body, ensuring isotopic similarity.

5.5. Testing Hypotheses

Each of these scenarios has testable consequences for geochemistry. Meier et al. (M.M.M. Meier, A. Reufer, and R. Wieler, manuscript in revision) considered the final fractions of Theia that end up in the Moon and in Earth (see also Wiechert et al. 2001) in published simulations. For example, in the standard model, ~1/10 of the finished Earth and ~4/5 of the Moon are made of Theia. From these fractions, isotopic trends are computed assuming complete silicate mixing of the planetary components inside each finished body.

Requiring the isotopes to be within the ranges of measurement places formal limits on the bulk composition of Theia. The analysis illustrates the formidable leverage that is available to constrain the problem. However, as noted above, one must question the validity of the geochemical assumption of perfect mixing, and must recognize that a very wide range of scenarios have yet to be evaluated—the forest of lines on the plot in **Figure 6**.

Only a very Earth-like Theia satisfies the standard model (**Figure 6**). Impact-triggered fission (Ćuk & Stewart 2012) fares better, if Theia's oxygen reservoir is similar to that of enstatite chondrites or of Mars. The hit-and-run model of Reufer et al. (2012) can be fitted by an enstatite chondrite Theia followed by an oxidized component. The theory of equal-mass planets (Canup 2012) works well for mass ratios very close to $\gamma = 0.5$, so long as the high angular momentum is not a problem. For unequal mass ratios, the Moon forms predominantly from the smaller body

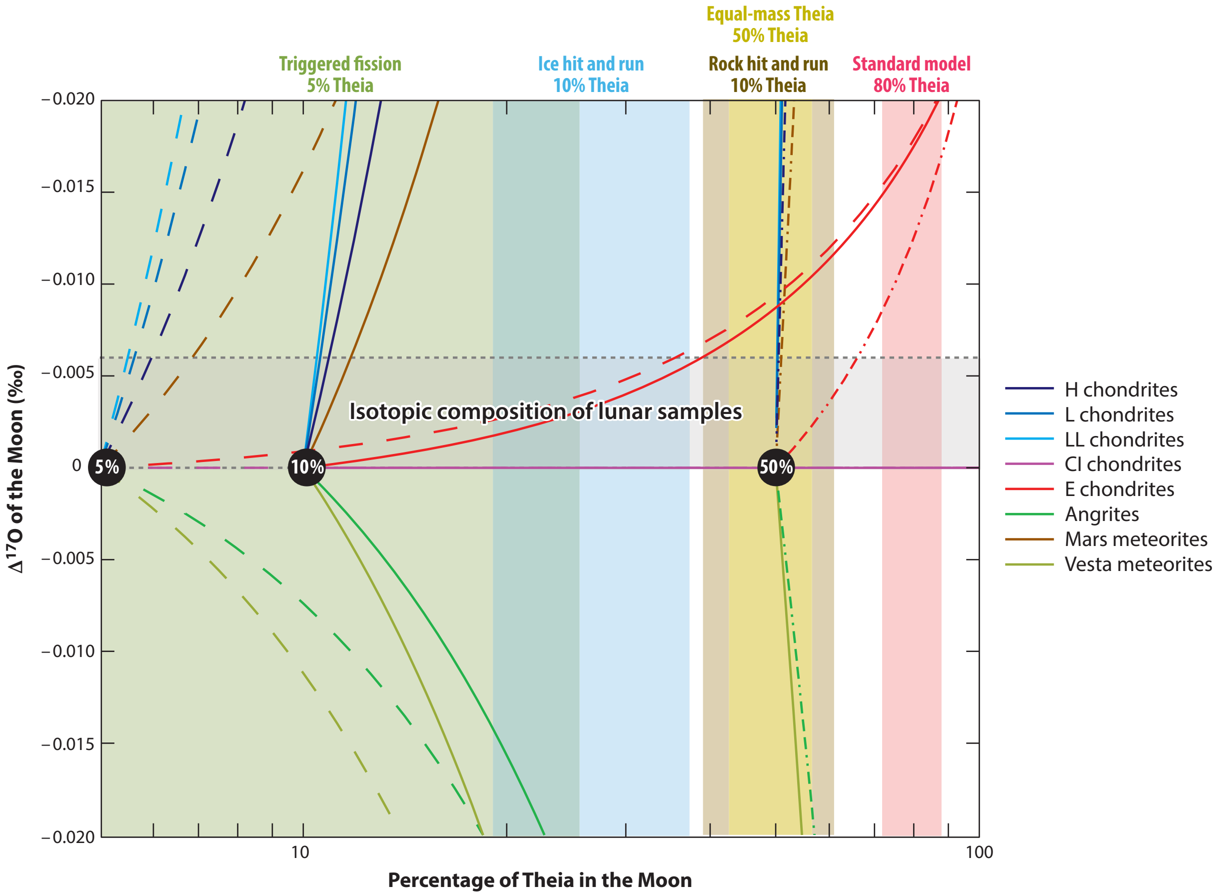

Figure 6

Oxygen isotopic composition of the Moon ($\Delta^{17}O$, in ‰) computed as a function of the percentage of Theia-derived material in the Moon and Earth, assuming complete mixing and no subsequent compositional exchanges. The gray band $\Delta^{17}O = 0.003‰ \pm 0.003‰$ defines the 2σ range of lunar samples (Wiechert et al. 2001). Black circles are starting points for Theia contributing 5%, 10%, and 50% of the Moon's final mass. Mixing lines emanating from each circle are for varying percentages of Theia in the Moon, for meteorite analogs as labeled. The proportion of Theia silicates found in the Moon is shaded by color for each scenario: green (triggered fission, 5% Theia; Ćuk & Stewart 2012), blue (ice hit and run, 10% Theia; Reufer et al. 2012); brown (rock hit and run, 10% Theia; Reufer et al. 2012); yellow (equal-mass Theia, 50% Theia; Canup 2012); and pink (standard model, 80% Theia; Canup & Asphaug 2001). Adapted from Meier et al. (M.M.M. Meier, A. Reufer, and R. Wieler, manuscript in revision) with permission.

for the angular momentum considerations described above, and so the isotopic signatures would be readily distinguished unless the planets were the same to begin with.

6. SIMULATIONS

The study of planetary collisions has evolved into a complex problem in computational fluid dynamics, so I conclude with a snapshot of the trials and tribulations. Unlike impact cratering, a planet-scale collision has no central locus, so the formalisms that lead to crater scaling relationships (e.g., Housen et al. 1983) do not apply. Progress in the field relies on hydrocodes, which are used to solve the continuum equations governing mass, momentum, and energy plus shocks and self-gravity. Like any remarkable tool they must be applied with caution, lest compelling visualizations take the place of argument.

A hydrocode solution is an approximation, incrementing from initial conditions at t_0 forward by dt according to the accuracy requirements of the integrator. Equations of state relate pressure to internal energy and density, closing the system of equations. Shock waves, represented as an instantaneous rise in energy, density, and pressure, are fit by a piecewise analytical solver or (more commonly) smeared over several resolution elements to reproduce the Hugoniot jump conditions, using artificial viscosity to suppress the associated numerical instability at the expense of resolution. Self-gravity is computed by adding up the pairwise attractions of N discretized elements, although an N^2 computation is usually approximated using an $N \log N$ "tree" approach. Also, temperature and pressure within the initial planets must be equilibrated gravitationally prior to a simulation, a task that can require significantly more computational effort than modeling the collision.

These numerical laboratories give us a remarkable view inside colliding planets (**Figure 7**), improving our intuition and allowing for validation and prediction. The most popular tool for this research has been smooth particle hydrodynamics (SPH), in part because the method conserves mass and angular momentum precisely—key quantities in modeling planet-scale collisions—and because it does not use a grid. This makes it well suited to the study of global-scale collisions expanding into space. For twenty years the code of Benz et al. (1986) was used by most groups who worked on the problem; other methods have given mostly consistent results (Wada et al. 2006, Canup et al. 2013), with important caveats.

Discretization and uncertain initial conditions introduce errors that can grow with every dt into major discrepancies. Although it is possible to demonstrate mathematically that a computational error will not grow exponentially (this controls dt), in practice, validation campaigns comparing simulations to laboratory experiments, theoretical solutions, and field observations are required (Pierazzo et al. 2008). Just as in traditional empirical research, we must be aware of situations in which error can dominate the signal.

Consider the problem of gravitational clumping, which is of great importance to Moon formation because the protolunar disk accretes rapidly around any massive "seeds." Early simulations obtained an immediate Moon-mass clump in the tail of the disrupted Theia (the cold Moon scenario). Higher-resolution simulations from the same initial conditions resulted in a sheared-apart disk with fewer, much smaller clumps. Artificial clumping is a recognized aspect of under-resolved SPH simulations, and the earliest models used only $N \sim 3{,}000$ particles in total, so only a few dozen particles for the proto-Moon.

Canup et al. (2013) found broad similarities across methods (SPH versus CTH, a grid-based hydrocode), as did Wada et al. (2006) using ZEUS (another grid-based hydrocode), but only for the mass and angular momentum of the captured protolunar disk. For clumping, there is no sense of convergence when resolution is increased or methods are compared. Furthermore, Wada et al. (2006) obtained powerful disk shocks, not seen in SPH simulations, that transport angular momentum and cause rapid spreading that might frustrate lunar formation.

Equations of state can be another source of error. The simple ones miss out on phase transformation, which can be vital to the outcome. The more complicated ones can be the trickiest part of a simulation. For example, thermodynamic variables can run off the end of a precomputed table, or an iteration can fail to converge (plummeting Earth's core to 0 K in one published study). Tricks done for expediency, especially the application of a low density cutoff or a sound speed minimum, can suppress shocks.

Because of the intense deformation, grid-based hydrocodes require an advective step, in which material moves across a grid; in this case angular momentum is poorly conserved. Also, a moving

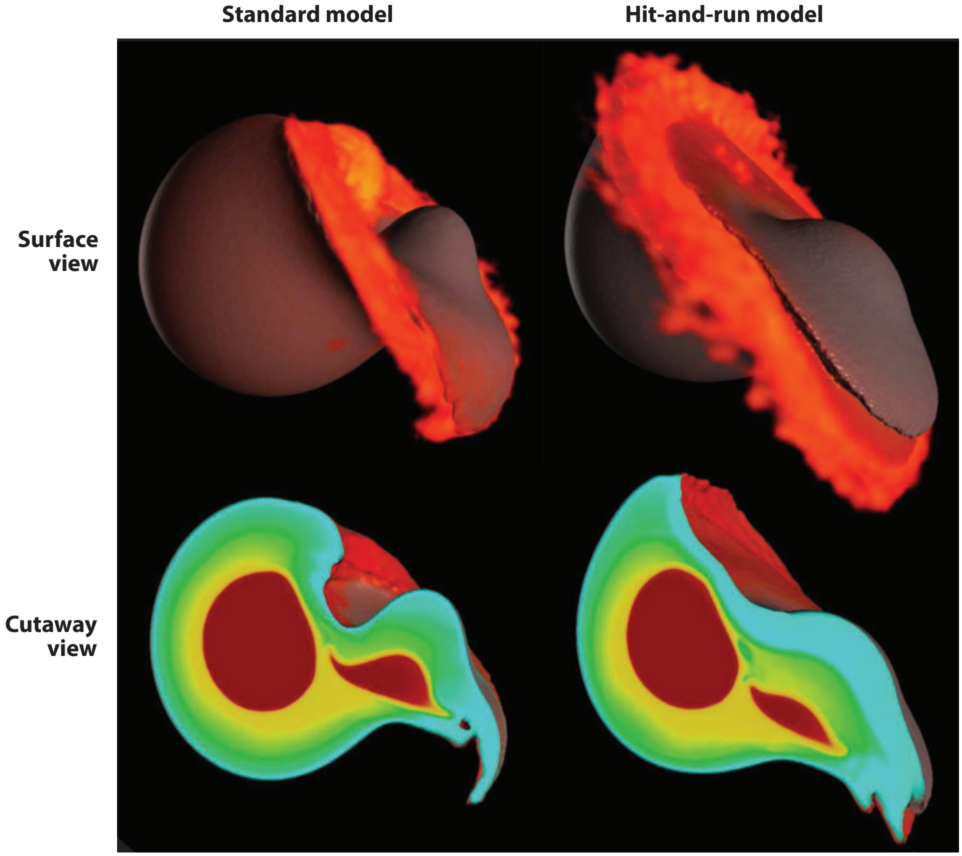

Figure 7

The standard model (*left*) and a hit-and-run model (*right*), 20 min after initial contact, in surface (*top*) and cutaway (*bottom*) views. In the standard model, $v_{\mathrm{imp}} \approx v_{\mathrm{esc}}$, and both bodies have 30 wt% Fe and 70 wt% SiO_2 (run cA08 in table 1 of Reufer et al. 2012). The result is an effective merger. In this version of the hit-and-run scenario, $v_{\mathrm{imp}} \approx 1.5v_{\mathrm{esc}}$ and the impactor is a water-rich planet with 15 wt% Fe, 35 wt% SiO_2, and 50 wt% H_2O (run iA13 in table 1 of Reufer et al. 2012). The surface views are isodensity plotted at $\rho = 1$ (*gray*), with lower-density material (*red*, *yellow*) colored by temperature. The cutaway views illustrate a slice through the symmetry plane, showing material $\rho > 1$ g/cm^3 colored by density. Courtesy of A. Emsenhuber, M. Jutzi, W. Benz, and A. Reufer (University of Bern).

planet gets smeared across space before the collision, and its core and mantle can become detached. Energy conservation is also problematic during rezoning, and numerical density fluctuations can be very problematic in a solid or liquid equation of state. Furthermore, the zeroing out of negative pressures can have unexpected consequences.

Given the pitfalls, it is important to run both kinds of simulations (Canup et al. 2013). Higher spatial resolution and finer time resolution can reduce these errors, but a 10× increase in resolution, in 3D, will require twenty years of advancement in computing at the current pace. Sadly, we are quite far from reliably simulating Moon formation as an end-to-end sequence of events. Determinations of accretion efficiency (**Figure 4**) are more robust, concerning only the coldest, least-accelerated fractions of the colliding matter. Accurately modeling Moon formation requires high numerical accuracy over the hottest and most tenuous percentile of the calculation, over the collisional timescale of days, followed by equally detailed and accurate calculation of the dynamical, thermal, and radiative evolution for perhaps thousands of years. That is a tall order.

7. SUMMARY OF MODELS

The standard giant impact model is a variant of capture, forming the Moon predominantly out of the Mars-sized accreted body. It is also fission, except that it is Theia that is fissioned by Earth (a graze-and-merge collision). It is also coaccretion, if Theia must derive from the same isotopic reservoir as Earth. So there is something for everyone.

But the standard model led us to the isotopic crisis. About a half-dozen principal new theories respond to that crisis, as summarized below. It is telling that each is based around a planet-scale collision; this most basic conclusion appears incontrovertible, although critics might say entrenched. They are listed below in order of their departure from the standard model, along with the major challenges they face.

- The standard model (mass ratio $\gamma \sim 0.1$, impact angle $\theta \sim 45°$, $v_{\rm imp} \sim v_{\rm esc}$), but with Theia being a close sibling of the proto-Earth (Belbruno & Gott 2005):
  - It may be impossible to make a Mars-sized Theia at one of Earth's Trojan points, and to make it isotopically indistinguishable.
- The standard model, followed by massive diffusion of the Earth and protolunar isotopic systems (Pahlevan & Stevenson 2007):
  - The required isotopic diffusion may be self-limiting due to mass and angular momentum transfer.
  - The highly refractory elements must also be homogenized (Zhang et al. 2012).
  - Isotopes must be homogenized without homogenizing water and semivolatile abundances (Desch & Taylor 2013).
- The standard model, followed by accreting layers of Earth-equilibrated interior disk material that bury an otherwise distinctive Moon (Salmon & Canup 2012):
  - Material would have been reprocessed by the lunar crust and mantle.
  - There is no evidence for a hidden lunar interior.
  - If there were late veneer collisions, then the original crust and upper mantle might be lost (Raymond et al. 2013), exposing the deep inside.
- A hit-and-run collision, with Theia escaping (Reufer et al. 2012):
  - The disk still ends up polluted by Theia, $\gtrsim$20% in the best case.
  - Escaping remnants of Theia would reaccrete onto Earth (what are the isotopic consequences?), or else disperse (are there remnants of Theia in the asteroid belt?).
- Energetic impact by a high-velocity icy plutoid (cf. Reufer et al. 2012):
  - No dynamically consistent solution has been found.
  - This scenario leaves the Moon with no iron.
- Merger of equal proto-Earths, with the Moon and Earth forming half from each (Cameron 2000, Canup 2012):
  - This produces twice the present $\vec{L}_{\rm EM}$; it is also the most energy-producing collision ($f \sim 0.37$).
  - Small (~5%) departure from equal mass results in a substantial isotopic asymmetry.

- A spinning oblate proto-Earth, close to rotational instability, followed by impact-triggered fission (Ćuk & Stewart 2012):
  - It might prove impossible for the target Earth to acquire ~2.3–2.7-h initial spin.
  - Afterwards the scenario requires substantial loss of $\vec{L}_{\mathrm{EM}}$.

Combinations lead to dozens of potentially viable scenarios. A hit-and-run collision might leave behind a faster-spinning proto-Earth plus a mantle-stripped escaping Theia that might accrete in a second Moon-forming collision. An icy plutoid could impact a fast-spinning proto-Earth. Hit and run might be followed by energetic diffusion.

Some maintain that the giant impact theory is doomed. My concern is different: that several of these scenarios will, with suitable modifications, end up satisfying the physical, geochemical, petrological, and dynamical constraints, and that we thus might never know how the Moon formed. But I take solace that in trying to find out, we will at least establish the very creative nature of the cataclysmic events that made the planets.

8. CONCLUSIONS

> *However great the probability of what I have advanced on the formation of the planets and their satellites . . . I do not pretend to convince the incredulous.*
>
> —Buffon, *Histoire Naturelle*, Vol. 1 (1749)

It is well established that the terrestrial planets finished growing in a late stage of pairwise accretionary collisions, also known as giant impacts. Being fundamentally stochastic, these collisions explain the compositional diversity of planets and their dynamical states. They are a natural final outcome of postnebula accretion. The giant impact hypothesis for the origin of the Moon is consistent with this principle but is by no means a corollary, and it faces considerable obstacles: Among terrestrial planets, why does only Earth have a sizable Moon? Why is the Moon isotopically indistinguishable from Earth?

8.1. Reconciliations Abound

Belbruno & Gott (2005) proposed that Theia is a close dynamical relative of the proto-Earth, sharing its chemistry. Pahlevan & Stevenson (2007) argued for wholesale diffusive exchanges between Earth and the hot lunar disk. Salmon & Canup (2012) proposed that an alien Moon is hidden beneath Earth-equilibrated debris that accreted after the Moon solidified. In these cases the geochemical record of Theia might be lost.

Others change the physical parameters. Reufer et al. (2012) explained the Moon as a hit-and-run collision by a somewhat faster Theia, whose mantle continues downrange after dredging up a much hotter, Earth-rich silicate disk. Ćuk & Stewart (2012) modeled an energetic impactor into an oblate proto-Earth that is already spinning close to disruption; Theia triggers fission from the equator, and its silicates are not retained. The resulting system would then have to be spun down. In these cases the dynamical record of Theia is also lost.

Impact-triggered fission and hit and run are much more energetic than the standard model, producing a much hotter protolunar disk that is considerably more Earth-like in initial composition. It remains to be seen whether they provide sufficiently accurate solutions to the geochemistry, but meanwhile they take us into uncharted dynamical territory and introduce new problems. Impact-triggered fission requires the loss of substantial angular momentum and a target Earth spinning

10 times faster than Earth spins today. Variations on hit and run are favored geochemically (M.M.M. Meier, A. Reufer, and R. Wieler, manuscript in revision), but they leave us wondering what happened to Theia.

8.2. Onward and Upward

Astronomers will someday put these hypotheses to the test by detecting Moon-like objects orbiting Earth-like planets around Sun-like stars. More immediately, we should remedy the sparse and incomplete sampling of lunar rocks, which are currently limited to nearside landing sites and random meteorites (Joy & Arai 2013). When modelers press on without sufficient data, it is a perilous and unmarked road, and science becomes reliant on simplifications and simulations. With so many new hypotheses on the table, and insufficient data available to choose from among them, the result is "a lumber-room of untested hypotheses" in need of a "spring-cleaning and bonfire" (Jeffreys 1929, p. 177).

New approaches to observation and exploration are ongoing. Missions to the Moon have flown from Russia, the United States, Japan, Europe, China, and India. Advances in microsatellite propulsion and communication have led universities and private consortia to consider their own lunar missions. China's Chang'e 3 last December was the first soft landing on the Moon since Russia's Luna 24 in 1976, which was the last lunar sample return.

Lunar exploration is sometimes criticized as something we can't afford. But Apollo paid for itself several times over with a stimulated economy (Bezdek & Wendling 1992), and the resources spent on the race to the Moon would have been burned up on other cold war technology had our forebears not been bold and imaginative. Those missions gave us the hope to answer the question, How was Earth created? The Moon hangs like a dot at the end of the question mark.

SUMMARY POINTS

1. The Moon is isotopically indistinguishable from Earth, presenting a grave challenge to the giant impact theory. Although the theory still stands, it is on uncertain footing.
2. The Moon is not blasted from Earth in the standard model of the giant impact. It is mostly a collisionally captured remnant of Theia's mantle, hence the isotopic crisis.
3. The standard model is the least energetic of the giant impact models, forming a much lower-temperature disk ($\sim$3,000 K) compared with those in competing models ($\sim$10,000 K), and it produces the most Theia-like Moon.
4. The standard model can be fixed, perhaps, by giving Theia and Earth a common origin or ancestor, or by plastering the Moon with Earth-equilibrated material, or by homogenizing Earth and the protolunar disk.
5. If Earth-Moon angular momentum was lost during evection, then previously discarded hypotheses are on the table (fission of a fast-spinning Earth, accretion of twin proto-Earths) while new ideas (like hit and run) can venture forth without this constraint.
6. A faster Theia might have continued downrange after dredging up a hot disk of primarily Earth-derived material. The hit-and-run scenario offers substantial advantages over other theories, although the ultimate fate of Theia requires study.
7. Hydrodynamic simulations are approximate at best and limited in scope, but they are of enormous utility in developing hypotheses, inspecting processes and outcomes, and validating models. They require ongoing connections to geochemistry and theory.

FUTURE ISSUES

1. The fundamental importance of Moon formation to so many aspects of Earth and planetary science requires that we remove the research bottlenecks by acquiring diverse new samples and obtaining detailed seismic imaging of lunar structure.
2. The giant impact hypothesis has weathered the isotopic crisis by morphing into new variants, the best of which still require considerable testing and validation.
3. Although there is substantial agreement on what happens during the first few hours of a planet-scale collision, we do not adequately understand how it spawns a disk and how a massive hot torus of silicate droplets and vapor becomes a satellite.
4. Linked models are needed that accurately compute the thermodynamic history of disk-forming material, volatile transport and loss, isotopic fractionation, and lunar coagulation.
5. The postimpact Earth, an irregular oscillating spheroid, would have driven the earliest dynamical evolution of the protolunar disk in powerful ways that have not been modeled.
6. If present-day Earth-Moon angular momentum is no longer a constraint on the giant impact, then the parameter space (preimpact rotations, impact vectors, planetary sizes and compositions) is enormous and awaits exploration.
7. Seismic imagery and geochemistry of Earth may reveal the record of planet-scale accretion in the core-mantle region, perhaps even the recognition of Theia deep below.
8. Astronomers have detected giant impact debris in planetary systems around other stars. In coming decades we might discover another moon, orbiting another Earth-like planet, giving us a direct way of answering the hardest of these questions.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I am grateful to NASA's Planetary Geology and Geophysics Program, which has sponsored my research into planet-scale collisions, and to valued colleagues for their illumination, advice, and graphical contributions. The time and effort to write this review was made possible by support of Arizona State University as their Ronald Greeley Chair of Planetary Science.

LITERATURE CITED

Agnor C, Asphaug E. 2004. Accretion efficiency during planetary collisions. *Astrophys. J.* 613:L157–60. doi: 10.1086/425158

Agnor CB, Canup RM, Levison HF. 1999. On the character and consequences of large impacts in the late stage of terrestrial planet formation. *Icarus* 142:219–37. doi: 10.1006/icar.1999.6201

Ahrens TJ. 1993. Impact erosion of terrestrial planetary atmospheres. *Annu. Rev. Earth Planet. Sci.* 21:525–55. doi: 10.1146/annurev.ea.21.050193.002521

Asphaug E. 2010. Similar-sized collisions and the diversity of planets. *Chem. Erde Geochem.* 70:199–219. doi: 10.1016/j.chemer.2010.01.004

Asphaug E, Agnor CB, Williams Q. 2006. Hit-and-run planetary collisions. *Nature* 439:155–60. doi: 10.1038/nature04311

Asphaug E, Reufer A. 2013. Late origin of the Saturn system. *Icarus* 223:544–65. doi: 10.1016/j.icarus.2012.12.009

Belbruno E, Gott JR III. 2005. Where did the Moon come from? *Astron. J.* 129:1724–45. doi: 10.1086/427539

Benz W, Cameron AGW, Melosh HJ. 1989. The origin of the Moon and the single-impact hypothesis III. *Icarus* 81:113–31. doi: 10.1016/0019-1035(89)90129-2

Benz W, Slattery WL, Cameron AGW. 1986. Short note: snapshots from a three-dimensional modeling of a giant impact. See Hartmann et al. 1986, pp. 617–20

Bezdek RH, Wendling RM. 1992. Sharing out NASA's spoils. *Nature* 355:105–6. doi: 10.1038/355105a0

Borg LE, Connelly JN, Boyet M, Carlson RW. 2011. Chronological evidence that the Moon is either young or did not have a global magma ocean. *Nature* 477:70–72. doi: 10.1038/nature10328

Bottke WF, Walker RJ, Day JMD, Nesvorny D, Elkins-Tanton L. 2010. Stochastic late accretion to Earth, the Moon, and Mars. *Science* 330:1527–30. doi: 10.1126/science.1196874

Bouvier A, Wadhwa M. 2010. The age of the Solar System redefined by the oldest Pb-Pb age of a meteoritic inclusion. *Nat. Geosci.* 3:637–41. doi: 10.1038/ngeo941

Brush SG. 1978. A geologist among astronomers: the rise and fall of the Chamberlin–Moulton cosmogony, part 1. *J. Hist. Astron.* 9:77–104

Buffon GLL. 1749. *Histoire Naturelle, Générale et Particulière*, Vol. 1. Deux-Ponts, Ger.: Sanson

Cameron AGW. 2000. Higher-resolution simulations of the giant impact. See Canup & Righter 2000, pp. 133–44

Cameron AGW, Benz W. 1991. The origin of the Moon and the single impact hypothesis IV. *Icarus* 92:204–16. doi: 10.1016/0019-1035(91)90046-V

Cameron AGW, Ward WR. 1976. The origin of the Moon. *Lunar Planet. Sci. Conf. Abstr.* 7:120–22

Canup RM. 2004. Dynamics of lunar formation. *Annu. Rev. Astron. Astrophys.* 42:441–75. doi: 10.1146/annurev.astro.41.082201.113457

Canup RM. 2005. A giant impact origin of Pluto-Charon. *Science* 307:546–50. doi: 10.1126/science.1106818

Canup RM. 2012. Forming a Moon with an Earth-like composition via a giant impact. *Science* 338:1052–55. doi: 10.1126/science.1226073

Canup RM, Asphaug E. 2001. Origin of the Moon in a giant impact near the end of the Earth's formation. *Nature* 412:708–12. doi: 10.1038/35089010

Canup RM, Barr AC, Crawford DA. 2013. Lunar-forming impacts: high-resolution SPH and AMR-CTH simulations. *Icarus* 222:200–19. doi: 10.1016/j.icarus.2012.10.011

Canup RM, Righter K, eds. 2000. *Origin of the Earth and Moon.* Tucson: Univ. Ariz. Press

Canup RM, Ward WR, Cameron AGW. 2001. A scaling relationship for satellite-forming impacts. *Icarus* 150:288–96. doi: 10.1006/icar.2000.6581

Chamberlin TC. 1916. The planetesimal hypothesis. *J. R. Astron. Soc. Can.* 10:473

Chandrasekhar S. 1969. *Ellipsoidal Figures of Equilibrium (the Silliman Foundation Lectures).* New Haven, CT: Yale Univ. Press

Ćuk M, Gladman BJ. 2009. The fate of primordial lunar Trojans. *Icarus* 199:237–44. doi: 10.1016/j.icarus.2008.10.022

Ćuk M, Stewart ST. 2011. The puzzle of lunar inclination. *EPSC Abstr.* 6:580

Ćuk M, Stewart ST. 2012. Making the Moon from a fast-spinning Earth: a giant impact followed by resonant despinning. *Science* 338:1047–52. doi: 10.1126/science.1225542

Dahl TW, Stevenson DJ. 2010. Turbulent mixing of metal and silicate during planet accretion—and interpretation of the Hf-W chronometer. *Earth Planet. Sci. Lett.* 295:177–86. doi: 10.1016/j.epsl.2010.03.038

Darwin GH. 1879. A tidal theory of the evolution of satellites. *Observatory* 3:79–84

Dauphas N, Pourmand A. 2011. Hf-W-Th evidence for rapid growth of Mars and its status as a planetary embryo. *Nature* 473:489–92. doi: 10.1038/nature10077

Desch SJ, Taylor GJ. 2013. Isotopic mixing due to interaction between the protolunar disk and the Earth's atmosphere. *Lunar Planet. Sci. Conf. Abstr.* 44:2566

Dobrovolskis AR, Peale SJ, Harris AW. 1997. Dynamics of the Pluto-Charon binary. In *Pluto and Charon*, ed. SA Stern, DJ Tholen, pp. 159–90. Tucson: Univ. Ariz. Press

Drake MJ. 1986. Is lunar bulk material similar to Earth's mantle? See Hartmann et al. 1986, pp. 105–24

Elkins-Tanton LT. 2012. Magma oceans in the inner Solar System. *Annu. Rev. Earth Planet. Sci.* 40:113–39. doi: 10.1146/annurev-earth-042711-105503

Elkins-Tanton LT, Burgess S, Yin Q-Z. 2011. The lunar magma ocean: reconciling the solidification process with lunar petrology and geochronology. *Earth Planet. Sci. Lett.* 304:326–36. doi: 10.1016/j.epsl.2011.02.004

England PC, Molnar P, Richter FM. 2007. Kelvin, Perry and the age of the Earth. *Am. Sci.* 95:342–49. doi: 10.1511/2007.66.3755

Garrick-Bethell I, Weiss BP, Shuster DL, Buz J. 2009. Early lunar magnetism. *Science* 323:356–59. doi: 10.1126/science.1166804

Genda H, Abe Y. 2005. Enhanced atmospheric loss on protoplanets at the giant impact phase in the presence of oceans. *Nature* 433:842–44. doi: 10.1038/nature03360

Gerstenkorn H. 1955. Über Gezeitenreibung beim Zweikörperproblem. *Z. Astrophys.* 36:245

Golabek GJ, Gerya TV, Morishima R, Tackley PJ, Labrosse S. 2011. Towards combined modelling of planetary accretion and differentiation. *EPSC Abstr.* 6:1140

Goldreich P. 1966. History of the lunar orbit. *Rev. Geophys. Space Phys.* 4:411–39. doi: 10.1029/RG004i004p00411

Goldreich P, Soter S. 1966. *Q* in the Solar System. *Icarus* 5:375–89. doi: 10.1016/0019-1035(66)90051-0

Halliday AN. 2008. A young Moon-forming giant impact at 70–110 million years accompanied by late-stage mixing, core formation and degassing of the Earth. *Philos. Trans. R. Soc. A* 366:4163–81. doi: 10.1098/rsta.2008.0209

Harrison TM. 2009. The Hadean crust: evidence from >4 Ga zircons. *Annu. Rev. Earth Planet. Sci.* 37:479–505. doi: 10.1146/annurev.earth.031208.100151

Hartmann WK, Davis DR. 1975. Satellite-sized planetesimals and lunar origin. *Icarus* 24:504–14. doi: 10.1016/0019-1035(75)90070-6

Hartmann WK, Phillips RJ, Taylor GJ, eds. 1986. *Origin of the Moon*. Houston, TX: LPI

Hauri EH, Weinreich T, Saal AE, Rutherford MC, Van Orman JA. 2011. High pre-eruptive water contents preserved in lunar melt inclusions. *Science* 333:213–15. doi: 10.1126/science.1204626

Hood LL, Mitchell DL, Lin RP, Acuna MH, Binder AB. 1999. Initial measurements of the lunar induced magnetic dipole moment using Lunar Prospector Magnetometer data. *Geophys. Res. Lett.* 26:2327–30. doi: 10.1029/1999GL900487

Housen KR, Schmidt RM, Holsapple KA. 1983. Crater ejecta scaling laws: fundamental forms based on dimensional analysis. *J. Geophys. Res.* 88(B3):2485–99

Ida S, Canup RM, Stewart GR. 1997. Lunar accretion from an impact-generated disk. *Nature* 389:353–57. doi: 10.1038/38669

Jackson AP, Wyatt MC. 2012. Debris from terrestrial planet formation: the Moon-forming collision. *MNRAS* 425:657–79. doi: 10.1111/j.1365-2966.2012.21546.x

Jacobsen SB. 2005. The Hf-W isotopic system and the origin of the Earth and Moon. *Annu. Rev. Earth Planet. Sci.* 33:531–70. doi: 10.1146/annurev.earth.33.092203.122614

Jacobson SA, Morbidelli A, Raymond SN, O'Brien DP, Walsh KJ, Rubie DC. 2014. Highly siderophile elements in Earth's mantle as a clock for the Moon-forming impact. *Nature* 508:84–87

Jeans JH. 1919. *Problems of Cosmogony and Stellar Dynamics*. Cambridge, UK: Cambridge Univ. Press

Jeffreys H. 1924. *The Earth: Its Origin, History and Physical Constitution*. Cambridge, UK: Cambridge Univ. Press. 1st ed.

Jeffreys H. 1929. The planetesimal hypothesis. *Observatory* 52:173–78

Johnson BC, Lisse CM, Chen CH, Melosh HJ, Wyatt MC, et al. 2012. A self-consistent model of the circumstellar debris created by a giant hypervelocity impact in the HD 172555 system. *Astrophys. J.* 761:45. doi: 10.1088/0004-637X/761/1/45

Jones JH, Palme H. 2000. Geochemical constraints on the origin of the Earth and Moon. See Canup & Righter 2000, pp. 197–216

Joy KH, Arai T. 2013. Lunar meteorites: new insights into the geological history of the Moon. *Astron. Geophys.* 54:4.28–32. doi: 10.1093/astrogeo/att121

Jutzi M, Asphaug E. 2011. Forming the lunar farside highlands by accretion of a companion moon. *Nature* 476:69–72. doi: 10.1038/nature10289

Kaula WM. 1971. Dynamical aspects of lunar origin. *Rev. Geophys. Space Phys.* 9:217–38. doi: 10.1029/RG009i002p00217

Kipp ME, Melosh HJ. 1986. Short note: a preliminary numerical study of colliding planets. See Hartmann et al. 1986, pp. 643–47

Leinhardt ZM, Marcus RA, Stewart ST. 2010. The formation of the collisional family around the dwarf planet Haumea. *Astrophys. J.* 714:1789–99. doi: 10.1088/0004-637X/714/2/1789

Matsui T, Abe Y. 1986. Origin of the Moon and its early thermal evolution. See Hartmann et al. 1986, pp. 453–68

McKeegan KD, Kallio APA, Heber VS, Jarzebinski G, Mao PH, et al. 2011. The oxygen isotopic composition of the Sun inferred from captured solar wind. *Science* 332:1528–32. doi: 10.1126/science.1204636

Melosh HJ. 1990. Giant impacts and the thermal state of the early Earth. In *Origin of the Earth*, ed. HE Newsom, JH Jones, pp. 69–83. New York: Oxford Univ. Press

Melosh HJ. 2009. An isotopic crisis for the giant impact origin of the Moon? *Meteorit. Planet. Sci. Suppl.* 72:5104

Melosh HJ, Vickery AM. 1989. Impact erosion of the primordial atmosphere of Mars. *Nature* 338:487–89. doi: 10.1038/338487a0

Morishima R, Watanabe S. 2001. Two types of co-accretion scenarios for the origin of the Moon. *Earth Planets Space* 53:213–31

Moulton FR. 1905. On the evolution of the Solar System. *Astrophys. J.* 22:165. doi: 10.1086/141260

Mukhopadhyay S. 2012. Early differentiation and volatile accretion recorded in deep-mantle neon and xenon. *Nature* 486:101–4. doi: 10.1038/nature11141

Nimmo F, Agnor CB. 2006. Isotopic outcomes of N-body accretion simulations: constraints on equilibration processes during large impacts from Hf/W observations. *Earth Planet. Sci. Lett.* 243:26–43. doi: 10.1016/j.epsl.2005.12.009

O'Brien DP, Morbidelli A, Levison HF. 2006. Terrestrial planet formation with strong dynamical friction. *Icarus* 184:39–58. doi: 10.1016/j.icarus.2006.04.005

Ohtake M, Matsunaga T, Haruyama J, Yokota Y, Morota T, et al. 2009. The global distribution of pure anorthosite on the Moon. *Nature* 461:236–40. doi: 10.1038/nature08317

O'Keefe JA, Sullivan EC. 1978. Fission origin of the Moon: cause and timing. *Icarus* 35:272–83. doi: 10.1016/0019-1035(78)90012-X

Pahlevan K, Stevenson DJ. 2007. Equilibration in the aftermath of the lunar-forming giant impact. *Earth Planet. Sci. Lett.* 262:438–49. doi: 10.1016/j.epsl.2007.07.055

Peale SJ, Cassen P. 1978. Contribution of tidal dissipation to lunar thermal history. *Icarus* 36:245–69. doi: 10.1016/0019-1035(78)90109-4

Pierazzo E, Artemieva N, Asphaug E, Baldwin EC, Cazamias J, et al. 2008. Validation of numerical codes for impact and explosion cratering: impacts on strengthless and metal targets. *Meteorit. Planet. Sci.* 43:1917–38. doi: 10.1111/j.1945-5100.2008.tb00653.x

Rasio FA, Shapiro SL. 1994. Hydrodynamics of binary coalescence. I. Polytropes with stiff equations of state. *Astrophys. J.* 432:27. doi: 10.1086/174566

Raymond SN, Quinn T, Lunine JI. 2006. High-resolution simulations of the final assembly of Earth-like planets. I. Terrestrial accretion and dynamics. *Icarus* 183:265–82. doi: 10.1016/j.icarus.2006.03.011

Raymond SN, Schlichting HE, Hersant F, Selsis F. 2013. Dynamical and collisional constraints on a stochastic late veneer on the terrestrial planets. *Icarus* 226:671–81

Reufer A. 2011. *Collisions in planetary systems*. PhD Thesis, University of Bern

Reufer A, Meier MMM, Benz W, Wieler R. 2012. A hit-and-run giant impact scenario. *Icarus* 221:296–99. doi: 10.1016/j.icarus.2012.07.021

Ringwood AE. 1986. Terrestrial origin of the Moon. *Nature* 322:323–28. doi: 10.1038/322323a0

Rubie DC, Nimmo F, Melosh HJ. 2007. Formation of Earth's core. In *Evolution of the Earth*, ed. J Stevenson, pp. 51–90. Treatise Geophys. 9. Amsterdam: Elsevier

Rudge JF, Kleine T, Bourdon B. 2010. Broad bounds on Earth's accretion and core formation constrained by geochemical models. *Nat. Geosci.* 3:439–43. doi: 10.1038/ngeo872

Safronov VS. 1972. *Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets*, transl. Isr. Program Sci. Transl. Jerusalem: Keter (from Russian)

Salmon J, Canup RM. 2012. Lunar accretion from a Roche-interior fluid disk. *Astrophys. J.* 760:83. doi: 10.1088/0004-637X/760/1/83

Schlichting HE, Warren PH, Yin Q-Z. 2012. The last stages of terrestrial planet formation: dynamical friction and the late veneer. *Astrophys. J.* 752:8. doi: 10.1088/0004-637X/752/1/8

Scott ERD, Haack H, Love SG. 2001. Formation of mesosiderites by fragmentation and reaccretion of a large differentiated asteroid. *Meteorit. Planet. Sci.* 36:869–91

Sharp ZD, Draper DS. 2013. The chlorine abundance of Earth: implications for a habitable planet. *Earth Planet. Sci. Lett.* 369:71–77. doi: 10.1016/j.epsl.2013.03.005

Shearer CK, Hess PC, Wieczorek MA, Pritchard ME, Parmentier EM, et al. 2006. Thermal and magmatic evolution of the Moon. *Rev. Mineral. Geochem.* 60:365–518. doi: 10.2138/rmg.2006.60.4

Stevenson DJ. 1987. Origin of the Moon—the collision hypothesis. *Annu. Rev. Earth Planet. Sci.* 15:271–315. doi: 10.1146/annurev.ea.15.050187.001415

Stewart ST, Leinhardt ZM. 2012. Collisions between gravity-dominated bodies. II. The diversity of impact outcomes during the end stage of planet formation. *Astrophys. J.* 751:32. doi: 10.1088/0004-637X/751/1/32

Terez EI, Gerasimov ME. 2009. Did the Moon form as a result of a thermonuclear explosion? *Bull. Crime. Astrophys. Obs.* 105:129–34. doi: 10.3103/S0190271709010148

Thompson C, Stevenson DJ. 1988. Gravitational instability in two-phase disks and the origin of the Moon. *Astrophys. J.* 333:452–81. doi: 10.1086/166760

Thomson W. 1864. On the secular cooling of the Earth. *Trans. R. Soc. Edinb.* 23:167–69

Touboul M, Kleine T, Bourdon B, Palme H, Wieler R. 2007. Late formation and prolonged differentiation of the Moon inferred from W isotopes in lunar metals. *Nature* 450:1206–9. doi: 10.1038/nature06428

Touma J. 2000. The phase space adventure of the Earth and Moon. See Canup & Righter 2000, pp. 165–78

Touma J, Wisdom J. 1994. Evolution of the Earth-Moon system. *Astron. J.* 108:1943–61. doi: 10.1086/117209

Touma J, Wisdom J. 1998. Resonances in the early evolution of the Earth-Moon system. *Astron. J.* 115:1653–63. doi: 10.1086/300312

Urey HC, ed. 1952. *The Planets: Their Origin and Development*. New Haven, CT: Yale Univ. Press

Wada K, Kokubo E, Makino J. 2006. High-resolution simulations of a Moon-forming impact and postimpact evolution. *Astrophys. J.* 638:1180–86. doi: 10.1086/499032

Ward WR, Canup RM. 2000. Origin of the Moon's orbital inclination from resonant disk interactions. *Nature* 403:741–43. doi: 10.1038/35001516

Warren PH. 1985. The magma ocean concept and lunar evolution. *Annu. Rev. Earth Planet. Sci.* 13:201–40. doi: 10.1146/annurev.ea.13.050185.001221

Weber RC, Lin P-Y, Garnero EJ, Williams Q, Lognonné P. 2011. Seismic detection of the lunar core. *Science* 331:309–12. doi: 10.1126/science.1199375

Wetherill GW. 1976. The role of large bodies in the formation of the Earth and Moon. *Lunar Planet. Sci. Conf. Abstr.* 7:3245–57

Wetherill GW. 1980. Formation of the terrestrial planets. *Annu. Rev. Astron. Astrophys.* 18:77–113. doi: 10.1146/annurev.aa.18.090180.000453

Wetherill GW. 1985. Occurrence of giant impacts during the growth of the terrestrial planets. *Science* 228:877–79. doi: 10.1126/science.228.4701.877

Wetherill GW, Stewart GR. 1993. Formation of planetary embryos: effects of fragmentation, low relative velocity, and independent variation of eccentricity and inclination. *Icarus* 106:190–209. doi: 10.1006/icar.1993.1166

Wiechert U, Halliday AN, Lee D-C, Snyder GA, Taylor LA, Rumble D. 2001. Oxygen isotopes and the Moon-forming giant impact. *Science* 294:345–48. doi: 10.1126/science.1063037

Wilhelms DE. 1993. *To a Rocky Moon: A Geologist's History of Lunar Exploration*. Tucson: Univ. Ariz. Press

Yu G, Jacobsen SB. 2011. Fast accretion of the Earth with a late Moon-forming giant impact. *Proc. Natl. Acad. Sci. USA* 108:17604–9. doi: 10.1073/pnas.1108544108

Zahnle K, Arndt N, Cockell C, Halliday A, Nisbet E, et al. 2007. Emergence of a habitable planet. *Space Sci. Rev.* 129:35–78. doi: 10.1007/s11214-007-9225-z

Zhang J, Dauphas N, Davis AM, Leya I, Fedkin A. 2012. The proto-Earth as a significant source of lunar material. *Nat. Geosci.* 5:251–55. doi: 10.1038/ngeo1429