The elephant in the room: Data from new missions by both traditional and new space faring nations and new measurements on Apollo-era samples and lunar meteorites are revolutionizing our knowledge of the Moon. Although many of the 'classic' views of the surface, interior, and atmosphere of the Moon, and their evolution through time have changed in light of these new data, the prevailing model of the formation of the Moon through a giant impact [1,2] continues to be virtually universally adhered to.

This is surprising, because a wide range of recent studies shows that our best estimate of lunar bulk chemistry is completely inconsistent with dynamical models of giant impacts that reproduce the current physical properties and dynamics of the Earth-Moon system. Such models predict the chemical composition of the Moon to differ significantly from that of the Earth. From a chemical point of view, an alternative Moon formation hypothesis that is much easier to defend would result on a Moon that is simply composed of terrestrial silicate material with a composition equivalent to the Bulk Silicate Earth (BSE). Here, we outline one such model. Our main aim is not to convince you of the validity of our alternative hypothesis (although admittedly that would be nice); our goal is to convince you that (a) the giant impact model is facing serious problems in light of a growing body of increasingly sophisticated chemical analyses and dynamical simulations and (b) alternative models should be developed and tested.

Recent chemical analyses and dynamical modelling: Measurements of the isotopic composition of lunar rocks [3-6], of their Sm-Nd [7] and Hf-W [8,9] systematics, and the water content of the lunar mantle [10,11], demonstrate that the bulk silicate Earth and the Moon show an surprisingly high degree of similarity. This is fully inconsistent with a giant impact model for the formation of the Moon. High-resolution smooth-particle hydrodynamic (SPH) simulations have shown that a grazing collision between the proto-Earth and a Mars-sized planet (Theia) can produce Earth-Moon systems with the correct masses, orbits, and angular momentum. Current simulations indicate that 70-80% of the mass of the Moon originates from Theia, not from Earth [12]. This discrepancy between geochemical data and dynamical models can not be resolved by proposing that the proto-Earth and the impactor formed at a similar distance from the Sun, for example in one of Earth’s Lagrange points [13]. The required size difference between Earth and impactor would lead to major differences in pressure-temperature-time conditions for core formation in these bodies (which predate Moon formation in all giant impact models). The resulting differences in Si isotope and Hf-W systematics would be resolvable with current analytical techniques, and are not observed.

Resolution of this discrepancy without changing one or more of the the main premises of the giant impact model requires complete isotopic homogenisation of Earth and Theia material after the impact. Turbulent exchange between partially molten and vaporised Earth and Moon shortly after the impact has been invoked to explain the similarity in O isotopes [14]. The effectiveness and dynamics of this mechanism are contested [15,16]. Even if this process could explain the O isotopes similarity, it is highly unlikely that such a mechanism can also fully homogenise initial differences in the isotopic compositions of much heavier, refractory elements including Si, Cr, Ti, Nd, Hf and W. For example, [17] suggest that this scenario would lead to differences in Si isotope compositions for Earth and Moon, which are not observed [6]. Given the large and growing uncertainty surrounding the feasibility of current versions of the giant impact model in light of geochemical data, alternative hypotheses for lunar formation need to be explored. In our view, a reasonable starting point for alternative hypotheses should be the notion that the Moon is so similar to the BSE because it was formed from the BSE. Below we outline a scenario that is consistent with this explanation.

Forming the Moon from the BSE: We revisit the hypothesis that the Moon was formed directly from terrestrial silicate material, as first proposed in the 'fission’ hypothesis by Darwin [18]. According to [18], the Moon originated from a hot, fast-spinning Earth. In Darwin’s model, centrifugal forces marginally exceeded equatorial attraction, and the Moon was formed from resonant effects of solar tides. In the 1960s, Ringwood and Wise updated Darwin’s hypothesis by including models for the thermal evolution and internal differentiation history of the Earth. They suggested that core-mantle differentiation led to a reduced moment of inertia of the Earth and hence to a larger angular velocity. The starting point for these modified models is a proto-Earth that is rotating rapidly (rotation period of ~ 2.7 h) with gravitational forces at the Earth’s surface only barely exceeding centrifugal
forces. In this situation, a slight increase in angular velocity would allow part of Earth’s equatorial mass to be launched into space. The main problem with this Darwin-Ringwood-Wise (DRW) model is the fact that the angular momentum of the current Earth-Moon system is only ~ 27% of that required for a 2.7 h rotation rate of the proto-Earth. In the absence of viable models for a decrease in angular momentum by a factor of four during Moon formation and subsequent Earth-Moon system evolution, the ‘fission’ hypothesis was abandoned before the first lunar sample was collected in the Apollo 11 mission.

An alternative hypothesis: We re-examined the dynamics of the Earth-Moon system and the energetics of initial Earth-Moon separation. In contrast to previous ‘fission’ models, our starting assumption is that the angular momentum of the proto-Earth before Moon formation must have been close to (within 10 per cent of) that of the present-day Earth-Moon system. This same assumption is made in recent three-dimensional hydrodynamic simulations of a giant impact origin for the Moon [12]. We then calculated the amount of energy that would be required to separate a Moon from the Earth in this case, and tried to identify processes other than giant impacts that could deliver this amount of energy in a very short time period. Nuclear fission is the only known natural process that we could identify, and we explored if and how a nuclear explosion in the interior of the Earth could lead to the ejection of silicate material to form the Moon. The iron-poor nature of the Moon requires that a fission-induced explosion leading to the launch of the Moon would have to have occurred outside Earth’s metallic core, with the core-mantle boundary (CMB) region as the most likely source region [19]. Depending on details of the triggering of a nuclear explosion the amount of energy required to launch the Moon is on the order of $0.5\times10^{50}$ J. This corresponds to a small fraction of the fissile content of the CMB at the time [20]. The dynamic effects of such an explosion have been explored by hydrodynamic modelling [21,22]. The Figure shows that the rapidly expanding plasma resulting from a deep nuclear explosion disrupts and expels overlying mantle, crust and gaseous materials, eventually resulting in the formation of a Moon-sized silicate body in an orbit around Earth. This scenario by definition results in an identical isotopic composition of Earth and Moon for both lighter and heavier elements as well as water, without relying on full post-collisional equilibration. Note that the decay products of the explosion do not end up in the Moon in these dynamical simulations. Our model is also consistent with the angular momentum and energy of the present-day Earth-Moon system. The recently proposed presence of two moons early in the history of the Earth-Moon system [23] does not require a giant impact either: a wide range of silicate fragments is produced in our proposed scenario (see Figure).

Conclusions: Moon formation models have to be consistent with lunar chemistry. Current versions of the giant impact model are not. Alternative models in which the Moon is formed from terrestrial material deserve more detailed study. We provide one (radical) alternative hypothesis – we hope others will follow.

Figure: Hydrodynamic model of the aftermath of a nuclear explosion at the boundary between metallic core (black) and silicate mantle (dark grey) of an Earth-sized planet. Top panel: $t\sim2700s$, middle panel: $t\sim4000s$, bottom panel: $t\sim13000s$ after explosion. Light-grey material is plasma / gaseous. Numbers denote sizeable silicate fragments. Fragment 8 reaches Earth orbit and has Moon-like mass. From [22].