The moon illusion – short history of a long-standing mystery of science

When the moon rises over the horizon on a beautiful summer evening, it looks larger than usual. Perhaps you have observed that in the past couple of weeks, this interesting summer phenomenon has been particularly prominent. Although some aspects of the phenomenon are well understood, others are not and an encompassing explanation is, amazingly, still missing.

The summer, the Sun, and the Moon

This year, the summer solstice (June 21) nearly coincided with the full moon (June 22) – ideal for observing the moon illusion. Since the full moon and the sun are opposite, and since in summer the sun is high, the full moon in summer is low at the horizon. When the full moon is near the horizon, we perceive it larger than when it stands near the zenith. Since June 1987, the full moon hasn't been as low in the sky as we see it in these weeks; consequently, the moon illusion is currently stronger than it has been for eighteen years.

In search of an illusion

When the moon illusion occurs, the moon looks some 50% larger than usual. It is a phenomenon that is undoubtedly occurring, yet cameras, unlike the human eye, cannot see it. It is at the same time real and unreal, fact and illusion. The circumstance calls for an explanation, but there is no entirely convincing theory that would explain it. The two most popular theories actually stand more for research hypotheses than for well-established findings. They are the "sky dome" (or distance illusion) theory and the "oculomotor micropsia" (or angular size illusion, a term to be explained in a moment) theory, but both offer only partial explanations.

So much we know: our perception of the moon varies with its position above the horizon, although its linear size (actual physical size as measured by its diameter and volume) and distance from the earth remain about the same. To be precise, the distance varies slightly: when we observe the moon at the
horizon, it is roughly one earth radius further away than at the zenith, which means its angular size (the angle of regard formed by the endpoints of its diameter) is a nuance smaller than when it is overhead – just the opposite of the moon illusion. According to Simanek (2002), the difference in angular size is 2%; but for our present purpose, it is accurate enough to assume that the moon's distance from a human observer, and thus its angular size, are constant.

Angular size, in distinction to linear size, includes information on distance; the angular size of an object decreases proportionally to its distance, that is, it is inversely proportional to distance. Since the "true" (undistorted) angular size of the moon on the surface of the earth is basically constant, there remain two basic options for explaining the moon illusion:

- Either we assume that its perceived angular size (as distinguished from its actual angular size) varies due to optical distortions, a variation that could be caused either by the refraction of the moon's light in the atmosphere or by the nature of the human eye (its optical processing capabilities), or by both. To the extent that this hypothesis holds, the moon illusion is an angular size illusion.

- Alternatively, we may assume that the moon's perceived angular size remains constant and it is the human mind (our cognitive apparatus for reading and interpreting visual signals) which produce the difference, by computing the distance information contained in the perceived angular size differently depending on the moon's position in the sky. To the extent that this hypothesis holds, the moon illusion is a distance illusion.

Regarding the angular size illusion approach, we can immediately rule out the "refraction" theory; for although refraction effects do effectively occur, they again work in the opposite direction: they make the horizon moon look slightly smaller (by about 1.7%, according to Simanek 2002) and in addition flatten its shape, thereby making its vertical diameter look even smaller. Hence, atmospheric refraction effects, far from explaining the moon illusion, actually imply that the effect that causes the moon illusion is even stronger than we perceive it. We thus need a refraction-independent explanation of that unknown effect (or combination of effects). I'll begin with the cognitive
apparatus (or distance illusion) approach, the so-called sky-dome theory, as it is the older and still more widely accepted explanation.

The "sky dome" theory This explanation assumes that the perceived angular size of the moon remains the same, namely, about half a degree. In agreement with this assumption, we know that high and low moons produce on the eyes's retina an image of identical size, about 0.15 mm wide. If this assumption is correct, the only way to explain the apparent variation of size is that the human mind, dependent on the moon's position above the horizon, judges its distance differently. We are dealing not with an illusion of angular size in the first place but rather with an illusion of distance. By analogy, if two balloons in the sky appear to have the same angular size but we judge one of them to be further away, the latter will "look" bigger to us. Our mind recalculates its size, as it were, so as to make up for its increased distance.

But why should our mind do this kind of differential distance computing for low and high moons, given that we know that the moon is always at the same distance? The "sky dome" theory explains this by assuming that our mental model of the sky (and thus, of the moon's orbit), shaped by both our everyday experience and our theoretical expectations, need not be identical with the "true" shape of the sky (the moon's actual orbit). Maybe our mental model of the sky is closer to the ancient view of the sky as a flattened dome rather than to current astronomical knowledge. In fact this traditional world view still conforms to a number of everyday experiences: bird watchers, for example, know that birds flying high are usually closer than those flying near the horizon. The same holds for airplanes or for clouds.

Our mental model of the sky may also be conditioned by our rather limited stereoscopic capabilities in judging distance cues. Basically, this conjecture makes sense because distance cues are not the same at the horizon and at the zenith. Furthermore, our stereoscopic capabilities have evolved for judging the distance of near-by objects rather than that of very remote objects in the sky. Both circumstances may have shaped our mental geometry of visual space in a way that treats objects above our heads differently from those on or near the ground.

A lot of sound reasons for taking the sky dome idea seriously! Note, however, that these reasons alone do not tell us what shape of the sky dome
we should properly assume to help us explain the moon illusion. I find it rather amazing that all the proponents of this approach of whom I am aware, have thus far assumed a flat dome model. It seems to me that determining the precise way in which our mental geometry of visual space differs from normal geometry is a (possibly complex) question of empirical science rather than of theoretical speculation only, and thus remains entirely open at present.

Now, if we do take the idea seriously, it has an important consequence: our cognitive apparatus then *somehow* needs to account for the ambiguity of visual signals about angular size with respect to of distance, perspective, and linear size. How can it achieve this? Since visual experience is not a reliable arbiter, it can only assess distance, perspective, and linear size by recalculating visual signals *according to its own inner model of visual space*. Because the process to a large extent occurs unconsciously, what matters is not so much our contemporary astronomical knowledge (although I would not preclude that theoretical expectations matter, in as far as they become intrinsic aspects of our world view) but rather our inner model of the sky dome, which may differ from one individual to another. If that model, whether we are aware of it or not, suggests a flat dome, the mind will assume that low moons are more distant and therefore will make them look larger. The distance illusion, then, is at heart a *sky dome illusion* (Figure 1).

![Figure 1: Flat sky dome theory of the moon illusion according to Kaufman & Kaufman (2000)](http://wulrich.com/picture_july2005.html)

Even though the (actual and perceived) angular size of low and high moons is the same, the distance illusion caused by the mind's flat dome model of the sky makes it compute different apparent sizes of the moon (black disks). This, then, is the core thesis of the sky dome theory in its popular flat dome version.
A modified sky dome theory Thus far, I find the approach rather illuminating. However, there is a basic difficulty of this explanation: it does not accurately describe the way most people experience the moon illusion. For me, at least, it is not true that I perceive low moons to be more distant than high moons; rather, the contrary is true. In fact, vision researchers have found that most people judge low moons to be just as close, if not closer than high moons (see, e.g., Boring 1962 and Gregory 1965). A few authors have noted this and have concluded that the sky dome theory is altogether mistaken (e.g. McCready 1965). This conclusion may be too hasty, though: the experimental finding in question merely implies that the flat sky dome model is empirically inaccurate, but not that the mental sky model approach as such is misguided in principle.

It seems too early, then, for throwing the constant angular size hypothesis definitely over board in favor of the alternative hypothesis, according to which the moon illusion depends on a changing perception of the moon's angular size. The empirical fact that the retinal size of the moon does not differ among low and high moons, along with the mentioned conjectures concerning our mental geometry of visual space, certainly suggest that the distance illusion hypothesis has some merits. Hence, before turning to the angular size illusion hypothesis, it may be worthwhile to consider whether a modified sky dome model might better describe our experience. To this end, I suggest to redraw Figure 1 as follows (Figure 2).

![Figure 2: Vertical sky dome theory of the moon illusion](http://wulrich.com/picture_july2005.html)

The modified model according to Figure 2 retains the basic idea that our mental sky model need not be identical with the true sky, and that the moon illusion may have something to do with this difference; but it replaces the
experimentally falsified flat dome with a more vertical dome model. I do not claim that this modified sky dome model provides a valid explanation of the moon illusion; it is certainly an insufficient, at best partial explanation. I merely suggest that under the assumption of constant angular size of the horizon moon as compared to the zenith moon, it corresponds better to the subjective experience of most people. And since I do not think the current stage of knowledge allows us to definitively drop the distance illusion hypothesis in favor of a one-sided angular size illusion hypothesis, it certainly makes sense to try to save the sky dome model from its overt inconsistency with the actual experience of a majority of observers.

Whether or not the constant angular size assumption is adequate, and what are its theoretical merits as compared to the angular size illusion approach, cannot be decided on the basis of the model itself but requires extensive empirical research. All that the modified model aims to suggest is that it is possible to reconcile the constant angular size hypothesis with the empirical finding that most people see the horizon moon closer than the zenith moon, a finding that is inconsistent with the conventional flat dome model usually associated with the constant angular size assumption.

As a welcome side-effect, the modification also responds to the fact that the ancient flat dome concept of the sky has longs since been replaced by a different view of the cosmos. Our notion of the nightly sky is no longer shaped by a flat horizon of expectations, to use Karl R. Popper's (1972, p. 345) pertinent phrase for describing the theory-impregnated character of all human experience. Nowadays, when we raise our eyes to the zenith, we imagine not a flat dome but rather, infinite space. (On the other hand, when we turn our eyes back to the ground, our visual horizon tends to be more limited than in earlier centuries, whether by the next group of ugly buildings or by polluted atmospheric conditions.) A vertical sky dome model certainly captures our contemporary mind set as well as the ancient flat dome concept.

Furthermore, I would suggest that the revised sky dome model may, but need not, assume that the perceived angular size of the moon indeed remains constant. This is, in fact, another unnecessary assumption; it suffices to assume that the constant angular size assumption has some part to play in putting together the unsolved puzzle. I see no reason why an adequately
calibrated sky dome model might not contribute to understanding the moon illusion even if it should turn out in the end that the constant angular size hypothesis cannot be upheld. Considering the role of our mental sky model – our horizon of expectations, as it were – may be just as meaningful in combination with the angular size variation hypothesis. Given the magnitude of the moon illusion, it seems not implausible that both effects, a distance illusion and an angular size illusion, are needed to explain the full extent of the moon illusion.

The "oculomotor micropsia" theory Let us, then, turn to this second of the earlier-mentioned basic hypotheses and assume that the perceived angular size of the moon varies with its position in the sky. To the extent that this assumption holds, we can and need not (at least not fully) explain the moon illusion by the role of an inaccurate mental sky model and by the illusion of varying distance that it causes; rather, we then primarily need to look at the visual processing capabilities of our eye.

Among the main proponents of this approach are McCready (1965, 1986, 2004), Enright (1989a, b), and Roscoe (1989). We need not know the details of their theory in order to understand the basic idea. It says that the perceived angular size of an object varies with the focal distance of our eyes. This is an experimental finding that you can approximately simulate in the following way. Watching your computer monitor, hold a small object such as a pen about 20 cm away from your eyes in front of the monitor. Now focus your eyes on the pen and watch what happens to the monitor – it looks smaller. It has changed its angular size!

There is thus an obvious connection between the focal point of our stereoscopic vision – the accommodation of the eye lenses and the simultaneous convergence of the eyes to a certain distance – and the perceived angular size of objects. The exact nature of the connection is complex and requires more research, but basically we can say that the closer our focal point is, the smaller look more distant objects. This is what vision researchers call oculomotor micropsia, an effect that probably results from a combination of both accommodative micropsia (reduction in angular size caused by increased accommodation) and convergence micropsia (reduction in angular size caused by increased convergence of the eyes). Micropsia is a
Greek noun meaning "small sight" or "small appearance" in the sense of a smaller-than-real illusion. The effect was first described by the English physicist Sir Charles Wheatstone (1872), the inventor of the stereoscope (a device that generates a three-dimensional view from two photographs of the same subject taken at slightly different angles). The reverse effect is called macroscopia; it occurs when our focal point is more distant than an object, and has us see the object bigger than we'd expect.

Enright (1989a, b) and Roscoe (1989) have demonstrated these effects experimentally. More recent research indicates that micropsia and macropsia may equally occur when eye lens muscles are paralyzed or absent, or when one eye is covered. Obviously, the two effects are controlled by neurological processes (the same ones that control the eye movements) rather than directly and only by the mechanics of the eye movements themselves; let us not misunderstand the name "oculomotor micropsia / macropsia" in this respect. The fact that angular size illusion can be produced experimentally by influencing the movements (accommodation and convergence) of the eyes need not mean this is the only way such illusions can arise. Lest we make unnecessary assumptions, let us consider micropsia and macropsia as primarily but not exclusively occurring at the level of perception, and keep in mind the role of the brain in processing all visual signals.

Applied to the moon illusion, the theory postulates that the zenith moon looks smaller, and consequently more distant. One possible explanation is the earlier-mentioned difference of available distance cues near the horizon and near the zenith. Near the horizon, we usually can rely on familiar distance cues such as houses, trees, mountains and so on, which help our eyes to focus on a point that is more remote than these objects. Near the zenith, however, no such cues are available to direct our stereoscopic vision. Experimental findings show that in the absence of distance cues, our eyes tend to focus on a default focal point – its resting focus – which is much closer than the moon's actual distance. Leibowitz et al. (1975) have found it to be some 2 m away only. Vision researchers speak of an empty field micropsia. Now take our previous little experiment and substitute the zenith moon for the computer monitor: since the default focus is much closer to your eyes than the moon's actual distance, you perceive the zenith moon as smaller than you would otherwise; and since your focal point is also closer
than the horizon where you watch the rising moon, the zenith moon looks smaller than the horizon moon, and vice-versa. The earlier-mentioned evolutionary difference between our capabilities of handling horizontal and vertical distances may also be part of the explanation. Finally, darkness may also induce micropsia, for obvious reasons: it renders the use of distance cues more difficult. Vision researchers speak of night micropsia.

**In conclusion** Despite many open questions, it appears safe to say that the following aspects play a major role in the moon illusion:

- **Shortcomings of our stereoscopic vision in dealing with large distances**: Our stereoscopic capabilities are not made for visual objects more than a few hundred meters away.

- **Shortcomings of our mental processing of visual signals coming from above the head**: Evolutionary conjectures suggest we do not handle such signals as well as signals coming from the ground.

- **Complex interdependence of perception and cognition**: The processing of visual signals moves at different, partly unconscious levels that interact in complex ways. In order to translate the visual signals received by the retina into geometric interpretations of the location and movement of objects in space, we depend on a (partly unconscious) inner mental model of visual space.

- **Shortcomings of our mental geometry of visual space**: Our mental model of visual space need not (and probably cannot) be the same as the "true" geometry of space (assumed versus "true" sky dome).

- **Importance of distance cues**: For all the previous phenomena, our handling of distance cues – or their absence – appears to play a crucial trigger role.

- **Inseparability of size and distance illusions**: All the mentioned difficulties of processing stereoscopic spatial information can be understood to cause either illusions of distance (sky dome theory) or illusions of size (oculomotor theory), or both. Since angular size and distance are interdependent, it seems plausible to assume that both kinds of illusions work together in causing the moon illusion.
• **Magnitude of the moon illusion as compared to the effects explained by any available explanation**: The last-mentioned conclusion is supported by empirical findings which suggest that angular size illusions or distance illusions *alone* do not produce effects nearly as large as the moon illusion. The micropsia/macropsia theory, for instance, accounts experimentally for angular size differences of less than 10% (Simanek 2002), while the perceived angular size difference of the horizon moon as compared to the zenith moon is around 50%.

It cannot surprise, then, that recent reviews of the moon illusion literature (Ross and Plug 2002, Simanek 2002) conclude that no single theory available today can really explain the mystery. For the time being, the moon keeps its secret.

**Some doubts and philosophical issues** As convincing as both basic approaches to explaining the moon illusion may appear at first, it soon becomes obvious that neither looks at the whole picture:

• *Theoretically* speaking, both theories look too simple to me. The sky dome theory appears too simple because it one-sidedly focuses on our cognitive apparatus, at the cost of rather neglecting the primary difficulties at the level of perception itself. The oculomotor theory, on the other hand, appears equally one-sided in seeking the explanation in the limitations of our stereoscopic vision, while rather neglecting the way our cognitive apparatus may have learned to handle some of the limitations.

• *Empirically* speaking, neither approach is currently able to account for more than a fraction of the full extent of the moon illusion.

Both theoretically and empirically, the complementary nature of the two explanatory approaches seems therefore rather obvious, even though the involved researchers tend to treat them as mutually incompatible. I certainly recognize the difficulties that researchers may face in combining them, but even so, I see no intrinsic reasons why neither side should be able to integrate what the other side has to contribute.

Rather, it's probably just not the way science works. Researchers are eager to
uncover the weaknesses rather than the merits of competing approaches. It seems to me that in the case of the moon illusion, the two competing approaches are involved in a sort of (unnecessary) paradigm war. Every war has victims; in this case it is the fundamental interdependence of perceptual and cognitive phenomena which somehow seems to have got out of focus. At least this is the impression conveyed to me by the two main sources that I have consulted, by authors who are major contemporary representatives of the two approaches: Kaufman & Kaufman (2000) for the sky dome paradigm, who seek the root cause of the moon illusion in a distance illusion, and McCready (2004) for the oculomotor paradigm, who seeks the root cause in an angular size illusion. I could not help but gain the impression that both articles are more concerned to argue why the other side has got it wrong than to support the competing theory with their own specific insights.

Apart from this general doubt concerning the isolationist tendencies of both sides, I have, of course, a number of more specific doubts. I find it difficult to understand, for instance, how the protagonists of the oculomotor microscopia/macroscopia theory can assume that the perceived angular size of the moon varies with the position of the moon in the sky, without carefully discussing the contrary experimental findings which suggest that the moon image on the retina has a constant size of about 0.15 mm in diameter. The theory thus in effect assumes what it proposes to explain, namely, deception.

Secondly, the theory depends on assuming persistent visual deception. I miss an explanation of this persistent nature of the deception, that is, of the apparent absence of learning. It makes sense to assume that children easily take micropsia or macropsia effects for granted – need I refer to Gulliver in Lilliput or to Alice in Wonderland? However, once we are adults, shouldn't we expect that over time, we will learn to focus more correctly, given that we know we are victims of a visual illusion? Why is there apparently no learning? Why does the theory tell us so little about this question? Is it perhaps because obviously, considering the role of learning means giving more weight to what happens at the cognitive level, the territory of the rival theory?

But of course, a similar doubt can be formulated against the sky dome theory. Once we are conscious of the difference between our inner mental
sky and the true outer sky, shouldn't we expect that we can eventually adapt our inner model so as to avoid obvious deception? Actually, when we pass from childhood to adulthood, some learning does seem to occur, as Leibowitz and Hartman (1959) report, but this apparently does not manage to match the full extent of the moon illusion.

In other words, the current state of research into the moon illusion leaves me with more questions than answers. For example, can persistently wrong eye convergence and accommodation (the hypothesis of a basic angular size illusion at the level of perception) really explain the full extent of the moon illusion? Hardly. Likewise, can a persistently inaccurate mental model of the sky (the hypothesis of constant angular size, with its implication of a distance illusion at the cognitive level) fully explain it? Hardly. Can the limitations of our stereoscopic capabilities in dealing with distant objects, which is a core issue of both approaches, really be explained by concentrating either on perceptive (optical) or on cognitive (mental) processes? Hardly. And so on. There is no way round it: the state of our knowledge about the moon illusion is hardly satisfactory.

Although this state of affairs certainly does not cause me sleepless nights, epistemologically speaking I find it thought provoking. If this is how science fails to explain a familiar everyday phenomenon in the sky, can we imagine how it must fail to do justice to more complex and less obvious issues? The unsolved mystery of the moon illusion thus at least helps us in avoiding another persistent illusion, that of the "secure path" (Kant 1781) and objective nature (Popper 1972) of science. The moon illusion as an epistemological warning signal, as it were! If we take it seriously, it invites philosophical reflection on our possibly distorted "horizon of expectations" (or should I say: mental model?) regarding science. I would like to conclude by hinting at just three possible reflections.

- **Socrates and science:** If science is unable thus far to elucidate and explain even an everyday empirical phenomenon such as the moon illusion, perhaps we should remind ourselves more than currently en vogue of the virtue of Socratic modesty in science? Could a more Socratic practice of science not have benefited moon illusion research, for example, by helping oculomotor and sky dome protagonists avoid
all-too one-sided claims in favor of more cooperation and multidimensional theorizing?

- **Constructivism, nothing new under the sun:** At least since Kant's "Copernican turn" away from naive realism to critical idealism, we know that we tend to see what we expect; that all knowledge is individually and socially constructed. Despite much fashionable talk about constructivism, critical realism, and other supposed epistemological insights of our time, Kant has given us the basic critical message long ago: "To avoid errors, one must search for their origin in illusion. Uncovering illusion is a much greater service to truth than any direct refutation of errors." (my transl. from Vorlesungen über die Logik, see Kant 1992). Should moon illusion research perhaps be taken much more seriously than it has been taken thus far, both scientifically and epistemologically?

- **Popper's horizon of expectations, turned critically:** Popper's insights into the epistemological importance of our theoretical horizon of expectations are rarely matched by similar lucidity about the importance of two other major factors that condition all our knowledge claims, namely, value judgments and boundary judgments. In my work on critical systems heuristics, I have attempted to provide a generic framework for boundary critique in contexts of applied science and professional practice, because boundary judgments underpin all our judgments of what are relevant facts and values. Could there perhaps be an equivalent critical heuristics for basic science?

And finally: what kind of a world would it be in which the failure of moon illusion research would be taken seriously?

**References**


Technical data
Digital photograph taken on 26 June 2004 around 9:30 p.m., shutter speed 1/60, aperture f/2.8, ISO 50, focal length 8 mm (equivalent to 36 mm with a conventional 35 mm camera). Original resolution 1158 x 759 pixels; current resolution 1255 x 764 pixels, compressed to 45 KB.

July, 2005
A case of micropsia? Summer moon rising into the evening sky

„The sky's the limit”
(An insight that certainly applies to our limited understanding of the moon illusion)

Write down your thoughts before you forget them!
Just be sure to copy them elsewhere before leaving this page.

Last updated 6 Dec 2009 (layout) and 3 July 2005 (text, first published 1 July 2005)