A letter to the Royal Society presenting

A new theory of light and colours

Isaac Newton

[Brackets] enclose editorial explanations. Small ·dots· enclose material that has been added, but can be read as though it were part of the original text. Occasional •bullets, and also indenting of passages that are not quotations, are meant as aids to grasping the structure of a sentence or a thought. Every four-point ellipsis . . . . indicates the omission of a brief passage that seems to present more difficulty than it is worth. Longer omissions are reported between brackets in normal-sized type.—The small-caps headings are added in this version.

First launched: May 2010

[The Royal Society received and published this letter in 1671, under this heading: A Letter of Mr Isaac Newton, Mathematics Professor in the University of Cambridge, containing his new theory about light and colours; in which he says •that light is not homogeneous or the same all through, but rather consists of rays of different forms, some of which are more deflectable than others; and •that colours are not states that light gets into derived from refractions of natural bodies (as is generally believed), but rather are basic properties that are different in different rays. Several observations and experiments are said to prove this theory.]

* * * * *

·A surprising observation and failed explanations ·

As I promised you, I am writing to tell you that early in 1666 (when I was working on the grinding of optical glasses of shapes other than spherical), I acquired a triangular glass prism with which I wanted to try the celebrated phenomena of colours. [This refers to the rainbow-like series of colours that appear on a surface when sunlight is shone onto it through a prism. Newton often called this phenomenon an 'image' or 'appearance'—Latin *spectrum*—but he was also coming to use 'spectrum' as an English word; just once in this letter, but the present version will use it oftener.] For this purpose I darkened my room, made a small hole in my window-blind to let in a convenient amount of sunlight, and placed my prism at this light-entry so that the light would be refracted by it onto the opposite wall. It was at first a very pleasing entertainment to see the vivid and intense colours produced by this procedure; but while I was attending more carefully to these colours, I became surprised to see that their shape was oblong; I had expected them to be circular, because that's the shape that the generally accepted laws of refraction say that they will have ·if they have entered the room through a circular hole·.

Their borders along the sides were straight lines, at the ends the light faded away so gradually that it was hard to say for sure what their shape was; but the ends seemed semicircular.

Comparing the length of this coloured spectrum [Newton's word] with its width, I found it to be about five times greater. This was such an enormous difference that it spurred me to an exceptionally probing attempt to find out what caused it. I could hardly think that the varying thickness of the glass prism, or the termination with shadow or darkness [referrng to Hooke's theory that colours are produced only towards the edges of a beam of light, nearest to the darkness], could affect light in such a way as produce such an effect; but I thought it would be as well to start by checking out those matters.

- •I sent light through parts of the prism of different thicknesses;
- •I sent it through holes in the blind of different sizes;
- •I put the prism outside the room, enabling the light to pass through it and be refracted before it reached the hole.

None of those details made any difference: the resulting spectrum was still the same through all these variations.

Then it occurred to me that this spreading out of the colours might be an effect of unevenness in the glass, or some other accidental •irregularity. To test this, I took another prism like the first one and placed it in such a way that the light, passing through both prisms, would be refracted in opposite ways, so that the second prism would restore the light to the course from which the first prism had diverted it. The upshot of this, I thought, would be that the regular effects of the first prism would be cancelled by the second prism, while the •irregular ones would be accentuated through being refracted and then refracted a second time. The upshot was that the light which the first prism spread out into an oblong shape was brought ·back· to a circular shape by the second prism, the circle being as regular as when the light ·came through that circular hole but· didn't pass through either prism. Thus, whatever it was that caused that length—that oblong shape—it wasn't any accidental irregularity ·in the materials·.

·Measurements and a more surprising observation·

I then proceeded to a more testing examination of what the results might be of differences in ·the angle· of the incidence of rays coming from various parts of the sun; and for that purpose I measured the various lines and angles relating to the image.

•distance from the hole or prism to the image: 22 feet;

•greatest length of the image: $13\frac{1}{4}$ inches;

•width of the image: $2\frac{5}{8}$ inches;

•diameter of the hole: $\frac{1}{4}$ of an inch;

•the angle formed by •the line of the rays reaching the middle of the spectrum and •the line that they would have followed if the prism weren't there: $44°56'$.

•vertical angle of the prism: $63°12'$.

•The angle between the light reaching the prism and that face of the prism was, as near as I could make it ·by gradually shifting the prism·, equal to the angle between the rays leaving the prism and *that* face of the prism. Consequently it was about 54°4′.

•The rays fell perpendicularly upon the wall.

•Subtracting the diameter of the hole from the length and breadth of the spectrum, there remains length = 13 inches, and breadth = $2\frac{3}{8}$ inches, comprehended by the rays that passed through the center of the said hole, and consequently the angle of the hole which that breadth subtended was about 31′, answerable to the sun's diameter [meaning that is about the angle that the sun subtends at any point on our planet], whereas the angle which its length subtended was more than five such diameters, namely 2°49′.

Having made these observations, I first computed from them the refractive power of that glass [i.e. the amount by which that glass tilts light coming into it and light going out of it]. . . ., and from that value I computed the refractions of two rays flowing from opposite edges of the sun's disc, given that there was a difference of 31′ in their angles of incidence [i.e. the angles at which they impinged on the prism], and I found that the emergent rays should have comprehended an angle of about 31′, as they did before they were incident.

This computation was based on the hypothesis of the proportionality of the sines of incidence and refraction [this refers to Snell's Laaw; its details needn't concern us now]; and my own experience of that made me think that it *couldn't* be so far off as to give the value 31′ to an angle which in reality was 2°49′. But my curiosity caused me again to take my prism, place it at my window, as before, and start turning it a little about its axis to and fro, so as to vary its angle to the light more than about 4° or 5°. I found that these rotations made no perceptible difference to the positions of the colours on the wall; which means that by varying the angle of incidence I made no perceptible difference to the quantity of refraction [i.e. the angle of refraction]. So I reached a conclusion that is strongly supported both by this experiment and by the previous calculation, namely that the difference in the incidence of rays flowing from different parts of the sun couldn't make them •diverge, after intersecting, at a perceptibly greater angle than the one at which they had converged; this angle was at most about 31′ or 32′; so some other cause had to be found that could make it 2°49′.

·CURVES?·

Then I began to suspect that perhaps the rays after passing through the prism move in curved lines, and move to different parts of the wall depending on *how* curved their trajectory is. And this suspicion was strengthened when I remembered that I had often seen a tennis ball move along a curved line after being struck with an oblique racquet. ·We know what causes that·. The stroke of the racquet gives the ball a forward motion and a rotating motion; on the side where the two motions are working together, the parts of the ball must press and beat the adjoining air more violently than the parts on the other side; that arouses correspondingly more resistance and reaction from the air on that side than on the other; ·and so the ball's path after being struck is curved·. By the same reasoning, if the rays of light are globular bodies which acquire a spin when they pass obliquely out of one medium (·the prism·) into another (·the air, or rather the ether·), they should encounter greater resistance from the adjoining ether on the side where the motions work together, and so be continually curved towards the other side. But plausible as this basis for suspicion was, when I came to test it I could see no such curvedness in the rays. And anyway I observed that the difference between the length of the image, and diameter of the hole through which the light was transmitted was proportional to the distance between them; and that was enough for my purpose, ·because it showed that the light was moving along straight lines·.

·THE CRUCIAL EXPERIMENT·

The successive failure of those suspicions eventually led me to the *experimentum crucis* [Latin = 'crucial experiment, the experiment that settles the matter'], which was this: [To the initial set-up of hole/prism/wall he added two boards, each with a hole in it, and a second prism. These were placed so that light passed

through the initial hole, then
immediately through the prism X, then some of it
immediately through the hole in board A, then all of that
across 12 feet to board B, then some of that
through the hole in B, and all of that
through prism Y, and all of that
across the remaining 10 feet to the wall.

For safety's sake, here is the set-up in his own words:]
I took two boards, and placed one of them close behind the prism at the window, so that the light might pass through a small hole made in it for the purpose, and fall on the other board, which I placed at about 12 feet distance, having first made a small hole in it also, for some of that incident light to pass through. Then I placed another prism behind this second board, so that the light, trajected through both the boards, might pass through that also, and be again refracted before it arrived at the wall.

I then took prism X in my hand and slowly turned it to and fro around its axis, enough to make changes in which parts of the image on board B passed through the hole in B, so that I could observe what parts of the wall prism Y refracted them onto. And I saw from the variation of those places that •light going to one end of the spectrum from the first prism was refracted much more by the second prism than was •the light going to the other end of the first prism's spectrum. So the true cause of the length of that ·initial oblong· image was found to be just this: **light consists of rays that are differently refractable**, so they were transmitted towards different parts of the wall according to their different degrees of refractability—quite independently of the angle at which they impinged on either prism.

·A DIGRESSION ON TELESCOPES AND MICROSCOPES·

When I understood this, I gave up the work on ·optical· glass that I mentioned at the start of this letter; because I saw that the perfection of telescopes had been limited not so much for lack of glasses accurately conforming to the prescriptions of the experts on optics (which is what everyone thought) as because the light itself is a heterogeneous mixture of differently refractable rays. So that if you did have a glass that was perfectly shaped to collect some one sort of rays into one point, it couldn't collect into the same point other rays which, though having the same angle of incidence upon the same medium, are refracted differently. Indeed, given how large I had found the differences of refractability to be, I was surprised that we have telescopes as good as the ones we do have.

[Newton proceeds with a fairly long 'digression' (his word) about telescopes and microscopes. His puzzle about telescopes being as good as they then were led him into various kinds of theorising and experimentation. A central feature of all this was the discovery that if the instruments are made with enough precision, you get a better telescope if you focus the light by reflecting it from a concave mirror than if you focus it through a lens. A notable personal remark in the course of this: 'Amidst these thoughts I was forced to leave

Cambridge by the plague, and it was more than two years before I proceeded further.' Details of 'these thoughts' are not given here: they are hard to follow, and have no bearing on the main point of this letter of Newton's.]

·Back to the properties of light·

But to return from this digression: I told you that light is not homogeneous, but consists of intrinsically unalike rays of which some are more refractable than others. Thus, of a number of rays that encounter the same medium at the same angle, some will be more refracted than others—not because of any property of the glass, or any other external cause, but because each individual ray has its own predisposition to be refracted at its own particular angle.

I shall now tell you about another, more notable, unalikeness in light-rays, which is the source of differences of colour. I shall set out the doctrine first, and then describe one or two of the ·supporting· experiments (there are others).

·The theory·

You will find the doctrine comprehended and illustrated in the following propositions.

(1) Just as the rays of light differ in degrees of refractability, so they also differ in what colours they are disposed to exhibit. Colours are *not* what they are generally believed to be, namely states that light gets into because of how it has been refracted or reflected by natural bodies. Rather, colours are basic properties of light, properties that come into existence when light does, and these properties are different in different rays. Some rays are disposed to exhibit a red colour and no other; some a yellow colour and no other, some a green colour and no other, and so on through the rest. And it is not only the more eminent colours that have their own particular rays; the same holds for all the intermediate shades.

(2) A given colour always has the same degree of refractability, and a given degree of refractability always goes with the same colour. The least refractable rays are all disposed to exhibit a red colour, and conversely the rays that are disposed to exhibit a red colour are all the least refractable. And the most refractable rays are all disposed to exhibit a deep violet colour, and conversely those that are apt to exhibit such a violet colour are all the most refractable. Similarly to all the intermediate colours in a continuous series belong intermediate degrees of refractability. This analogy between colours and refractability is very precise and strict; any two rays either exactly agree in both or are different to the same extent in both.

(3) The kind of colour and degree of refractability that any particular sort of ray has can't be changed by refraction, or by reflection from natural bodies, or by any other cause that I have so far found. When rays of any one sort have been well separated from rays of other kinds, they then obstinately retained their colour despite my best efforts to change them. I have

- refracted them with prisms, and reflected them with bodies which in daylight were of other colours;
- made them pass through a coloured film of air trapped between two compressed plates of glass;
- transmitted them through coloured mediums, and through mediums irradiated with other sorts of rays, and
- terminated them differently;

and yet I could never produce any new colour out of it. [Regarding 'terminated them differently': This refers to different ways of defining the edges of a beam of light, something that Newton cared about because of Hooke's theory that colours depend on what happens near those edges.] By contracting or dilating it I could make the colour more brisk or more faint, and by reducing the number of rays I could sometimes make it very obscure and dark; but I could never see any change in the shade.

(4) But apparent changes of colours may be made by mixing different sorts of rays. In such mixtures, the component colours appear only through their mixing with each other, resulting in an intermediate colour. And therefore if, by refraction or any other of the causes I have listed, the intrinsically unalike rays in such a mixture are separated, •colours will emerge that are different from the colour of the composition. Those •colours are not newly generated, but only made apparent by being separated; for if those rays are again entirely mixed and blended together, they will again compose the colour that they did before separation. And for the same reason, changes made by bringing different colours together are not real; for when the unalike rays are again separated, they exhibit the very same colours as they did before entering into the composition. You can see that blue and yellow powders, when finely mixed, appear green to the naked eye, yet the colours of the component corpuscles are not really changed but only blended. When you look at them with a good microscope you will see them as blue and yellow scattered amongst one another.

(5) So there are two sorts of colours: •original and simple colours and •colours made by compounding these. The original or primary colours are red, yellow, green, blue, and a violet-purple, together with orange, indigo, and an indefinite variety of intermediate shades.

(6) Colours that are exactly the same as these primary ones may be also produced by composition: a mixture of yellow and blue makes green; a mixture of red and yellow makes orange; a mixture of orange and yellowish green makes yellow. And in general if you mix any two colors that are not too far apart in the spectrum, then they jointly produce the colour that is intermediate between them in the spectrum. But that is not what happens when you mix two colours that are too far apart in the spectrum: orange and indigo don't produce the intermediate green, scarlet and green don't produce the intermediate yellow.

(7) But the most surprising, and wonderful composition was that of whiteness. There is no one sort of ray that can on its own exhibit whiteness; it is always compounded, and the compound has to include all the primary colours that I have listed, mixed in the right proportions. I have often seen with admiration that when all the colours of the spectrum are made to converge, so that they become mixed as they were in the light before it encountered the prism, they produce light that is entirely and perfectly white, with no perceptible difference between it and light directly from the sun. . . .

(8) This explains why light is usually white: light is a confused aggregate of rays with all sorts of colours, launched randomly from the various parts of luminous bodies. And that kind of confused aggregate (I repeat) generates whiteness, if there proportions of the ingredients are right; but if any one ·kind of ray· predominates, the light must incline to the colour ·that it always produces·. That is what is happening with the blue flame of brimstone, the yellow flame of a candle, and the various colours of the fixed stars.

(9) When all this is taken into account, it becomes obvious how colours are produced by the prism. The light reaching the prism is a mixture of rays each associated with its own special colour; because they differ in refractability, they leave the prism heading off in different directions, each taking its colour with it. The result is ·the spectrum·, an oblong of light-colours in an orderly series from the least refracted (scarlet) to the most refracted (violet). And this also explains why objects looked at through a prism appear coloured. The unalike rays, being refracted unequally, diverge ·from the prism· towards different parts of the retina, and there present images of coloured things, like the spectrum made by the sunlight on the wall. Through this inequality of refractions they become not only coloured but also very confused and indistinct.

(10) Why the colours of the rainbow appear in falling drops of rain is also explained in this way. [Newton's version of the explanation is condensed and hard to grasp. Here it is at greater length. Sunlight hits a raindrop, passes through it and is refracted out, continuing its journey to the earth. The angle at which a given ray is refracted is determined by what sort of ray it is—purple-carrying rays will exit the drop at one angle, red-carrying rays at another, and so on. At a given instant, what reaches your eye from a given drop will be just one sort of ray (say the purple-carrying one), so that drop will look purple to you: the other rays from it will be angled in such a way that they don't hit your eye. Of course a single drop won't be visible to you on its own; but you will get the purple effect from *all the drops that are at that same angle from your eye*; they will form a crescent or 'bow'; so you'll see the agglomeration of all those drops, i.e. you'll see the purple band in a rainbow. Repeat the story for (say) red: at a given instant the drops that refract the red-carrying rays to your eye will be a little higher in the sky than the ones exhibiting the purple 'bow'; so you'll see the red 'bow' above the purple one. Those in fact are the extremes. The other colours will appear in between them; and to each the same Newtonian explanation applies. Into his already condensed account of this Newton crams an explanation of double rainbows, in which a secondary rainbow appears above the primary one, with the order of colours reversed.]

(11) The odd phenomena in which an infusion of Lignum Nephriticum [a tropical hardwood with fluorescent properties], gold leaf, fragments of coloured glass, and some other transparently coloured bodies appears of one colour in one position and of another in another are no longer riddles. Those are substances apt to •reflect one sort of light and •transmit another. To see this, put them in a dark room and then illuminate them with homogeneous uncompounded light, ·i.e. with light all of whose rays are of one kind·. What you will see then is that they appear to have only the colour belonging to those rays, but they will be more vivid and luminous in one position than in another, according to whether they reflect more or transmit more of the incident colour.

(12) We can now clearly understand the facts of a surprising result that Mr Hooke reports having found (the report is somewhere in his *Micrographia*). He had two wedge-shaped transparent vessels, one filled with a red fluid and the other with a blue fluid: each of these on its own was transparent, but when they were put together they were jointly opaque. Explanation: one transmitted only red and the other only blue, so no rays could pass through both.

(13) I could add more examples like this, but I'll settle for ending with this general point: The colours of all natural bodies have no other origin than this, that they are variously qualified to reflect one sort of light in greater plenty than another. I have tested this in a dark room by illuminating those bodies with uncompounded—·i.e. homogeneous·—light of various colours. By that means, any body can be made to appear of any colour. They exhibit there no colour of their own, but always appear to have the colour of the light that is shone onto them; the only difference being that this displayed colour is at its most brisk and vivid when it is that body's own daylight-colour. And therefore minium [lead tetroxide] reflects rays of any colour but reflects most plentifully the ones associated with red; and consequently in daylight—·i.e. when lit up by all sorts of rays jumbled together·—-the red-associated ones will be the most plentiful in the reflected light, and their prevalence will cause the minium to appear red. [Newton repeats the same story for some stuff called 'Bise', which is blue. Then:] And it's clear that this is the whole cause of their colours, because they have no power to change or alter the colours of any sort of rays that fall on them, but present themselves with any colours they are illuminated with.

·CONCLUSIONS·

These things being so, there's no longer room for argument about •whether there are colours in the dark, or about •whether colours are qualities of the objects we see. Perhaps we can even put an end to the dispute over •whether light is a body. Colours are qualities of light; the entire immediate subject of a colour—·the item that *has* the colour·—is a ray of light; and we can't think of light-rays as qualities, unless one quality can be the subject of another quality, the item that *has* the other quality, and that amounts to calling it a 'substance'. How do we know that bodies are substances? Through their sensible qualities. If now the principal one of those sensible qualities turns out to belong to something else, we have as good reason to believe that the 'something else' is a substance also.

Besides, whoever thought that a quality could be a heterogeneous aggregate such as light is discovered to be? But it's not so easy to settle more absolutely •what light is, •how it is refracted, and •how it produces in our minds the images of colours. And I shan't mingle conjectures with certainties.

On reading what I have written up to here, I can see that this report will lead to various experiments by which it can be tested. So I shan't trouble you further ·with these· except to describe just one, which I have already mentioned in passing.

·ONE EXPERIMENT·

In a darkened room make a hole in the blind of a window—give it a diameter of about a third of an inch—to admit a convenient quantity of sunlight. Next to the hole place a clear colourless prism, to refract the entering light towards the opposite wall of the room, where it will be diffused into an oblong coloured image, ·the spectrum·. Then about four or five foot from the prism place a lens with a radius of about three feet and a thickness of about $2\frac{1}{2}$ or 3 inches. . . ., so that all the colours from the prism are transmitted through the lens and caused by its refraction to come together at a further distance of about ten or twelve feet. If at that distance you intercept this light with a sheet of white paper, you will see the colours converted back into whiteness by being mingled. The prism and the lens must be held steady, and the paper on which the colours are projected should be moved to and fro ·along the line from the centre of the lens to the centre of the paper·. As you move the paper slowly away from the lens you will not only find what the distance is at which the whiteness is most perfect, but will also see how the colours gradually come together into whiteness and then, having •crossed one another at the place where they make whiteness, are again separated and cut off from one another, again having the colours they had before they entered the composition, except that now their order is •inverted.

You can also discover that if any of the colours aren't caught by the lens, the whiteness ·on the paper· will be changed into the other colours. For the composition of whiteness to be perfect, take care that none of the colours misses the lens. [Newton now provides a diagram of this experiment; but it is hard to reproduce, and his description is clear enough without it.]

In my propositions **(3)** and **(13)** I have said that an uncompounded colour can't be changed. If you want to test that, you need to work in a room that is *very* dark; otherwise the colour you are testing will get mixed with others, which will defeat the purpose of the experiment. You will also need to separate the colours more completely than can be done—as in the previous experiment—by the refraction of a single prism. Anyone who attends to the discovered laws of refractions will find it easy enough to get this more complete separation. [Newton then goes on at some length about how experiment could be adapted to a situation in which perfect separation isn't achieved. He concludes:]

I think this is enough of an introduction to experiments of this kind. If any member of the Royal Society is interested enough to conduct some of them, I would be very glad to be informed about what the outcome was, so that if anything seems to be defective, or to falsify anything in the report I have given in this letter, I can have an opportunity to give further directions about it or to acknowledge my errors if I have committed any.