1

Historical perspective

The history of the sociological and scientific relationship between man and the Sun is far too lengthy to be chronicled in a single chapter. Therefore, in the following paragraphs I will try to summarize those events which in my experience have generated the majority of questions pertaining to the history of solar observations and research. Since much of the interest in the past and present is centered on the study of sunspots and phenomena which are associated with them I have emphasized this field, although I have tried to present these aspects in context with other major discoveries about the Sun.

The earliest records show us that primitive cultures frequently thought of the Sun as an individual, or as a person carrying a huge ball of fire. Since the Sun appeared regularly and didn't seem to change, a number of fanciful tales arose in connection with its behavior, which to the ancients was more like that of a slave than that of a master. In his 1949 book about the Sun, the famous solar astronomer, Donald H. Menzel, relates one legendary explanation for this apparent contradiction: 'It seems that the Sun was once very erratic. Sometimes he hurried too fast on his journey; at other times he dawdled. On occasion he came too close to the Earth; often he was too far away. Sometimes he failed to appear at all. Finally, after great difficulties have been surmounted, the Sun is caught in a trap or net, beaten into submission, and thereafter performs his duties without remonstrance.'

Later civilizations believed that the Sun was a deity and pictured it in extraordinary ways, often as a god transported across the daytime sky within a ship or chariot (Noyes, 1982), or regarded it as proof of the validity of their philosophical beliefs. Nowhere is this more evident than in the glorification of the Sun so loudly proclaimed by the powerful philosopher Aristotle, whose insistence upon its perfection profoundly influenced the thought and direction of early western society.

Because the Sun plays such an enormous role in our lives, it has always been viewed with fascination and curiosity. To its earliest observers, the Sun generally appeared to be nearly the same each day; however every once in a while a dark smudge would appear on its face which would cause great consternation before departing as swiftly as it came. According to Schove (1983), the earliest of all references to these large spots is contained in a translation of a Chinese oracle bone from the twelfth century BC: 'Will the Sun have marks? It really has marks ...'

Ironically, the first written record of a sunspot was made by a pupil of Aristotle named Theophrastus, who observed a spot on the Sun in the fourth century BC (Noyes, 1982). The Chinese, and pre-Historic Peruvian astronomers began to make and record their observations of the spots and other solar phenomena within a few centuries of Theophrastus's sighting, but for a variety of reasons the attempt to record them from Europe came much later. For that matter, even the historical observations made by the Chinese were not available in a European language until after 1873 (Giovanelli, 1984).

Today, many historians believe that the long delay in western reports can be traced to the teachings of Aristotle and his claim that the Sun was literally a ball of pure fire which could not be blemished. Despite the identification of sunspots as features on the Sun's surface by Aristarcus in the fifth century AD (Noyes, 1982), the strength of Aristotle's instruction combined with the unwillingness of academics to refute it, hid the truth for over a millennium. This unfortunate situation was not fully resolved until sometime after the first telescopic observations of the Sun.

The first observations of the Sun's spots were made with the unprotected eye during the morning or evening and when the Sun was obscured by smoke or haze so that its brilliance was somewhat diminished. (**The reader is cautioned against observing the Sun without adequate eye protection!**) The large sunspot groups which can be seen in this manner (Figure 1.1) typically occur

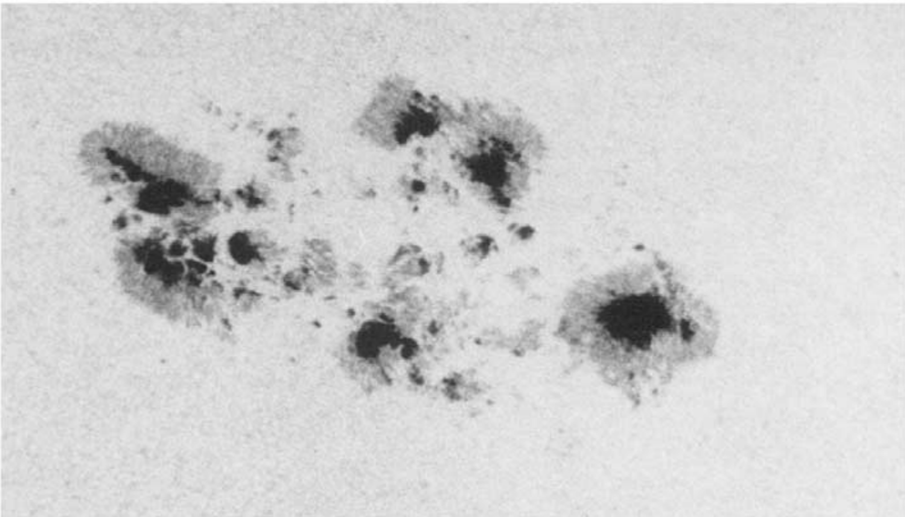

Figure 1.1 This large sunspot group grew to encompass an area of over 5.5 thousand million square kilometers during May 1990, making it easily observable with the unaided (but suitably protected) eye. Photograph courtesy of Jean Dragesco.

within a year or so of the sunspot cycle maximum. This aspect has proven to be a valuable one for those who study the spot cycle since it has allowed astronomers to determine the approximate dates of cycle maxima and minima which occurred a century or more before sunspots were first viewed telescopically.

Prior to the time that their physical connection with the Sun was universally acknowledged, observations of these 'naked-eye' sunspots were usually attributed to transits of known or undiscovered interior planets, or to phenomena which occurred in the Sun's or Earth's atmosphere. (A *transit* occurs on those occasions when an object passes directly between the Sun and Earth.) Even Galileo's friend, the famous mathematician and astrologer Johannes Kepler, incorrectly attributed a spot which he viewed in 1609 to a transit of Mercury, although he eventually acknowledged his error, allowing that he was 'mistaken.'

Although the spots were recorded in this simple manner for over two-thousand years, it was not until 1610 and the advent of the astronomical telescope that the real character of these secular features became apparent. At that time, Galileo Galilei, surely one of the greatest intuitive scientists that the world has ever known, seized upon and perfected the newly invented telescope and began to make routine observations of the Sun and other astronomical objects.

Shortly thereafter use of the device spread to several of Galileo's contemporaries, among them Goldsmid (Fabricus), Scheiner and Harriot. Their investigations, coupled with those by Galileo, soon began to unlock the secrets of the Sun. Unfortunately, because of social and political pressures brought about by the jealousy of his fellow astronomers and by his own combative nature, Galileo felt compelled to delay the announcement of his studies. As a result Fabricus became the first to publish findings which showed that sunspots were physically associated with the Sun (Bray and Loughhead, 1965). Interestingly, Fabricus observed the spots by projecting the Sun's image into a darkened room with a camera obscura, or pinhole camera, instead of viewing them directly.

Perhaps the real reason behind the reluctance of the powerful leaders of the day to acknowledge the true nature of these features had more to do with granting a scientist the right to teach and publish their views than with an affront to the supposed perfection of the Universe (Drake, 1957). But whatever the cause, the effect was the same: to temporarily retard progress in the area of scientific discovery and inhibit the flow of knowledge in the western world.

Even though Galileo delayed the dissemination of his conclusions regarding the nature of the spots, he is rightfully credited with an amazing degree of perception concerning many of their characteristics. He was almost certainly the first to interpret the Sun's spots as solar, rather than atmospheric or planetary phenomena, and realized early on that they almost always occurred in circum-solar bands which extend thirty degrees or so from the equator.

Furthermore, by 1612 Galileo had tracked the spots and the bright cloud-like

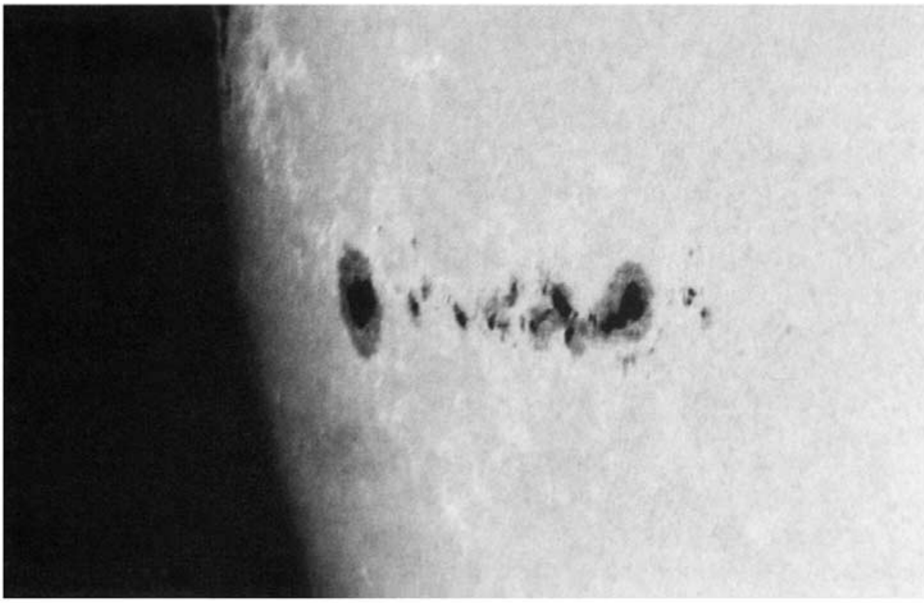

Figure 1.2 The light-appearing active features of the Sun called faculae are more easily seen in association with sunspot groups near the Sun's limbs, such as demonstrated in this fine photograph taken by Jean Dragesco.

features of the Sun called 'faculae' (Figure 1.2) across its disk, and by doing so had demonstrated that the Sun rotated on its axis. Moreover, he had skillfully argued these points (and effectively demolished his critics while doing so) in his replies to Scheiner, who believed that the spots were objects located between the Sun and Earth. Galileo accomplished this in a way which was typical of him; through a series of letters written to the wealthy merchant and amateur scientist Mark Welser, in response to those written anonymously by Scheiner (Drake, 1957). (Any reader who is interested in the enormous impact of Galileo upon society and the scientific community is urged to obtain Stillman Drake's fine discourse on this amazing individual.)

Most of Galileo's work on the Sun took place in the few years which followed the introduction and initial development of the telescope. On the other hand, Scheiner continued his observations for many years and eventually compiled the information, along with a number of fine engravings which showed the spots and faculae as they traversed the Sun's disk, into his great work, *Rosa Ursina sive Sol* (1630).

While Scheiner does demonstrate an understanding of the tilt of the Sun's axis in the publication, it is likely that Galileo was aware of the orbital relationship between the Sun and Earth many years beforehand. In the *Dialogue*, Galileo uses this argument to explain the movement of the Earth according to circumstances which could not have taken place after 1613 (Drake, 1957). In light of this, Scheiner's most important contribution to the study of the Sun is likely to have been the observation of the spots themselves, since these data have been used to help define the spot cycle; particularly the maximum which occurred in 1626.

Unfortunately, after the flood of new information which arose from the use of the telescope between 1610 and 1630, the Sun itself intervened in the discovery process with a long lull in activity that lasted into the following century. Hardly any spots were seen during this period which extended from 1645 until 1715, and eventually came to be known as the 'Maunder Minimum' after the English astronomer who investigated it in detail.

Then in 1769 a University of Glasgow professor and astronomer named Alexander O. Wilson completed what is regarded as the first scientific investigation into the *properties* of sunspots, rather than simply registering their number or location (Bray and Loughhead, 1965). Wilson found that the appearance of large spots near the Sun's limb was saucer-like, and explained his findings as depressions in the Sun's surface which were caused by a lack of material covering a dark and solid inner-core. In 1770 Wilson won the astronomical award given by the Science Society of the University of Copenhagen for this project. (The effect is now thought to be caused by the higher transparency of the spot material relative to the Sun's 'surface,' or photosphere.)

At the time that Wilson offered his findings, he believed that the Sun was surrounded by two atmospheres; an inner one blanketing a presumably cooler planet-like object where life might well exist (!), and a hot outer shell (Chambers, 1890). The outer atmosphere was luminous, and consequently it came to be called the *photosphere* (the luminous, or light-sphere). The photosphere is now known to be the visible exterior of the Sun; a region which is several hundred kilometers thick with a surface temperature that is approximately 5800 kelvin.[1]

Of course the photosphere is not the smooth glowing shell which Wilson and other early scientists envisioned. Instead it is composed of many bright areas, each of which is several hundred kilometers in diameter and is surrounded by darker lanes of cooler material (Giovanelli, 1984). These irregularly shaped features are known as the solar 'granules' (Figure 1.3), a term which is thought to have originated with Dawes in the nineteenth century who described the effect as 'irregular luminous clouds surrounded by less brilliant lanes ...' (Todd, 1899). A similar view was expressed in 1792 by the Danish astronomer, Thomas Bugge, who may have actually been the first to observe the effect, and also by William Herschel at about the same time (Webb, 1893).

The granules actually form the tops of huge currents which rise from beneath the Sun's exterior at a rate somewhere near 500 meters per second (Zirin, 1988) and erupt at the surface in a manner which is similar to a thick bubbling liquid. Typically they are short-lived; most have lifetimes which are measured in

[1] The kelvin scale is derived by adding the centigrade temperature to 273.15. Thus 100 degrees centigrade is equivalent to 373.15 kelvin.

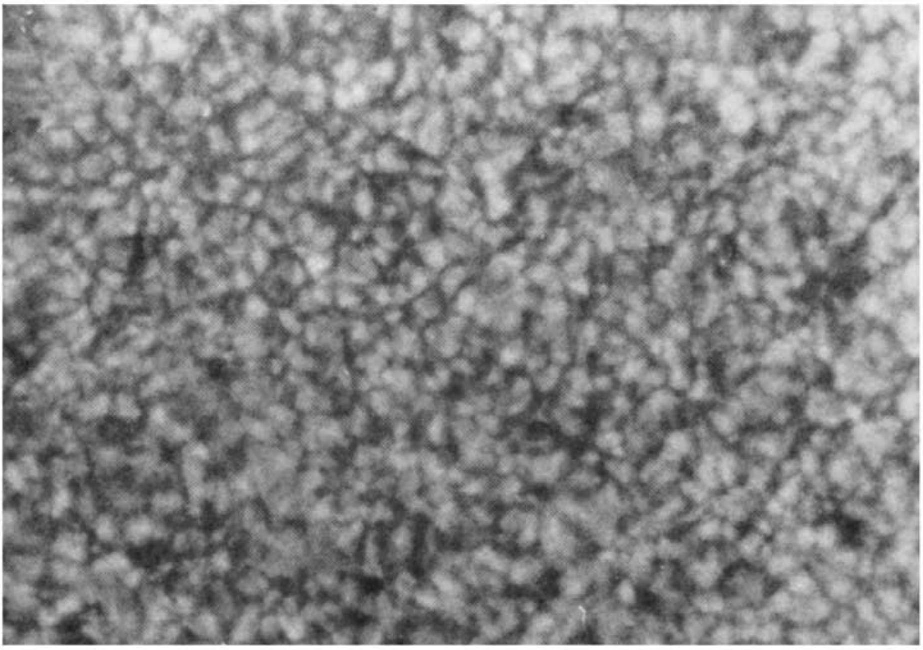

Figure 1.3 Most of the Sun's surface, or photosphere, is made up of a network of cells known as the solar granulation. Note the dark intergranular lanes which surround each of the cells. Photograph courtesy of Thomas G. Compton.

minutes before they cool and decay. When seen with lower magnifications, the granules produce a visual effect known as the 'rice-grain' pattern, a description which was first suggested by Huggins in his summary of the phenomenon at the Royal Astronomical Society meeting of 1866 (Huggins, 1866).

The Danish astronomer and director of the famous Round Tower Observatory, Christian Horrebow (1718–1776), was probably the first to suspect a periodicity in the numbers of sunspots (Vitinskii, 1965). In the last year of his life, Horrebow entered the following remarks into the observatory protocol, indicating that he believed a more systematic observation of sunspots might lead to 'the discovery of a period, as in the motions of the other heavenly bodies ...' and added, 'then, and not till then, it will be time to inquire in what manner the bodies which are ruled and illuminated by the Sun are influenced by the sunspots ...' (Young, 1888).

The study of sunspots was a subject of recurring interest at the Round Tower Observatory from the time that it was founded in 1637 on the recommendation of the astronomer and only pupil of Tycho Brahe, Longomontanus, until its eventual demise as a scientific institution. Peter the Great viewed sunspots from the Tower on several occasions under the guidance of Christian's father, Peder Horrebow, during his tenure as observatory director. News reports of the day indicate that the Czar actually rode the observatory's long wooden ramp to the top of the tower on horseback, accompanied by the Czarina who was driven up in a carriage drawn by four horses! (Thykier, 1988).

Unfortunately, the majority of Christian Horrebow's work was lost for a considerable period of time in the destruction of Copenhagen during the Napoleonic Wars. As a result, it was not until the mid-nineteenth century that the existence of a *sunspot cycle* was recognized. At that time the German amateur astronomer and apothecary, Heinrich Schwabe, published his long series of sunspot observations and suggested that their number appeared to vary regularly with a period of about ten years (Schwabe, 1849, 1851). His important discovery forms the basis for an often-told tale of scientific serendipity.

It seems that Schwabe's real purpose was to discover an additional planet orbiting the Sun within the path of Mercury. His forty-three-year record of sunspot observations was made in support of that research, so that he could differentiate between sunspots and the expected transit of his proposed planet. Although he never found his missing world, Schwabe's dedication and perception guaranteed him a place in astronomical history and caused him to react to his unexpected discovery by stating, 'I can compare myself to Saul, who went out to find his father's asses and found a throne ...' (Noyes, 1982).

Soon after Schwabe's announcement, the director of Bern Observatory, Rudolf Wolf, became interested in Schwabe's work and also began to study the spots. Wolf would have preferred to compute the spot group's areas rather than count the number of spots, but his equipment was not suited to this purpose. Consequently, Wolf developed an arbitrary counting system and called the resulting index 'Universal' sunspot numbers (Wolf, 1858). Eventually these came to be known as relative sunspot numbers, which today form the longest continuous record of the Sun's activity.

Through an extensive search of the early records, Wolf was able to derive a more accurate value of a little over eleven years for the average length of a cycle (Wolf, 1852). It is interesting to note that according to Schove (1983) Chinese observations of naked-eye sunspots between AD 188 and 1638 also show a period of around eleven years. During Wolf's exploration of the old observatory records, Horrebow's unpublished diaries were finally recovered, and the material was published shortly thereafter (Thiele, 1859). Horrebow's notes clearly showed the cycle minima which transpired in 1766 and 1775, and also indicated that a maximum had taken place around 1770 (the maximum actually occurred during the latter half of 1769).

In 1855 Wolf became director of the Swiss Federal Observatory in Zurich, Switzerland, where he and his successors, Wolfer, Brunner and Waldmeier, continued an expanded program of sunspot observation and research. During his term as director of the observatory, Wolf also instigated an international collaboration of observers whose data were employed when the Swiss weather was poor. However, the long series of values known as *Zurich Relative Sunspot*

Numbers are mainly the result of observations which were secured at the observatory in Zurich or by its official sub-stations at Locarno and Arosa, all of which employed identical instrumentation.

Almost a decade after Schwabe published his discovery, an English amateur astronomer named Richard Carrington presented the results of his research on the positions of the spots. Carrington, the son of a brewer, had originally intended to become a minister but became interested in astronomy while attending Trinity College in Cambridge. His interest in the Sun was sparked by Schwabe's identification of a period in the numbers of sunspots, and continued until he was forced to abandon observations and assume control of the family brewery in 1861 on the death of his father (Bray and Loughhead, 1965).

Carrington's observations showed that a relationship existed between the heliographic (solar) latitude of emerging sunspots and the phase of the newly discovered spot cycle (Carrington, 1858). This phenomenon, in which new spots erupt at high sunspot latitudes at the beginning of a cycle and then at progressively lower locations as the cycle matures, has come to be regarded as one of the principal characteristics of the sunspot cycle. During the time that Carrington compiled these observations, he also became one of the first to detect the Sun's differential rotation (Webb, 1893), which results from the Sun's gaseous make-up.

In the second investigation, Carrington found that the Sun rotates more slowly at polar latitudes where a single revolution requires about a month to complete, than it does at the equator where the rotational rate is some five days less. It is interesting to note that, according to Webb, Carrington determined the existence of both of these effects by using only a simple cross-wire within his instrument; no micrometer or other special measuring equipment was employed.

Eventually Carrington's findings concerning the emergence of spots during a cycle were corroborated by Gustav Spörer, and thoroughly investigated by Edward W. Maunder who demonstrated the existence of the latitude-effect beyond all doubt (Maunder, 1904). Consequently, the phenomenon is generally called *Spörer's Law*, while the graph which results when sunspot locations are plotted according to time is known as a *Maunder butterfly-diagram*. The original butterfly-diagram was drawn by Maunder and his wife Annie (Maunder, 1940), just prior to the Royal Astronomical Society meeting in 1904 where Maunder presented his famous paper, and is currently displayed in the library of the US National Center for Atmospheric Research.

It is intriguing (and not a little unusual!) that Carrington was also the first to observe and record a solar flare (Carrington, 1859). These rare, but spectacular events are known as solar white-light flares when they are seen in the normal visual portion of the spectrum, as this flare was. Fellow Englishman Richard Hodgson independently observed the same event which was followed by an intense geomagnetic storm and brilliant aurorae. Although Wolf had previously

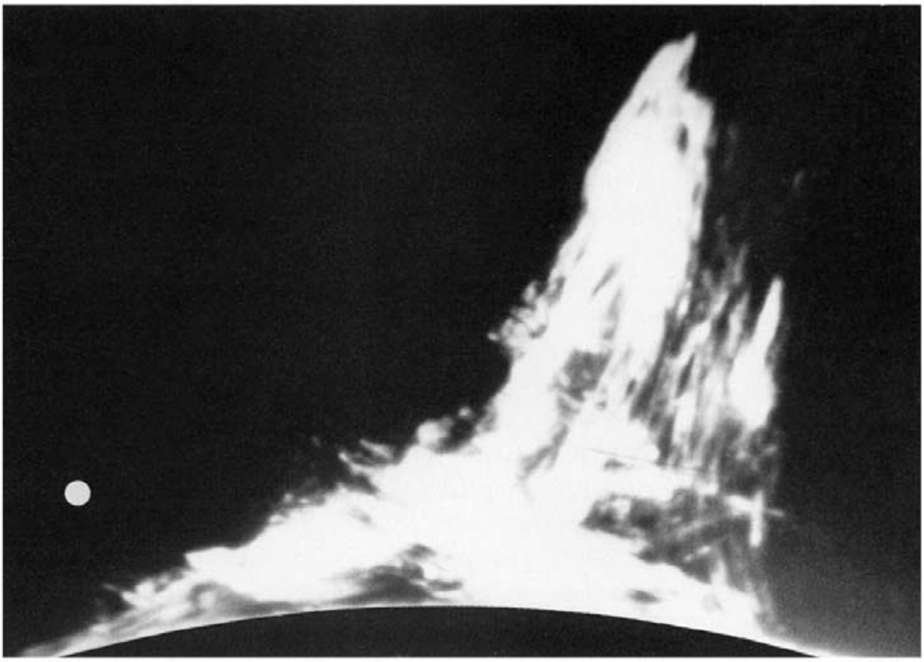

Figure 1.4 Near the Sun's limbs, the huge eruptions of solar gasses called prominences can be viewed in the red light of atomic hydrogen. The small filled circle indicates the relative size of the Earth in this photograph by Jean Dragesco.

uncovered a correlation between the sunspot cycle and geomagnetic disturbances, the dramatic effect of these combined occurrences formed the impetus for future investigations into the true nature of the solar–terrestrial relationship (Noyes, 1982).

At the 1868 eclipse of the Sun in India, French astronomer P. J. Janssen made another important discovery. During the eclipse, Janssen had observed a huge eruption of gas on the Sun's limb, a feature known as a solar prominence (Figure 1.4). He had viewed the phenomenon spectroscopically, carefully noting the presence of bright spectral lines near the wavelength of atomic hydrogen in the red portion of the spectrum.

After the eclipse was over Janssen again pointed his instrument towards the location of the prominence, and again he found a bright area. After the spectroscope's slit was opened slightly, the shape of the prominence could clearly be seen. The identical discovery was made independently and virtually simultaneously by the English astronomer Norman Lockyer (who did not attend the eclipse), and ultimately led to the invention of the spectrohelioscope. After this instrument was developed, details on the Sun's disk could be discerned for the first time, and Janssen was rewarded with the directorship of France's new Meudon Astronomical Observatory.

While he was director of the Meudon Observatory, Janssen obtained a number of extraordinary photographs of the solar granulation. (No photographs of the Sun of any type were made until after 1870.) A few of the photographs appeared to show even the finest granular structure, but eventually this was found to be a consequence of distortions produced in the Earth's atmosphere. As a result, the first high-resolution photographs of the granular network were delayed until the balloon experiments which were conducted by Schwarzschild (1960).

At about the time that Janssen was photographing the solar surface, American physicist Jonathan Lane developed an idea which showed that the Sun was not the planet-like object which Herschel and Wilson had envisioned. Instead, Lane suggested that the Sun was entirely composed of gas held together by the force of its own gravity, and driven by an internal source of energy. The origin of this dynamic mechanism was originally characterized as nuclear in nature by Arthur Eddington in 1926, but a complete definition of the process was delayed until after the end of World War II (Giovanelli, 1984).

George Ellery Hale, the guiding force behind the construction of the Yerkes, Mount Wilson and Palomar Observatories, and a developer of the spectrohelioscope, discovered the magnetic nature of sunspots in 1912 (Hale, 1912). Hale accomplished this by noting that the spectral lines which arise from the spots are split into two or three separate lines. The same spectral line-splitting and polarization in sunspot spectra is also seen in laboratory spectra when their source is exposed to a strong magnetic field. Hale correctly deduced that the presence of this condition (known as the *Zeeman-effect* after the Dutch physicist who discovered it in 1896) in sunspot spectra indicated that the spots are magnetic phenomena.

By 1919, Hale and his co-workers at Mount Wilson had gone on to describe the polarities of individual sunspots and their reversal from cycle to cycle (an effect which is termed the *Hale–Nicholson Law*), and had established the Mount Wilson magnetic classifications (Hale *et al.*, 1919). In spite of its age the Mount Wilson system continues to be used by professional astronomers, although because special equipment is required to obtain these data, it is not suitable for use by amateur astronomers.

In 1938 M. Waldmeier devised the evolutionary sunspot-group classification scheme for visual observers which continues to be in widespread use today (Waldmeier, 1947). This technique for grouping sunspots into clusters according to their physical appearance and size is generally referred to as the 'Zurich Classification' system. The categories which are assigned to spot groups according to this method have proven to be especially valuable for studies of the growth and decay of spot clusters as well as for research into the physical characteristics of sunspots.

In the early 1950s, Ludwig Biermann (1951) suggested that there was