

A Zoo of Galaxies

Karen L. Masters^{1,2}

¹Institute for Cosmology and Gravitation, University of Portsmouth, Dennis Sciama Building,
Burnaby Road, Portsmouth, PO1 3FX, UK
email: karen.masters@port.ac.uk

²South East Physics Network, www.sepnet.ac.uk

Abstract. We live in a universe filled with galaxies with an amazing variety of sizes and shapes. One of the biggest challenges for astronomers working in this field is to understand how all these types relate to each other in the background of an expanding universe. Modern astronomical surveys (like the Sloan Digital Sky Survey) have revolutionised this field of astronomy, by providing vast numbers of galaxies to study. The sheer size of these databases made traditional visual classification of the types of galaxies impossible and in 2007 inspired the Galaxy Zoo project (www.galaxyzoo.org); starting the largest ever scientific collaboration by asking members of the public to help classify galaxies by type and shape. Galaxy Zoo has since shown itself, in a series of now more than 30 scientific papers, to be a fantastic database for the study of galaxy evolution. In this Invited Discourse I spoke a little about the historical background of our understanding of what galaxies are, of galaxy classification, about our modern view of galaxies in the era of large surveys. I finish with showcasing some of the contributions galaxy classifications from the Galaxy Zoo project are making to our understanding of galaxy evolution.

1. What are Galaxies?

To say our understanding of the “zoo” of galaxies that are found in our Universe has changed a lot over the last century or two is a bit of an understatement. In 1845 the state of the art picture of an external galaxy, was an image of M51, or the Whirlpool galaxy drawn by William Parsons, Third Earl of Rosse (1800-1867), looking through what at the time was the largest telescope in the world - the “Leviathan of Parsonstown” at his castle in Ireland (Left panel Figure 1; for a discussion of the history of picture see Steinicke 2012). At the time no-one fully understood that this was an external galaxy made of billions of stars. Although one of the motivations Parson is said to have had to build the telescope, was to resolve structure in distant nebulae to see if they could be “island universes”, rather similar in structure to our own galactic home. Today we have images of millions of galaxies from large surveys like the Sloan Digital Sky Survey (e.g. the last imaging release in Data Release Eight, Aihara *et al.* 2011), and extraordinarily high resolution images of hundreds of galaxies, including the Whirlpool, from the Hubble Space Telescope (Fig 1, right), as well as information on galaxies extending back more than half of the age of our Universe.

It took a long time for astronomers to understand that stars in the universe are organized in collections we now call galaxies. I can’t help thinking that it was a major leap in understanding for astronomers to connect the uneven distribution of stars in the night sky with the three-dimensional structure of the galaxy that we live in. The first published example of this idea is the map shown in Figure 2, published by William Herschel in 1785 (Herschel 1785), and based on star counts made by himself and his sister Caroline. This diagram demonstrates an understanding of the Galaxy as a collection of stars, and while there is a lot wrong with it (for example the Sun is at the centre, and the whole thing

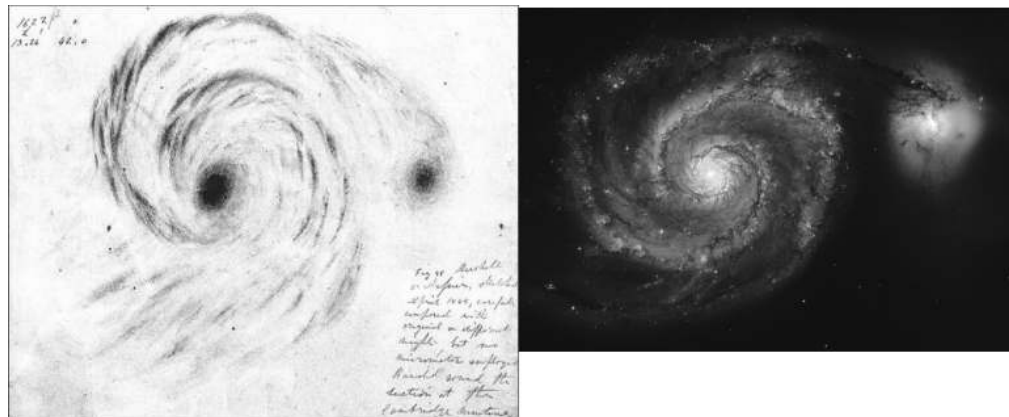


Figure 1. Two views of M51 (the Whirlpool Galaxy). On the left is the original 1845 drawing by William Parsons (Figure 8 from Steinick 2012). On the right is an image from the Hubble Space Telescope (Credit: NASA, ESA, S. Beckwith (STScI) and the Hubble Heritage Team (STScI/AURA)).



Figure 2. The first map of the Milky Way showing it as a collection of stars, and based on observational data. (Herschel 1785)

was much too small at ~ 7000 light years across) its an extraordinary piece of work as the first example of such a map based on actual astronomical data.

By 1900 astronomers understood quite a lot more about the basic structure of our galaxy. The map made by Cornelius Easton in 1900 (Fig 3) was the first to show our galaxy as having spiral structures (Trimble 1995). Easton used pictures of other spiral galaxies he saw in the sky to suggest the Milky Way might have this structure, although he still incorrectly placed the Sun in the centre of the Galaxy, and it's still too small.

The variety of different galaxies observed in the sky naturally caused people to wonder what they were. The scientific arguments surrounding this question at the start of the 20th century, are best represented perhaps, by the public debate held in 1920 between Heber Curtis and Harlow Shapley. Many other authors have described this debate, in much more detail that I intend to (for example Trimble 1995), and as is well known, both astronomers had it partly right. While Curtis held the opinion that the spiral nebula represented external galaxies, his measure of the size of our Milky Way was much too small (only 10 kpc across) and he still placed the Sun at its centre. Shapley believed the Milky Way was much larger, and the Sun offset (and in this was right), but because of the vast size of our Galaxy seemed unable to conceive that the Universe could be large enough to contain millions of other similar galaxies.

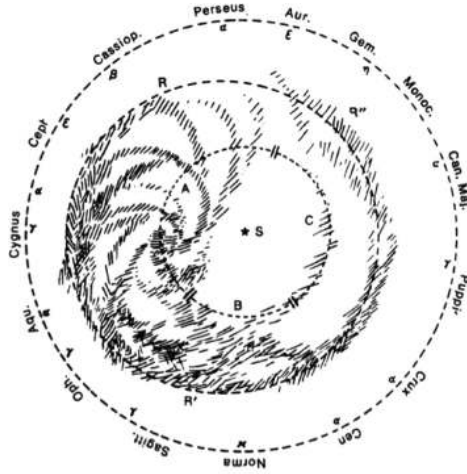


Figure 3. Cornelius Easton's model of the Galaxy in 1900 (reproduced from Fig 2 of Trimble 1995)

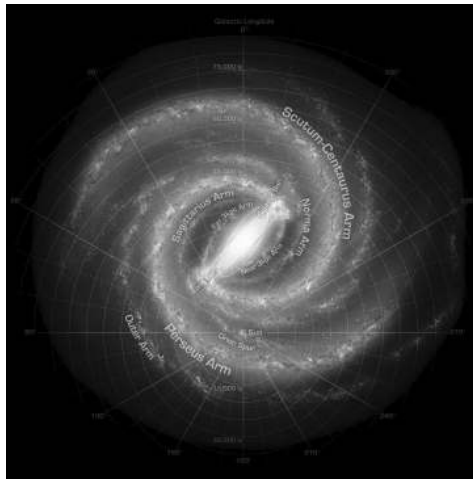


Figure 4. Modern view of the structure of our Galaxy. Image Credit: NASA/JPL-Caltech

The modern view of the structure of our Galaxy is presented in Figure 4, as an artists impression based on counting stars in data from the Spitzer Space Telescope observations (Benjamin 2008). Our Galaxy is a classic example of what we call a spiral galaxy, with most of its stars in a large, flat and very thin disc like structure which shows spiral arms, and a rounder region in the centre we call the “bulge”. Our Galaxy also shows evidence for an elongated “bar” of stars across its central parts, a structure which is seen in many, but not all spiral galaxies. The Milky Way demonstrates just one of the two main kinds of big galaxies that are found in our Universe. Particularly in high density regions of the Universe, for example in the core of the nearby Virgo cluster of galaxies, many galaxies are large, smooth and spheroidal (or elliptical on the sky) in shape, and we call these types “elliptical” galaxies.

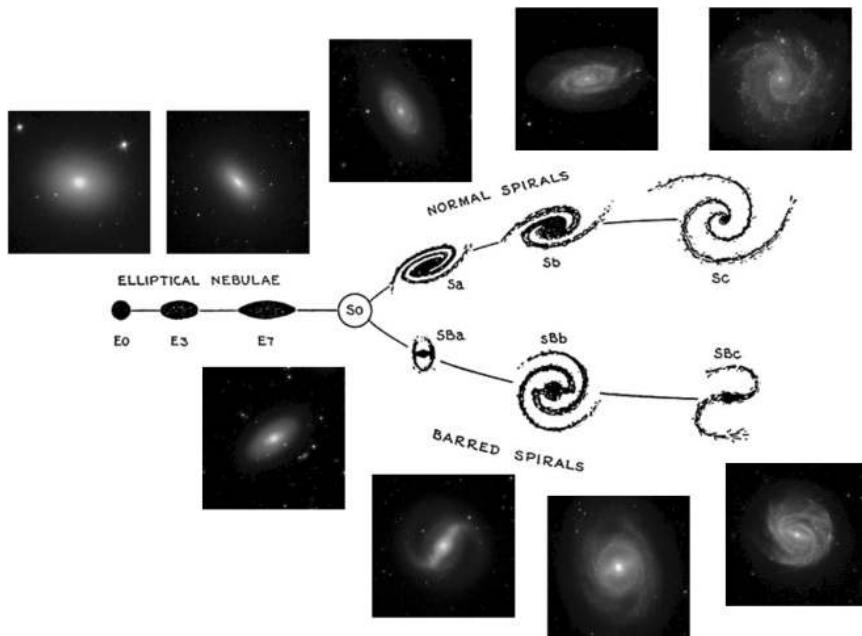


Figure 5. Hubble's original tuning fork as he presented in Hubble (1936), with *gri* images of galaxies from the SDSS (similar to those used in the Galaxy Zoo project) added to further illustrate each galaxy type.

2. Classification of Galaxies

A common scientific approach to understanding a large collection of objects better is to classify it, then make categories. Much of the early effort in the field of extragalactic astronomy went into classifying the types of galaxies seen in the sky.

Probably the most famous of the early galaxy classifiers was Edwin Hubble (1889–1953). He was not the only person at the time to develop a classification scheme, but his scheme has been the most long lasting. It has been suggested that this was because he made such broad categories, that most galaxies can fit into one of them (Buta 2011). In a defence of the scheme against criticisms made by J. H. Reynolds (Reynolds 1927), Hubble claimed to have looked at more than a thousand images of galaxies, with only a small fraction not fitting the scheme, and uncertainty in placement in “less than ten per cent” (Hubble 1927).

In this same paper, Hubble described the scheme (which he first published in diagram form in *The Realm of the Nebula* in 1936, as shown in the centre of Figure 5). Hubble writes: “The classification under discussion arranges the extra-galactic nebulae in a sequence of expanding forms. There are two sections in the sequence, comprising the elliptical nebulae and the spirals respectively, which merge into one another.” Ellipticals were ordered in how elliptical they appeared (from E0 being roundest, to E7 most elliptical), and spirals were ordered with reference to “conspicuous structural features”, notably (1) the size of the central bulge, (2) the extent to which the arms wound, and (3) how distinct the spiral arms were. Hubble additionally notes that the “barred spirals” made up a distinct type of spiral ordered in the same fashion, but comprising “a small fraction of the numbers of normal spirals” (at least in the photographic plates he used). In between the spirals and ellipticals was the (then hypothetical) S0, or lenticular type,

which shows a disc, but no spiral arms. The small number of galaxies Hubble found not fitting this scheme were called “irregular”.

With only minor modifications and extensions (see e.g. Buta 2011 for an extensive discussion) the Hubble classification scheme as presented is still widely used in modern extragalactic astronomy. We don't just keep using it because its simple and it works well, but also because we have discovered that the actual physical properties of galaxies tend to vary in predictable ways along the sequence (e.g. as shown in Roberts & Haynes 1994). It's a common misconception that Hubble labelled galaxies at the left of the diagram “early” and those on the right “late” in an attempt to imply an evolutionary sequence. In fact Hubble himself writes that no such conclusion should be drawn (saying that “temporal connotations are made at one's peril” and that he set up the classification “without prejudice to theories of [galaxy] evolution”, Hubble 1927), but rather he used the labels by analogy to what was then a commonly used classification scheme for stars. These terminology are still commonly used by astronomers today to discuss relative positions of galaxies on the diagram, with “early-type” galaxies used to mean elliptical and lenticular types, and “late-types” for spiral galaxies. Even among spirals, early and late will be used interchangeably with the Sa, Sb, Sc notation to describe relative location along the spiral sequence.

3. Physical Properties of Galaxies

A galaxy is a massive collection of stars, gas and interstellar dust. To first order, the optical colour of a galaxy tells about the types of stars which live in the galaxy (ignoring the complications of dust which can redden some types of galaxies quite significantly e.g. Masters *et al.* 2003, 2010a). When stars form, they form in a variety of masses. The most massive stars are extremely bright, and hot and so they look blue/white (in the same way very hot metal glows blue/white). They also live for very short times (astronomically), by which we mean just a few 100 million years, and when they are visible in a galaxy the galaxy will look blue because they totally outshine everything else.

If star formation ceases, after a short time all the massive hot blue stars will die, and all that is left in the galaxy will be the less massive, cooler, red stars which can have lifetimes of tens of billions of years. So a galaxy which hasn't made new stars for while (more than a few 100 million years) will look red, as only red stars still shine in that galaxy.

Back to the Hubble diagram – the ellipticals on average are more massive (i.e. have more stars), redder, and tend to be found in clusters, while the spirals tend to be less massive, bluer and not in clusters (and increasingly so as you move along the Hubble Sequence away from the ellipticals). This tells us that the spiral galaxies are still forming stars, while the ellipticals have stopped, and more generally the Hubble sequence reveals a sequence of increasing star formation (and potential for future star formation) from left to right (Roberts & Haynes 1994), as well as decreasing (average) total galaxy stellar mass.

Hubble claimed almost all galaxies he looked at fitted his sequence, however we have observations of a lot more galaxies now. Thanks to improvements in detector technology and computers able to automatically process the data, very large surveys of the sky revealing millions of galaxies have become possible, including many much fainter than those available to Hubble, and revealing extra categories of dwarf, irregular and low surface brightness galaxies.

4. How Many Galaxies are There?

In his 1927 defence of the classification scheme, Hubble claimed to have examined “upward of a thousand galaxies” in its construction. However for true physical understanding of a galaxy, more than just an image is required – we also need an estimate of its distance in order to reveal its size and mass. Hubble is most famous for a discovery which revolutionised extragalactic astronomy in its ability to provide a relatively quick and easy way to make these distance estimates. In 1929, Edwin Hubble used a sample of just 24 nearby galaxies to make this discovery. He used the Doppler shift of spectral lines to measure recessional velocities and used distances to these galaxies estimated based on the brightness of certain types of stars in them. By plotting these results against each other he demonstrated that the further away a galaxy is the faster it appears to be moving away from us (a finding now called “Hubble’s Law”, Hubble 1929). He got the proportionality constant drastically wrong (finding what we now call Hubble’s constant to be 465 ± 50 km/s/Mpc, while around 70 km/s/Mpc is the current accepted value, e.g. as we measured in Masters *et al.* 2006), but the main finding persisted, and in proceeding decades has generated an industry of using recessional velocities of galaxies (or “redshifts” since they are revealed by a reddening of known spectral lines via the Doppler shift) to map the universe. Even to this day it remains challenging to estimate redshift independent distances to galaxies, but the spectral measurements to indicate redshifts have become relatively routine. In the first 57 years after Hubble’s original findings the number of galaxies mapped had increased by a factor of 50. In 1986 the state of the art in this field was the first slice of the CfA Redshift survey (de Lapparent, Geller & Huchra 1986). Using a sample of 1100 galaxies with redshifts this survey demonstrated that galaxies in our universe are not uniformly distributed, but rather clump in “large scale structure”. Only 22 years after this survey, the state-of-the art was a sample 1000 times larger than this – the final release of the SDSS Main Galaxy Sample (Strauss *et al.* 2002; Abazajian *et al.* 2009) consisting of almost 1 million galaxy redshifts. Ongoing surveys (e.g. the Baryon Oscillation Spectroscopic Survey of SDSS-III, Dawson *et al.* 2012) are working on increasing this even more.

5. The “Zoo of Galaxies” in the Era of Large Surveys

Modern large surveys of hundreds of thousands of galaxies with detailed images and spectra (like the SDSS Main Galaxy Sample) have inspired a new field in extragalactic astronomy – the use of statistical analyses to reveal demographics of the population of galaxies. Perhaps the most famous example of this in the field of extragalactic astronomy is the use of the colour-magnitude diagram as a basic observation of galaxies (e.g. Strateva *et al.* 2001, Bell *et al.* 2003, Kauffmann *et al.* 2003, Baldry *et al.* 2004, Balogh *et al.* 2004). The galaxy population show a striking bimodality in this diagram, with galaxies mostly found in two regions, which have become known as the “red sequence” and the “blue cloud”, with a sparsely populated “green valley” in between (e.g. a fraction of the galaxies of the SDSS Main Galaxy Sample are shown on this diagram in Fig 6). This diagram demonstrates a general trend among galaxies that bigger (or intrinsically brighter) galaxies tend to be optically redder (ie. having an older stellar population). This trend is most striking in the red sequence, but also apparent in the blue cloud. Hubble Space Telescope surveys have demonstrated that this colour magnitude diagram is in place since at least $z \sim 1$ (Bell *et al.* 2004). They also show that at earlier times in Universe more galaxies were found in the blue cloud – demonstrating that on average galaxies move from blue to red as cosmic time progresses. Much of extragalactic

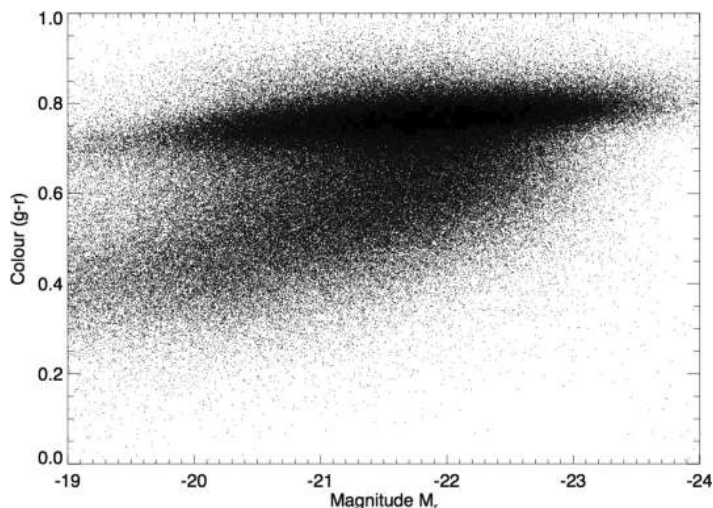


Figure 6. A colour magnitude diagram of a random selection of the SDSS Main Galaxy Sample (Strauss *et al.* 2002). This plots the $(g-r)$ colour versus the r -band absolute magnitude of almost 300,000 randomly selected galaxies in the Data Release 6 version (Adelman-McCarthy *et al.* 2008).

astronomy in recent decades has focused on developing an understanding of the mechanisms which shape the locations galaxies are found in diagram and the physical processes which move them around on it.

The bimodality in colour mirrors the two galaxy types discussed above, with the blue cloud hosting mostly spiral galaxies, and the red sequence mostly elliptical and lenticular galaxies. What had been missing in making this statement though was morphology (or type) for more than a few thousand objects (e.g. Fukugita *et al.* 2007 classified ~ 2000 of the SDSS galaxies). It seemed impossible to do that for the sample sizes being generated by astronomers in the early part of the 21st century (although the MOSES project tried – visually inspecting 50,000 galaxies to search for blue ellipticals, Schawinski *et al.* 2007). Without reliable information on galaxy morphology however, it's not possible to have the full picture of galaxy evolution. The colour of a galaxy is driven by the stellar (and gas and dust) content of the galaxy, while the shape or morphology of a galaxy reflects its dynamical history which could be very different (and have a different timescale). Therefore, one of the central motivations for the original Galaxy Zoo project was to construct a large sample of early and late type galaxy classifications from SDSS that were independent of colour.

6. Galaxy Zoo

Galaxy Zoo in its original form was launched on July 11, 2007 and introduced in a BBC online article that same day[†]. The original site simply asked volunteers to classify galaxies as either spiral or elliptical – the most basic morphological split among galaxies. Something in Galaxy Zoo resonated extraordinarily with the general public. The original projection estimated that if a few thousand members of the public got involved the 1 million galaxies might be classified in a couple of years. However, within twelve hours of the launch, the Galaxy Zoo site was receiving 20,000 classification per hour. After forty

[†] *Scientists seek galaxy hunt help*, by Christine McGourty
 (<http://news.bbc.co.uk/1/hi/sci/tech/6289474.stm>)

hours, the classification rate had increased to 60,000 per hour. After ten days, the public had submitted ~ 8 million classifications. By April 2008, when the Galaxy Zoo team submitted their first paper (Lintott *et al.* 2008), over 100,000 volunteers had classified each of the $\sim 900,000$ SDSS galaxy images an average of 38 times.

The popularity of Galaxy Zoo and the number of classifications received enabled science which simply would not have been possible without its contribution. Not only does Galaxy Zoo give a classification for each galaxy, but also by collecting ~ 40 independent classifications of each galaxy Galaxy Zoo produces an estimate of how likely that classification is to be true (a classification “likelihood”). All of the data from this first phase of Galaxy Zoo were published in Lintott *et al.* (2011) and are available to download via the SDSS servers[†]. For an in-depth history and analysis of the Galaxy Zoo project see Fortson *et al.* (2012).

6.1. Red Spirals

One of the biggest contributions of Galaxy Zoo has been in finding large samples of rare classes of extragalactic object in the SDSS sample. An example of this are the “red spirals” – relatively rare spiral galaxies which are found on the red sequence (first discussed in Bamford *et al.* 2009, Skibba *et al.* 2009). Due to the reddening effects of dust, a significant number of edge-on spirals are found in the red sequence, and spiral galaxies with large central bulges can also be intrinsically very red due to old stars in the bulge (Masters *et al.* 2010a), however Galaxy Zoo revealed a significant fraction of even late-type spirals found on the red sequence (Masters *et al.* 2010b). These are galaxies which have small bulges, but intrinsically red discs yet still show clear spiral arms. They provide a direct probe of evolution affecting star formation but not morphology, revealing that processes exist which can turn spiral galaxies red without disturbing their morphology. Even though these are relatively rare objects in the galaxy population, studies have suggested they form a significant part of the route for most of the evolution from the blue cloud to the red sequence both with increasing density in the local universe (Bamford *et al.* 2009) and with redshift (Bundy *et al.* 2010). In Masters *et al.* (2010b) we revealed that while these objects are most common in intermediate density regions (as also shown in Bamford *et al.* 2009 and Skibba *et al.* 2009) they are found even at very low densities. We demonstrated that compared to blue cloud spirals, red spirals become more common as spiral galaxies become more massive, that they are not significantly more dusty (as revealed by Balmer decrements) but are significantly more likely to host LINER (Low Ionization Nuclear Emission Region) like emission and obvious bars. The Galaxy Zoo red spirals are not completely passive, but at fixed stellar mass show significantly less star formation and an older stellar population than their blue cousins.

The red spirals have provided part of the evidence prompting a move from the view of major mergers as the main route of galaxy evolution (e.g. as presented in Steinmetz & Navarro 2002), to more gentle and slower (secular) processes playing a significant role for most galaxies. In order to make a spiral galaxy red (by stopping star formation), something has happened to these galaxies to exhaust, or remove their supply of atomic hydrogen gas (the raw material for star formation). It’s been discussed for many years that in our galaxy the amount of gas in the disc will be used up by the current rate of star formation in much less than a Hubble time (Larson, Tinsley & Caldwell 1980), and evidence exists that this gas is being replenished by the infall and cooling of hot gas from the halo (for a review see Putman *et al.* 2012). If this gas supply is shut off (either by being removed, or heated by processes like strangulation or harassment, Larson, Tinsley

[†] <http://skyserver.sdss3.org/CasJobs/>

& Caldwell 1980, Balogh, Navarro & Morris 2000, Bekki *et al.* 2002) star formation will cease and the disc redden on a timescale of ~ 1 Gyr. In fact recent work modelling the star formation histories of the Galaxy Zoo red spirals compared to blue spirals and red ellipticals supports this picture, showing that the average star formation history of red spirals only differs from blue spirals in the last 1 Gyr (Tojeiro *et al.* 2012).

One of the most striking observations about the red spirals was that such a large fraction of them were very strongly barred (75% Masters *et al.* 2010b). In much of the study of galaxy evolution, the division between barred and unbarred spirals has been ignored. It has often been argued that these types of galaxies are intrinsically the same, just caught with or without a bar (e.g. as recently discussed by van den Bergh 2011), even though most theoretical considerations now suggest bars must be very long lived. Around 30-60% of massive spirals host bars, with the exact value depending on how strict you are with the definition of what a bar is, which can vary from a linear feature stretching across most of the galaxy (as in the classic example NGC 1300), to mild oval distortions in the central regions.

6.2. Galaxy Zoo 2 and Barred Spirals

At the time we noticed the large bar fraction in red spirals, data from the second phase of Galaxy Zoo was starting to become available. Galaxies can show all sorts of interesting structures beyond the simple split between spiral and elliptical, and these features reveal more clues to the formation histories of the galaxies. Galaxy Zoo was so popular, and the results so reliable (agreeing with experts just as well as experts are able to agree with each other, as discussed in Lintott *et al.* 2008) that a new version was developed, asking for significantly more detail for the brightest quarter of the original sample. If a galaxy was identified as being a disk or showing features, this version asked questions about the number of spiral arms, size of the bulge, and most importantly for us, the presence of a bar. The full classification scheme for Galaxy Zoo 2 (GZ2) is shown in Figure 7, reproduced from Masters *et al.* (2011). This version of the site launched in February 2009 and ran for fourteen months, collecting in that time 60 million individual classifications of the images.

The first science result from GZ2 were a study of the bar fraction of disc galaxies as a function of colour. This work (Masters *et al.* 2011) demonstrated that the types of bars revealed by GZ2[†] were significantly more common in redder, more massive disc galaxies – with the extreme of this being the red spirals, among with 75% showed bars. This work also revealed a split in bar fraction between disc galaxies with large and small bulges (as revealed by fits to the light profile from SDSS) such that bulge dominated spirals showed more bars. These types of trends of bar fraction with galaxy colour have also been seen in the local universe by Nair & Abraham (2010) and Giordano *et al.* (2010) and at higher redshift by Sheth *et al.* (2008). The GZ2 results also hinted at an upturn of the bar fraction in the bluest and lowest mass disc galaxies as observed more clearly by Barazza *et al.* (2009) and Aguerri *et al.* (2009) in disc galaxy samples dominated by these kinds of galaxies.

The striking thing about our theoretical understanding of bars in disc galaxies, is that forming a bar in a disc galaxy is extremely easy. The question is not why some galaxies make them, but why some do not (or have not yet).

While a bar is forming it enables the exchange of angular momentum in a disc galaxy – basically moving material around. This has the effect of growing central

[†] which have been shown to be very similar to the classical “strong” bars identified in early visual catalogues - see Masters *et al.* (2012) for a discussion.

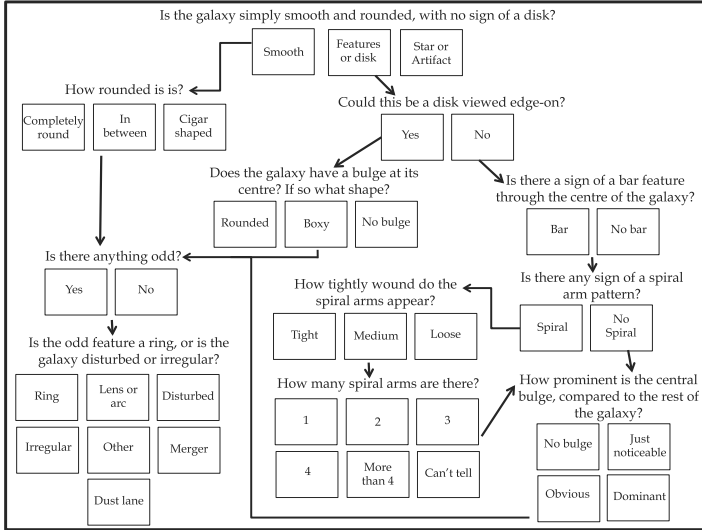


Figure 7. Full classification tree from Galaxy Zoo 2. First published in Masters *et al.* (2011).

concentrations (Kormendy & Kennicutt 2004), sparking central star formation (e.g. as observed by Sheth *et al.* 2005, Ellison *et al.* 2011), and possibly helping to feed central active galactic nuclei (although this is more controversial e.g. Ho *et al.* 1997, or Oh *et al.* 2012, Cardamone *et al.* in prep.). Theoretical considerations indicate that the gas content of a galaxy is important in this process (e.g. Athanassoula 2003; Combes 2008). By virtue of being able to dissipate energy, gas is an important sink of angular momentum, and any galaxy with a significant gas fraction in simulations struggles to develop a large bar (e.g. Villa-Vargas *et al.* 2010), while in a simulation of a gas poor disc, bars can be very stable and long lived.

The effective forces produced by the bar instability act to drive gas inwards from co-rotation (the point at which stars in the disk rotate with the same speed as the pattern speed of the bar) to the central regions. This gas loses its angular momentum which is transferred to the stars in the bar. Interestingly, the forces outside the ends of the bar may also act to inhibit inflow of gas from the outer regions of the disc, so that gas inflow of external gas onto a disc galaxy is inhibited in the presence of a strong bar (Combes 2008). It is possible that this process could play a role in the global reddening of disc galaxies, and while the concentration of gas in the centres of galaxies caused by the bar will spark central star formation, it evacuates gas from the rest of the disc (preventing star formation there) and potentially helps to use up the total gas content of the galaxy quicker.

6.3. Bars and Atomic Gas Content

The discovery that the type of strong, or obvious bars easily identified in GZ2 classifications are more common in more massive, redder disc galaxies, already implies that bars are more often found in gas poor disc galaxies, since optical colour is a good proxy for current star formation rate, which correlates well with gas content. However we wanted to investigate more directly the role of gas content on the likelihood of a galaxy hosting a bar.

The Arecibo Legacy Fast ALFA (Arecibo L-band Feed Array) survey (ALFALFA; Giovanelli *et al.* 2005) provide the ideal data to do this, matched well to the available GZ2 bar sample. ALFALFA has mapped all of the high Galactic latitude sky visible