REPORT and *RESPONSE*

The paper "The 2dF Galaxy Redshift Survey: The b_J-band galaxy luminosity function and survey selection function" is a nice analysis of the luminosity function. There are several details which **must** be fixed, however, before publication.

1 Requirements:

1) The calibration of the 2dFGRS has to be described SOON, preferably in this paper. In this paper and others, the authors have noted that the calibration is designed so that $< m_{APM} - m_{CCD} >$ is zero and constant with apparent magnitudes. Maddox et al 1990b discusses an OVERALL correction to the magnitude using CCD magnitudes (ie. it tells us that $b_j = 29.4 - m$), but does not discuss anything else using CCD magnitudes. Please devote a paragraph or two to discussing in DETAIL what is done and how someone could fairly mimic its effects, either in simulated or real data. If the calibration method is so important to understanding the comparison between your experiment and those of others, it must be explained carefully and completely. Otherwise it will be difficult to perform the fundamental scientific task of comparing the results of different experiments. Since theory can in principle predict galaxy profiles, luminosities, and colors, one would also like to be able to predict the outcome of your experiment from a theory of galaxy formation. It is not acceptable to wait for Maddox et al 2001 unless that paper is in press right now; at any rate, the closer that paper is to publication, the easier it will be to include more information on the calibration in this paper.

We agree with the referee that is far from ideal to have the calibration of the 2dFGRS described in a paper that is not yet available. As the Maddox et al. 2001/2002 paper is still not complete we have included, in Section 2, a new, quite detailed, description in the revised version of this paper. There are three parts to this new material:

i) **Section 2.1:** *A brief review of the steps involved in the calibration of the APM magnitudes. This is very brief and mainly refers the reader to the extensive description and variety of tests that are presented fully in the APMII paper (Maddox et al. 1990a). This section now also includes a reference to another paper (Maddox etal 1990b). that explicitly shows the relation between the APM magnitudes and the calibrating CCD data. While there is nothing new in this section it is perhaps useful as it highlights the fact that APM magnitudes cannot be treated as isophotal magnitudes. This seems to be a common misconception.*

ii) **Section 2.2:** *A short section describing revisions to the calibration and the dust corrections that were applied to define the input catalogue of the 2dFGRS.*

iii) **Section 2.3:** *The final section describes a number of tweaks to the calibration that were done prior to the release of 100k data. These changes are minor in the SGP region which was in the original APM survey, but are quite important for the NGP region.*

The referee suggests we describe how the APM data maybe mimicked. To do this properly is a very tall order. It would require both mimicking all the non-linear processes involved in generating the photographic images and all the data reduction and calibration steps that were applied in an attempt to model these processes. The difference between APM magnitudes and true total magnitudes arises from the failure of the calibration processes to work precisely or not to work at all, as happens in certain regimes (eg on saturated images). In practice, what one needs to know when working with APM/2dFGRS magnitudes is the distribution of magnitude errors and to what extent they correlate with other parameters such as magnitude, surface brightness, etc. This is shown in some detail by the empirical comparisons made in Sections 2 and 3.

To avoid duplication some related text and the original top panels of Fig 3 have been removed from Section 3.

2) The authors are rather glib about the importance of surface brightness effects. They quote Cross et al, saying that the surface brightness effects are only important for the LF at luminosities less than -16. In addition, by comparing Folkes et al to Cross et al, you conclude that the effect of surface brightness selection on the luminosity density is only 10%. However, in the internal comparison in Cross et al (much more robust than comparing the different samples of Folkes and Cross), they found that correcting properly for surface brightness is a 30-40% effect in the luminosity density. In addition, the authors don't look for any systematic trends with surface brightness, or evaluate their completeness as a function of surface brightness. These trends will be MUCH stronger with surface brightness than they will be with apparent magnitude, and furthermore surface brightness is better correlated with luminosity than is apparent magnitude. The paper should at least show the magnitude residuals and incompleteness between 2dF and the CCD surveys as a function of surface brightness (as measured in the CCD surveys).

At the time of writing the original draft of this paper a much more thorough investigation of the effects of surface brightness was being undertaken by 2dFGRS team members Nick Cross and Simon Driver. This work is now essentially complete, but thesis writing and changing jobs and continents has delayed the production of this paper.

We have, therefore, expanded section 3 to include a more thorough investigation of the effect of surface brightness as measured by the SDSS on both magnitude errors and incompleteness. We also report on the new work of Cross, Driver etal which makes use of the still deeper MGC data and supercedes their earlier work.

A new panel in Fig. 3 shows how the magnitude residuals correlate with surface brightness and this is discussed in the text. We conclude that while a significant fraction of the 2dFGRS magnitude measurement error correlates with surface brightness it does not dominate the overall error. Thus there are other shortcomings of the photographic data that are equally or slightly more important in determining the measurement error. We point out that when determining the overall galaxy luminosity function the main thing to know is the overall distribution of magnitude errors. One case where the correlation with surface brightness has to be taken into account is when the sample is split according to some other parameter that correlates with surface brightness. This was done in the Madgwick etal (2002) paper on the luminosity function of different spectral classes, but even here it was a very small effect.

We have also included a new panel in Fig. 3 which demonstrates there is little or no correlation of the magnitude residual with SDSS absolute magnitude.

A new panel in Fig 4 shows how the completeness varies with surface brightness. This demonstrates that over most of the range of surface brightness populated by galaxies the completeness is uniformly high. Only at the extremes of the surface brightness distribution does it begin to drop. The overall incompleteness due to this variation is 2%, with another 7% being caused by processes unrelated to surface brightness. We note also that much of the incompleteness is at the high surface brightness end where objects can be lost due to mis-classification.

Our estimate of the completeness of the 2dFGRS parent catalogue, shown in Fig 4, has increased from 87% to 91%. The reason for this is that we now compare our completeness relative to the SDSS objects selected to be part of the main galaxy survey. Previously we were comparing to a catalogue of all extended SDSS objects. The later is slightly contaminated by faint stars and ghost images. We comment on this in Section 7 and Fig 10 where we compare with the Yasuda etal "galaxy" counts, which did not exclude this contamination. The net effect of this revision to our estimate of the 2dFGRS completeness is an approximate shift of 4% in the value of $\phi_\star$ *and quantities derived from it.*

The referee was right to point out the result of Cross et al. (2001) for the effect of surface brightness limits on the estimated luminosity density. In the Cross et al. (2001), the effect of the surface brightness limit was calculated by making the same inaccurate assumption as Blanton et al. (2001). Namely, Cross et al. treated the APM magnitudes as gaussian corrected isophotal magnitudes without taking account of the way they are calibrated to total CCD magnitudes. It therefore overestimates the correction. In Nicholas Cross' thesis (and the soon to be submitted Cross et al. 2002) a more empirical approach has been adopted using the deeper MGC data as a reference. This shows (like Fig 3 and 4) that only for objects of extreme surface brightness do the magnitude errors and completeness become poor. The net effect is that our LF counts all galaxies brighter than $M_{b_J} = -16.5$ and brighter than this magnitude agrees well with estimates from the deeper MGC data. The resulting systematic error in the luminosity density is of the order of 5-10%.

In the light of this new analysis we have modified several of our references to the Cross et al. (2001) paper and have rewritten the section in the conclusions that discusses our estimate of the luminosity density.

3) You do not specify whether you apply the reddening corrections provided by the SDSS EDR or your own when you plot Petrosian magnitudes. The reddening is significant compared to the differences you are finding, so please be specific about this. Note that the Petrosian magnitudes in the EDR are NOT corrected for reddening automatically (though the SDSS reddening values are given).

We thank the referee for pointing out this oversight. We in fact correct both 2dF and SDSS photometry using the Schlegel etal dust maps. This is now stated explicitly and appropriately referenced.

2 Other General Comments:

1) The authors make reference to the improvements in the luminosity function due to greater numbers of objects with redshifts and larger volumes covered. While this might have been true until the first several thousand redshifts were measured, since that time the determination of the luminosity function has been limited by imaging quality. For example, the big difference between the quality of the luminosity function in the blue of 2dFGRS and, say, the CfA2 survey is NOT the fact that 2dFGRS has 10 times the number of objects, but the fact that one has APM magnitudes rather than Zwicky magnitudes.

We agree that accurate photometry (or failing that accurate characterisation of the photometry errors) is of prime importance. Having said this the value of a wide angle, large volume, survey in determining $\phi_\star$ should not be underestimated. We haven't been able to identify a specific sentence or comment that could lead to the above objection. We note that we do make it clear that sheer numbers of galaxies is no longer what determines the accuracy of the luminosity function estimates. For example:

In section 8 we say: "The large number of galaxies used in our estimates makes this statistical error very small and so it is never the dominant contribution to the overall error."

In section 9 we write: "We conclude that, when normalized in the same way, there is excellent agreement between the SDSS and 2dFGRS luminosity functions and that the dominant remaining uncertainty in the present day b_J-band LF is due to residual uncertainties in evolutionary corrections."

In the conclusions we write: "With over 110 500 redshifts, the statistical errors in our estimate are negligible compared to the systematic errors ..."

2) The references should be broadened in general. For example, no mention at all is made of the LCRS luminosity function of Lin et al. 1996, which is quite an important work in the

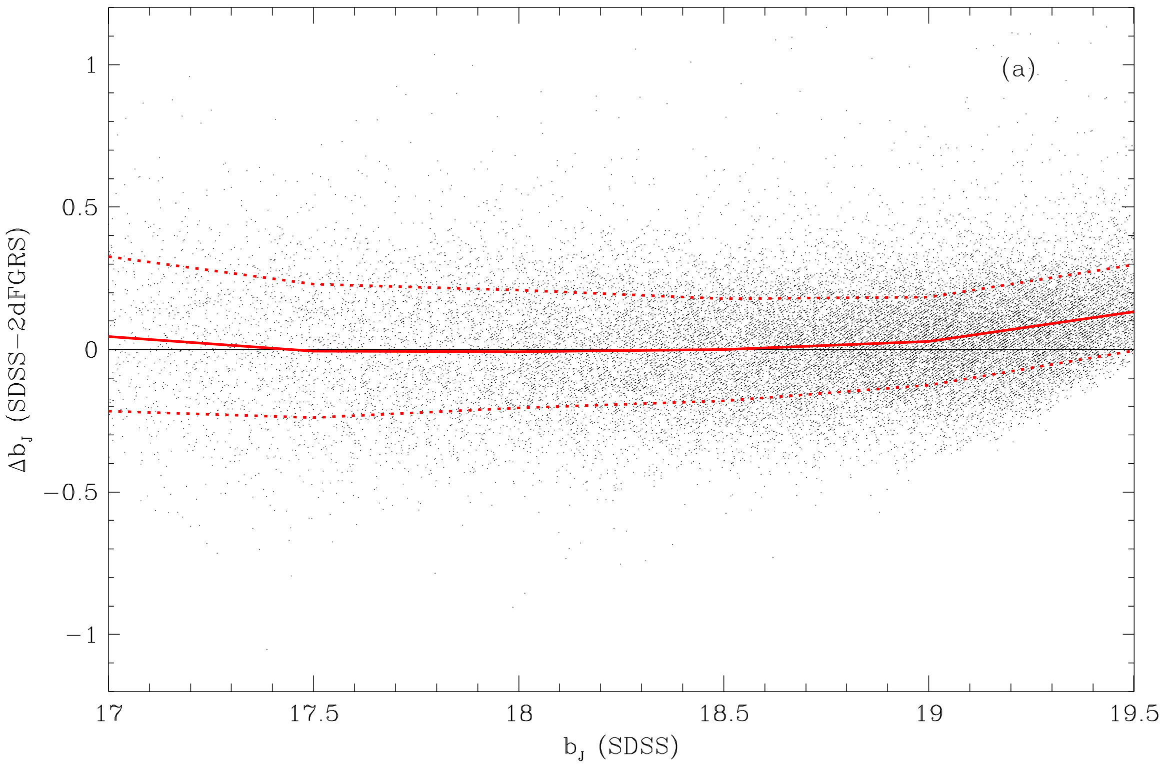

Figure 1: The 2dFGRS-SDSS magnitude residuals plotted against SDSS magnitude rather than 2dFGRS magnitude. The 2dFGRS magnitude limit causes a significant bias in the median residual at faint magnitudes.

literature. It does turn out to underestimate the luminosity density by a large factor, but I don't think this diminishes its importance. At least some of the hydrodynamical models of galaxy formation should be quoted. There are a number of places where the text must be fleshed out and referenced more thoroughly. These are discussed under "Textual Comments" below.

> *We have updated and expanded the references. We now reference the LCRS luminosity function of Lin et al., but since it is solely in the r-band we cannot compare directly and therefore do not have an extensive discussion.*

3) In general the authors need to remember that when one plots trends to check for systematic effects, it is best to regress along an axis with the highest possible signal to noise. For example, in Figure 2, when you plot the differences with respect to EIS or SDSS, the x-axis should be the EIS or SDSS magnitudes, which are far more accurate magnitudes. Otherwise the noise in the 2dF magnitudes can wash out trends.

> *We understand the referee's point and generally have done precisely what is suggested. However, in the plots showing residual versus magnitude one also has to consider selection effects and any biases they introduce. As the CCD surveys (EIS and SDSS) go much deeper than the 2dFGRS parent catalogue then at each 2dF magnitude we essentially have an unbiased sample. (The CCD magnitude limit does produce an upper envelope to the distribution of points, but has negligible effect on the median and plotted percentiles.) However, if we make the equivalent plot with CCD magnitude on the x-axis [see Fig 1 above] then the 2dFGRS magnitude limit causes a strong bias at faint magnitudes.*

> *The rms error of 0.164 magnitudes on the 2dFGRS magnitudes would not hide a significant non-linearity, if one existed over the 2.5 magnitude range shown in these figures.*

3 Textual Comments:

Abstract: ——–

1) correct to "$\rho_L = (1.9 \pm 0.18) \times 10^8$ ..." (page 1)

> *Fixed.*

Section 1:

Section 2:

1) How were the extinction corrections determined? (page 2)

> *We use the dust maps of Schlegel et al. (1998). This is now stated explicitly and appropriately referenced in the extended Section 2.*

2) The explanation of the calibration needs to be MUCH clearer. It should be clear in the first paragraph that *two* calibrations relative to CCD photometry are performs: first with the "external CCD photometry" and second with the ESO data in the SGP. It should also be clear that the $b_J = B - 0.28(B - V)$ relation is enforced by the CCD calibration. (page 3)

> *We have made major changes to section 2 such that all stages of the calibration are discussed in much more detail. We are not sure what the referee means by "the $b_J = B - 0.28(B - V)$ relation is enforced by the CCD calibration". However, we do explicitly state that the ESO data give empirical support for this colour equation.*

3) Quote the χ^2 difference between the color terms 0.152 and 0.088. I am sure it is large, but the plot is visually unconvincing (especially since the relation is clearly nonlinear), so it would forestall criticism to make this point quantitative. (page 3)

> *A least squares fit gives a slope of* 0.152 ± 0.004 *for the relation between* $\mathrm{b_J - g}$ *and* $\mathrm{g - r}$. *A statement to this effect has been inserted near the end of the 2nd paragraph of Section 3.*

4) What is the source of the estimate 0.04 of the accuracy of the 2dFGRS overall zeropoint? I find it optimistic. (page 3)

> *This zeropoint is essentially set using the EIS. This is now discussed fully in the new Section 2.*

5) I don't think Blanton et al. showed the plot corresponding to Figure 2e for the "APM" magnitudes, only for the isophotal magnitudes. (page 3)

> *We thank the referee for pointing this out and have revised the text.*
>
> *The plots in Blanton et al. show the difference between total (Petrosian) and isophotal magnitudes for various isophotal limits. Blanton et al. do consider a small gaussian correction to the isophotal magnitudes and call these corrected magnitudes "APM like", though since they are not calibrated in the same way as APM magnitudes this label is misleading. Blanton et al. find that a luminosity function synthesised from the SDSS isophotal magnitudes matches quite well the APM LF. This is a coincidence (because they used the wrong colour equation and normalized in a different way) along with the fact that their gaussian corrections are small lead them to conclude that APM magnitudes miss as much of the light as their isophotal magnitudes. But this is wrong.*

6) In Section 2.2, why do you compare to the SDSS spectroscopic sample at all, which has its own incompleteness to worry about? Just compare to the target list in the photometric catalog. It is the only important comparison to make, and the only "completeness" worth talking about. (page 6).

> *This has been modified as discussed above and has led to our estimated incompleteness falling by 4% . We were unable to extract the target list in the SDSS photometric catalog for many months due to problems with access through a firewall! It took considerable correspondence with SDDS technical people before they*

fully determined what ports sdssQT utilised before successful access was achieved through our firewall. In the meantime the best we could do was to use the redshifts as an indicator of inclusion in the spectroscopic sample. While we now use the full target list as our primary estimate of incompleteness the comparison to the sample with redshifts is still of interest and so we have chosen to retain it.

7) In Section 2.3, you compare to the SDSS and conclude that 1% of the objects are hugely discrepant. I'll just note that the estimates of the SDSS galaxy redshift error rate are roughly at this level, so this is really an upper limit on your own error rate.

We've added a a comment that some of the discrepancies could be due to the error rate in SDSS, but we are unaware of a SDSS paper that quantifies this rate.

Section 3:

1) Properly speaking, the spectrum is redshifted, not the bandpass. In the restframe of the galaxy, the bandpass is actually blueshifted. (page 7)

Thanks. We've corrected this.

Section 4:

1) It is great that you've made these mock catalogs. One detail for the future is, instead of using the errors estimated as a function of apparent magnitude, use them as a function of absolute magnitude, since this is the thing you are really trying to measure AND it is more closely related to the likely systematics (since it is more closely related to surface brightness than is apparent magnitude). I doubt this is an important enough issue to justify holding up this paper, though.

The raw mock catalogues which will hopefully be described in a paper in the not too distant future have no magnitude errors. Thus the user of the mocks is free to add errors in whatever way they feel appropriate. This makes the catalogues more flexible. In this paper we added errors that were independent of apparent magnitude and so also independent of absolute magnitude. This seems justified to a good approximation by the scatter plots shown in Fig. 3.

Section 5:

1) It is likely that the differences in Figure 8a are just due to the missing 5% of objects in the "k+e by type" sample. The low signal- to-noise objects are probably low surface brightness, and thus low luminosity. (page 10)

The referee is absolutely correct. We have made new LF estimates using the mean k+e correction but only the "k+e by type" sample. This estimate agrees very accurately with the estimate from the same sample but using the type dependent k+e corrections. Consequently we have revised the Fig 9a (old 8a) to compare to the two different k+e corrections for identical samples. The fact they these samples produce a LF that misses objects at the faint end is discussed in the modified text.

Section 6:

1) Hmmm, perhaps a clarification is in order as to why the 2dFGRS star-galaxy separator does poorly at bright magnitudes? This seems counterintuitive — is saturation the problem? (page 13)

There seems to have been some confusion here. In the original draft the only mention of star-galaxy separation on this page was in regard to the star-galaxy separation we used on the SDSS EDR data. The criterion we used was $\mathrm{r_{PSF}} - \mathrm{r_{model}} > 0.242$ *and this does break down at bright magnitudes. Presumably diffraction spikes, bleeding and/or scattered light cause even the images of stars to become extended when very bright.*

Now that we have been able to extract all the flags used by SDDS to reject stars and classify galaxies for the main SDSS galaxy sample this section has been modified in several ways. We now show and discuss both the "extended" source counts and those of SDSS classified galaxies.

Section 7:

1) State what absolute magnitude range you assumed when you calculated the number counts. I know it matters very very little, but you should at least say it. (page 14)

> *We use the absolute magnitude range* $-13 > M_{\mathrm{b_J}} - 5\log_{10} h > -24$*, which is more than adequate for accurate convergence of the integral that defines the counts. These limits are now stated in Section 8 (old Section 7).*

Section 8:

1) There is some mix-up near the end of the second paragraph. You imply that the area of overlap is 173 deg^2; this is the area used by Yasuda et al. It is the area of overlap (about 60 deg^2 or so, if indeed you used the 2dF public dataset) which is 5% overdense, right? Perhaps I have misunderstood, but in any case you need to clarify the issue. (page 16)

> *Our numbers are correct. The northern part of the EDR data lies entirely within 2dFGRS northern strip. Thus the overlap area is almost identical to that used by Yasuda etal in the north. The only difference is due to holes in the 2dF coverage cut out around satellite trials etc. Perhaps the referee was thinking of the overlap with the portion of the 2dFGRS survey whose redshifts have been measured. However this isn't relevant as for the counts we don't need the redshift and can use the full area of the parent catalogue (all of which was released in the 100k public data).*

Section 9:

1) What range of absolute magnitudes do you use in defining the selection function? (page 18)

> *There are no explicit limits in the code used to calculate the selection function. The integral over the selection function is done analytically.*

2) You find large differences in the selection functions in the NGP and the SGP, but note that these can be explained with small tweaks to the calibration. You should note the difference in the luminosity density for the NGP and SGP separately. This would be as good an error estimate as any other. (page 19)

> *The luminosity densities estimated individually from the NGP and SGP data are* 2.04×10^8 *and* 1.64×10^8 h $L_\odot Mpc^{-3}$*, each differing by close to* 1σ *from the combined estimate of* 1.82×10^8 h $L_\odot Mpc^{-3}$*. These values are now quoted in penultimate paragraph of Section 8 (old Section 7) where we first estimate* ρ_L*.*

Section 10:

1) Your comparison to the 2MASS LF completely ignores the work of Kochanek et al, which, since it agrees with Cole's result, suggests that LSS is not to blame for the discrepancy (I think).

> *We have inserted a reference to Kochanek et al. (2001). Their result does loosely support that of Cole et al. , but it should be noted that their uncertainty is dominated by the uncorrected isophotal magnitudes that they adopt. We now mention this and include a reference, Andreon (2001), to more recent work on the accuracy of 2MASS photometry.*

4 Additional Changes:

Here we describe some additional modifications we have made to the manuscript to address issues raised by others.

1) Ivan Baldry has been added to the list of authors.

2) The magnitude range in the abstract over which the LF is claimed to be well fit by a Schechter function has been reduced to $-16.5 > M - 5\log_{10} h >$ *-22 . A Schechter function fits our results to fainter magnitudes but fainter than* -16.5 *we know that the surface brightness limits of our survey begin to play role.*

3) A comment has been added in section 9 regarding the Zucca et al. explanation of the difference between their and Loveday's LF estimates.

4) We have extended the acknowledgements

References

[1] Blanton M.R., et al. 2001, AJ, 121, 2358 (astro-ph/0012085)

[2] Cole S., et al. (The 2dFGRS Team) 2001, MNRAS, 326, 255

[3] Cross N.J.G., et al. (The 2dFGRS Team) 2001a, MNRAS, 324, 825

[4] Kochanek C.S., Pahre M.A., Falco E.E., Huchra J.P., Mader J., Jarrett T.H., Chester T., Cutri R., Schneider S.E., 2001, ApJ, 560, 566

[5] Maddox S.J., Sutherland W.J., Efstathiou G., Loveday J., Peterson B.A., 1990b, MNRAS, 247, 1

[6] Maddox S.J., Efstathiou G., Sutherland W.J., Loveday J., 1990b, MNRAS 243, 692 (APMI)

[7] Maddox S.J., et al. , (The 2dFGRS Team) 2002, in preparation.

[8] Schlegel D.J., Finkbeiner D.P., Davis M., 1998, ApJ, 500, 525

[9] Zucca E., et al. 1997, A&A, 326, 477