1 Title: ULTRA-VISTA: Observing Beyond Re-ionization

PI: M. Franx (Leiden Observatory)

Co-I's: H.-W. Rix, I. Smail, E. Bell, R. Somerville, B. Leibundgut, P. v.d. Werf, H. Rottgering, I. Labbe, J. Dunlop, A. Cimatti, P. van Dokkum, A. Omont, J. Emerson, M. Irwin

1.1 Abstract: (10 lines max)

For many years to come, VISTA will be peerless in space and on the ground in its efficiency of ultra-deep near-IR imaging of cosmologically-large areas. On this basis, we propose the ULTRA-VISTA survey to open up the possibly last redshift frontier in galaxy formation, $z = 6.5 - 10$. This requires multi-band imaging totalling 1760 hours (200 nights) at a *single pointing*, best within the COSMOS field, with t_{exp} driven by the requirement to sample the rest-frame UV luminosity function to L_* at these epochs (extrapolating the $z = 6$ luminosity function to higher redshift). ULTRA-VISTA will yield a thousand $z > 6.5$ galaxies, and map the evolution of the sources that caused reionization; it will do so well before JWST. Combined with results from other telescopes (ALMA, IRAC), the masses and star formation rates of these galaxies can also be estimated. At the same time, ULTRA-VISTA will provide a definitive assessment of the role of red and/or passive galaxies at $2 < z < 3.5$, and hence define the onset of downsizing in galaxy formation.

2 Description of the survey: (Text: 3 pages, Figures: 2 pages)

2.1 Scientific rationale:

Probing the Onset of Galaxy Formation: Current Limitations

Progress in studies of the high redshift universe has been enormous in the past decade. Whereas in the early nineties the term high redshift galaxies meant galaxies at redshifts of about 0.5, we now study galaxies at $z = 6$ routinely, and have candidate high redshift galaxies up to z=10. The revolution has been made possible by both large groundbased telescopes, and very deep imaging by the Hubble Space Telescope. The Hubble Space Telescope gave us the beautiful images like the Hubble Deep Fields, and the Hubble Ultra Deep Field, which allowed astronomers to reliably select high redshift galaxies on the basis of their colors. The new 10-m class telescopes on the ground enabled astronomers to confirm the redshifts of galaxies selected in this way.

Using these facilities, we can currently reach to about a redshift of 6.5. Above that redshift, progress becomes much harder: the break in the spectra have moved to $\lambda_{\rm rest} = 1216$Å, due to the Gunn-Peterson absorption of the partially neutral medium at high redshift. Hence in order to detect the continuum of the high redshift galaxies, one has to observe at 1 μm or above, where existing observational capabilities have been dramatically less effective. For example, the perhaps best ultra-deep near-IR imager, HST's NICMOS, has a field-of-view of 50x50 arcsec (NIC3), undersamples the PSF significantly and is out of focus; ISAAC on the VLT has 2.5x2.5 arcmin, but faces much higher backgrounds. *With the status-quo instrumentation, comprehensive studies of the galaxy population ($z > 6$) that caused re-ionization and ended the "dark ages" are simply not feasible.* This is about to change with VISTA.

VISTA: Revolutionizing Ultra-Deep Cosmological Near-IR imaging

To understand these earliest phases of galaxy evolution ($z \sim 6 - 10$) one needs to devise an imaging survey that not only reaches flux limits that correspond to characteristic galaxies (L_*) at those epochs (e.g $J_{AB} \approx 27$), but does so over a volume large enough to cover all relevant galaxy/density environments. Both cosmological simulations and observations at redshifts 1-5 suggest that the role of environment was tremendously important at the earliest epochs: galaxies used to be much more biased (e.g. Springel & Hernquist 2003). It turns out that VISTA will be unmatched in this regime of ultra-deep imaging over $\sim 100Mpc$ co-moving size fields until the advent of JWST.

In any single pointing, Vista will cover a field-of-view of 2000 sq arcmin, distributed over 16 adjacent fields. This is 300 times faster than ISAAC on the VLT, and still 40 times faster than the VLTs next generation IR-imager, HAWK-I. But this 1-2 orders of magnitude gain will hold until 2013-2014 even when compared to space: if the new Wide-Field Camera 3 will be installed on HST, it can image to J=27 in 7 hours over a field-of-view of 5 sq arcmin. VISTA will image 2000 sq arcmin in the J-band to that depth 9 times faster; just the J-band image of the survey proposed here would require > 1000 orbits of imaging with WFC3 on HST.

Of course this survey speed advantage will only hold, if really a field of $\sim 1^\circ$ on a side is needed, which it is. A great deal of recent theoretical work, and empirical verification, have shown that at redshifts $z \sim 1$ a field size of $0.5^\circ - 1^\circ$ is needed to sample the relevant density environments fairly (e.g. Somerville et al 2004); this has lead to survey designs such as COMBO-17/GEMS (Wolf et al 2004; Rix et al 2004), COSMOS (Somerville et al; in prep.) or VVDS (Le Fevre et al 2004), all with such field sizes. In a LCDM cosmology the co-moving volume per unit redshift and solid angle at $z = 8$ is within 25 % the same as at $z = 1$: hence, to sample the early phases of the full range of environments we see at lower redshifts, we need to survey a similar solid angle: ~ 1 square degree, or one VISTA pointing.

While 16 disjoint fields, adding up to 1 square degree, may lead somewhat faster to the cosmic average, the near-contiguous field of VISTA provides an advantage: an improved ability to estimate the (projected) correlation function of sources within broad redshift bins. This may be the most immediate link to connect any detected galaxy population to the halos they live in (e.g. van den Bosch et al 2003). Existing work with small samples and poor redshift estimates (e.g. Daddi et al 2003 for the FIRES data) has demonstrated that this is a very powerful tool.

For a given total amount of observing time, e.g. 1500 hours, any distribution of the time over several fields or pointings, with a corresponding reduction in exposure time per "pixel" would threaten our goal of reaching L_*.

TAKEN TOGETHER, THESE ARGUMENTS DEMONSTRATE THAT VISTA WILL BE A PEERLESSLY EFFICIENT LARGE AREA IMAGER FOR VERY HIGH REDSHIFT STUDIES

As a consequence, we can finally access the $z > 6.5$ regime with VISTA, by doing multi-color imaging in the Near-IR bands. This will allow us for the first time to do a large area high redshift survey, sensitive enough to find galaxies of characteristic luminosity during and even before re-ionization, and before JWST.

2.2 Immediate objective:

Here we propose to do an ultra-deep imaging survey with VISTA, surpassing all previous and ongoing surveys in depth and area. We propose to spend 1600 hours of integration time on 1 single pointing (with small dithers), in 5 filters (Z,Y,J,H,K), with comparable integration times in each band. This will produce 16 ultra-deep tiles of 11x11 arcmin. The depths that can be reached are 27.4, 27. 26.8, 26.3, 25.8, in Z,Y, J, H and K respectively (in AB magnitudes, as used through the full text). The data is supplemented with deep I-band imaging (existing for our proposed field at the COSMOS field). Many different science applications can be done with this survey:

Sources of reionization: the high redshift universe ($z > 6.5$).

The most secure way to find high-redshift galaxies from photometry is by the "dropout" technique, which relies on the 1216 Å break present in $z > 6.5$ galaxies. Galaxies can be selected as I, Z, Y and J-dropouts. The selection criteria are shown in Fig. 1. The redshift intervals are 6-6.5, 6.5-7.5, 7.5-9.5, 9.5-11 for I, Z, Y, and J-dropouts respectively. Using the luminosity function derived for $z = 6$ I-dropout galaxies by Bouwens et al (2006), we find that we expect 1200, 740, 242, and 10 galaxies as I, Z, Y, and J-dropouts. This demonstrates that we can effectively access galaxies to $z = 9.5$, obtaining a sample of several hundred galaxies in the bin to $7.5 < z < 9.5$. Obviously the number of $z = 10$ and above galaxies is too low to be very meaningful, unless the luminosity density increases compared to $z = 6$.

The 8-sigma magnitude limits correspond to 0.5, 0.8, 1.2 L_* at $z = 6.3$, 7 and 8.5, assuming the $z = 6$ Luminosity Function from B06. Hence the Luminosity Function is well sampled, and can be accurately determined at these redshifts. Furthermore, the area is large enough to give a result not influenced strongly by large-scale structure. The total field has a size of 195x156 comoving Mpc at $z = 8$, much larger than the scale of fluctuations. Hence the determination of the luminosity density will be mostly free of the uncertainties of large-scale structure. Angular correlation functions can also be constructed.

Once the time is awarded, the field will become the standard Near-IR deep field. We (and others) will propose for a variety of follow-up with other instruments to address the following questions: i) VLT spectroscopy with KMOS and other spectrographs can detect Lyα from the high redshift galaxies. Obviously, continuum spectroscopy is beyond the capabilities of the VLT, but it is possible to do Ly-α spectroscopy of the galaxies with equivalent widths of more than 20 Åin the rest frame. Approximately 30% of the Lyman break galaxies have such strong emission out to $z = 6.5$ (Dow-Hygelund et al 2006, submitted). This requires observions at 1 μ and beyond, which will become possible with KMOS on the VLT, and Lucifer on the LBT. Long integration times up to 100 hours will be needed to determine the redshifts.

ii) Ultra- Deep IRAC imaging will be needed to image the galaxies in the rest-frame optical, in order to derive more reliable masses. At $z = 6.5$, the Balmer-break/4000Å-break has moved beyond 3μm, and only IRAC will be able to access this regime. Previous work by Eyles et al. (2005) and Yan et al. (2005) has detected several $z = 6.5$ galaxies with IRAC, leading these authors to conclude that the galaxies are massive ($> 10^{10} M_\odot$) and old (several hundred million years), implying they started to form at very high redshift. The combined deep VISTA + IRAC imaging will be able to push this back to $z = 9.5$, when the universe was only 500 Myear old.

iii) ALMA imaging will produce very accurate sub-mm fluxes. Star formation rates of 10 $M_\odot$/year can be detected in the continuum to $z = 8$ with integration times of 30 minutes. This will give star formation rates completely independent of the UV determined star formation rate, and will answer the question whether these galaxies had build up a reservoir of dust.

Galaxies at z=2-4.5

The deepest near-IR survey until now is the FIRE Survey (PI Franx). The FIRE Survey has uncovered a large population of red galaxies at $z = 2 - 3.5$. These galaxies are called "DRGs" for Distant Red Galaxies, and are selected by the simple color criterion $J - K > 2.3$, equivalent to $U - V > 0.3$ (Franx et al 2003). They occur in large numbers. Using all available fields with deep Near-IR imaging, van Dokkum et al. (2006) found recently that 69 % of the massive galaxies at $z = 2 - 3$ were DRGs, most of which are not selected by the usual Lyman break criterion.

One of the main uncertainties at this moment is how prominent DRGs are at lower masses, and at higher redshifts. The mass limit of the study by van Dokkum et al. was 10^{11} $M_\odot$, driven by the shallowest survey included in the study. The redshift limit was similarly driven by the depth of the survey. The survey proposed here goes more than 2.5 magnitudes deeper, and hence will sample the mass function to below 10^{10} $M_\odot$ at $z = 2 - 3$, well below the typical mass for Lyman break galaxies. Hence it will finally be able to give the relative contribution of "red" galaxies to the total mass budget. Similarly, combined with deep Alma imaging it will give the relative contribution of red galaxieds to the overall star formation rate.

The area of the survey is finally large enough to determine these numbers with sufficient accuracy. Currently, only very small areas have been imaged to this depth. The comoving size of 100x150 Mpc produces a sufficiently large sample of the universe. In addition, we expect to find 6000 DRGs or more, extrapolating from the FIRE Survey, demonstrating that studies of the correlation function should be possible.

Furthermore, we can push this study to redshifts significantly higher. Galaxies with masses of 10^{11} $M_\odot$ will be detected out to $z = 5$. Again, the large area will produce a proper sample of the universe, and the build-up of massive galaxies can be traced over a large redshift range.

In addition that data contain detections of several hundred distant supernovae, using the predictions from

Dahlen and Fransson (1999) for our yearly limits of J=26.2, K=25. These will be distributed from z=0 to z=2. While real-time triggers may not be efficient, II's and Ia's can be analyzed (position, photometric redshifts and light-curve shape) in hindsight, with a global fit to the data. This valuable science is beyond the immediate survey goals, the coadded image, but should be borne in mind.

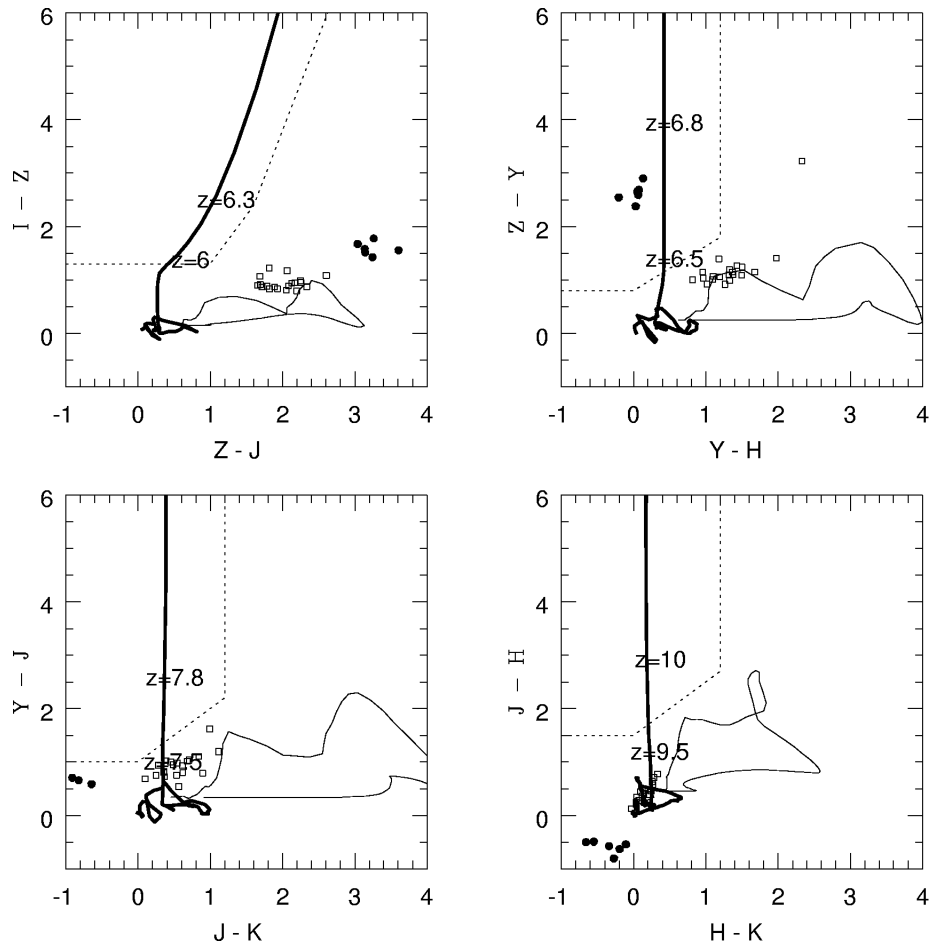

Figure 1: Color selection criteria for I, Z, Y, and J-dropouts, respectively, in AB magnitudes. The thick curves show the expected colors for a starburst galaxy. The thin curves show the expected color for early-type galaxies at lower redshift. The selection boundaries (dashed) are chosen to minimize contamination by low redshift ellipticals (where the 4000Å break is mistaken for the 1216Å break), and stars (points). The stars plotted here are brown dwarfs, the main contaminants of concern. Note that T-dwarfs produce some contamination for Z-dropouts, but can be identified by their blue $J-K$ colors. The redshift labels refer to the color-space location of high-redshift starbursts on the (thick) tracks.

References

van den Bosch et al. 2003, MNRAS, 340, 771
Bouwens et al. 2006, ApJ, accepted, astroph 0509641 (B06)
Daddi et al. 2003, ApJ, 588, 50
Dahlen & Fransson, 1999, AA, 350, 349
van Dokkum et al. 2006, ApJ 638, L59
Eyles et al. 2005, MNRAS, 364, 443
Forster Schreiber et al. 2004, ApJ, 616, 40
Franx et al. 2003, ApJ, 587, L79
Labbe et al. 2003, AJ, 125, 1107
Labbe et al. 2005, ApJ, 624, L81
Somerville et al. 2004, ApJ, 600, L171
Springel & Hernquist, 2003, MNRAS, 339, 312
Yan et al., 2005, ApJ, 634, 109

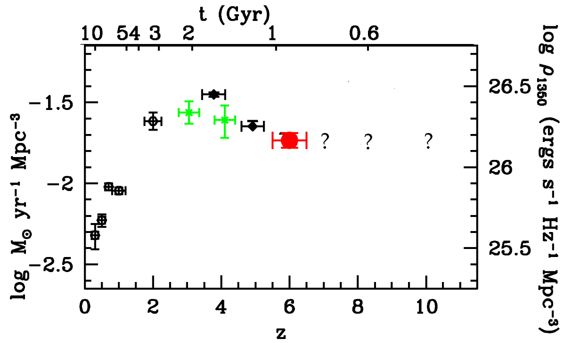

Figure 2: Integrated luminosity density, and star formation rate density derived to $z = 6$ (adapted from Bouwens et al 2006). This survey will provide an accurate determination of the UV-luminosity density to 1 L_* at z=9.5, possibly extending beyond z=10 depending on the evolution of the Luminosity Function.

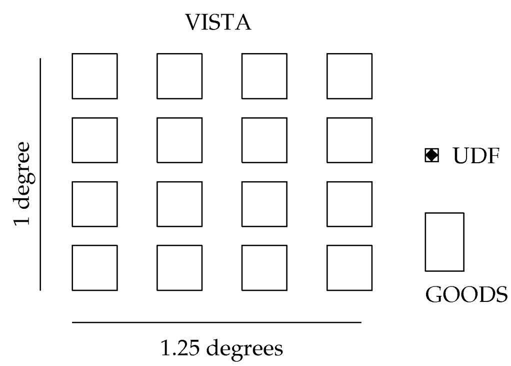

Figure 3: The field size of VISTA compared to deep fields at this moment. The UDF is the only deep field with Z-K band imaging of comparable depth as the survey proposed here. The area covered by the Near-IR imaging indicated by the black area and is 300 times smaller than the proposed ULTRA-VISTA size. The GOODS survey has Near-IR imaging available over 1 field of nearly 150 sq arcmin, but shallower by 2 magnitudes or more than the survey proposed here. The difference in size is a factor of 13. The field size of the VISTA pointing corresponds to a comoving size of 195 by 156 Mpc at $z = 8$.

3 Are there ongoing or planned similar surveys? How will the proposed survey differ from those? (1 page max)

We are not aware of any other survey planning to do this work. The UKIDSS-UDS possibly comes closest, but note that it is significantly shallower than this survey, and lacks the depth in the Z and Y bands. The planned magnitude limits in J and H for the UDS are 25.9 and 25.37, 1 magnitude shallower than the survey proposed here. **The difference is critical**. Even if the UDS has Y-band imaging to similar depth as its J-band depth, it would find only 20 Z-dropouts and 10 Y-dropouts.

HENCE IT IS CRITICAL TO GO TO THE DEPTH PROPOSED HERE. Other surveys simply will not go deep enough to allow this kind of work.

4 Observing strategy: (1 page max)

We propose to observe only 1 pointing. We will dither in a box of 30 arcsec to allow a proper correction for bad pixels and structure in the dark. We do not propose to construct a "filled" image as this would cost the depth needed to achieve our science aim for t_{exp}(total)$<$ 2000 hours: it would reduce the integration time per pixel by about a factor of 3.

We note that one of the detectors suffers from small patches of bad pixels. The total relative area of these patches is so small (<< 1% of total field size) that we do not adapt our strategy, as the cost in terms of depth would be significant if we were to make much larger steps, e.g., by making large steps in the y-direction to make a filled column.

We will also make half-pixel dithers and contruct an output image which is resampled at half the scale of the VISTA imager.

We note that the total length of the survey is 200 nights, comparable to that of UDS, which runs on a telescope with both dark-time and bright-time instrumentation. Hence relatively speaking, our request is more modest by a factor of 2.

The field proposed here is the COSMOS field. It is optimal for a variety of reasons. The most important advantage is the availability of deep HST-ACS I-band imaging, which goes to a magnitude limit of 27.7 (5 sigma). We can use this imaging directly to select I-dropouts at $z = 6$ Even though this redshift regime has been sampled well already, this overlap is very important, as it will enable us to compare the derived luminosity function at $z = 6$ with previous work, and hence enable us to provide a baseline for the luminosity functions derived at higher redshifts.

Other advantages include the wealth of data relevant for lower redshifts, including very deep optical imaging, and spectroscopic redshifts. This is extremely valuable for studies of the $z = 2 - 5$ redshift regime.

Obviously, other fields could be considered too, if practical reasons make it very difficult to schedule this deep exposure. The second best field would the UKIDSS UDS field (ra=2 hours). It will be the main "other" large field (3 times larger than the GEMS - COMBO-17 ECDFS), and hence observations done in the optical for UDS will be usable for this survey (although covering only half the VISTA area). Optical imaging will need to be obtained for the other half, but we have no doubt that it will be possible to get this.

5 Estimated observing time:

Period	Time (h)	Mean RA	Moon	Seeing	Transparency
P79	66	10h	Z, Y: dark, J, H K: all conditions	< 0.8	clear
P80	286	10h	Z, Y: dark, J, H K: all conditions	< 0.8	clear
P81	66	10h	Z, Y: dark, J, H K: all conditions	< 0.8	clear
P82	286	10h	Z, Y: dark, J, H K: all conditions	< 0.8	clear
P83	66	10h	Z, Y: dark, J, H K: all conditions	< 0.8	clear
P84	286	10h	Z, Y: dark, J, H K: all conditions	< 0.8	clear
P85	66	10h	Z, Y: dark, J, H K: all conditions	< 0.8	clear
P86	286	10h	Z, Y: dark, J, H K: all conditions	< 0.8	clear
P87	66	10h	Z, Y: dark, J, H K: all conditions	< 0.8	clear
P89	286	10h	Z, Y: dark, J, H K: all conditions	< 0.8	clear

We request dark time (moon down or illuminated < 0.2) for the Z-band and Y-band. It is documented that the background is lower in Z during dark time, it is not documented on the ESO website concerning Y, and hence we are conservative and request dark time for Y. No constraints are necessary for J, H, K, hence our survey is designed to make equal use of all moon conditions.

We request all seeing conditions better than 0.8 as measured on the detector. We estimate that this corresponds to the best 75% of the time. This is based on the distribution of DIMM seeing for 2005, and the conversion of DIMM seeing to near-IR (true) seeing, to seeing as measured on the detector. The seeing averaged over all seeing conditions better than 0.8 is 0.67 arcsec. We used this value to calculate our magnitude limits.

The ratio of time over even and odd periods is adapted to produce “equal pressure” at any time the field is visible. The survey will run for 5 years, hence from 2007 - 2011.

5.1 Time justification: (1 page max)

The aim of this proposal is to go as deep as possible on 1 field. Our current integration time allows us to make a > 1 magnitude improvement in previous/ongoing studies, and allows us to sample the luminosity function to roughly L_* at $z = 7 - 9$.

We request an exposure time of 320 hours per band. This leads to a 5-sigma depth of 27.4, 27, 26.8, 26.3, 25.8 in Z, Y, J, H, and K respectively (AB magnitudes). As desirable for dropout work, the limits become progressive deeper with bluer band. This is preferred as it helps to select the “breaks” in the blue, required for reliable high redshift selection.

We estimate the overhead for the observations at 10%. This is an average of the overheads produced by readouts, jitter moves, and guide lock. The overhead due to telescope preset and filter moves is negligeable as we observe only 1 field, and take hour-long exposures per filter. Hence the total time required is 5 x 320 x 1.1 = 1760 hours (approximately 200 nights).

We note that the uniqueness of this project is its depth. A shorter integration time will substantially affect the science:

1) the number of detected galaxies at $z > 6.5$ goes down quickly, and may well approximate a very low number if the luminosity function evolves between $z = 6$ and $z = 9$. For example, the number of expected galaxies goes down by a factor of 3-4 if the depth is 0.5 mag shallower, leading to 60 $7.5 < z < 9.5$ galaxies. If the luminosity function were to evolve at the same rate as between $z = 3$ and $z = 6$, this number would go down by another factor of $2 - 3$. In short, in order to obtain meaningful samples or a meaningful test of the evolution of the luminosity function, we need the depth requested here.

2) the depth will become close to that of the UDS, which is 1 mag shallower. Hence the attractiveness of the survey will suffer.

6 Data management plan: (3 pages max)

6.1 Team members:

Name	Function	Affiliation	Country
CASU (VDFS) team	Pipeline processing	University of Cambridge	UK
CASU (VDFS) team	Data Quality Control-I	University of Cambridge	UK
J. Emerson	VDFS Coordinator	Queen Mary University of London	UK
WFAU (VDFS) team	Science Archive	University of Edinburgh	UK
WFAU (VDFS) team	Data Quality Control-II	University of Edinburgh	UK
N. Walton	VO Standards	University of Cambridge	UK
	This ULTRA-VISTA Survey specific tasks		
ULTRA-team	Data Quality Control-III	Leiden/Heidelberg	NL/D

The CASU (VDFS) team consists of Irwin, Lewis, Hodgkin, Evans, Bunclark, Gonzales-Solares, Riello. The WFAU (VDFS) team consists of Hambly, Bryant, Collins, Cross, Read, Sutorius, Williams. The ULTRA-team consists of Franx, Rix, Bell, and 2 postdocs (one based in Heidelberg and one based in Leiden).

6.2 Detailed responsibilities of the team:

We will use the VISTA Data Flow System (VDFS; Emerson et al. 2004, Irwin et al. 2004, Hambly et al. 2004) for all aspects of data management, including: pipeline processing and management; delivery of agreed data products to the ESO Science Archive; production of a purpose-built IVOA compliant science archive with advanced datamining services; enhanced data products including federation of VISTA survey products with SDSS survey products. Standardised agreed data products produced by VDFS will be delivered to ESO, with copies remaining at the point of origin.

The VDFS is a collaboration between the UK Wide Field Astronomy Units at Edinburgh (WFAU) and Cambridge (CASU) coordinated by the VISTA PI (QMUL) and funded for VISTA by PPARC. The VDFS is a working systems-engineered system that is already being successfully employed for the UKIRT WFCAM surveys as a test bed for the VISTA infrared surveys, and which is sufficiently flexible as to be applicable to any imaging survey project requiring an end-to-end (instrument to end-user) data management system. We emphasise the track record over the last decade of both the Cambridge and Edinburgh survey units in processing and delivering large-scale imaging datasets to the community as exemplified by the WFCAM Early Data release (EDR, (*http://surveys.roe.ac.uk/wsa/dboverview.html*) Lawrence et al 2006, Dye et al 2006).

6.3 Data reduction plan:

The data reduction will be using the VDFS, operated by the VDFS team, and augmented by individuals from Leiden and Heidelberg, especially for product definition and product Quality Control. We divide the plan into two distinct but intimately related parts: pipeline processing and science archiving. Much greater detail can be found in the SPIE papers cited previously.

6.3.1 Pipeline processing

The Cambridge Astronomy Survey Unit (CASU) are responsible for the VDFS pipeline processing component which has been designed for VISTA and scientifically verified by processing wide field mosaic imaging data from UKIRT's NIR mosaic camera WFCAM and is now routinely being used to process data from the WFCAM at a rate of up to 250GB/night. It has also been used to process ESO ISAAC data e.g. the FIRES survey data and a wide range of CCD mosaic camera data.

The pipeline is a modular design allowing straightforward addition or removal of processing stages and will have been tested on a range of input VISTA datasets. The standard processing includes: instrumental signature removal – bias, non-linearity, dark, flat, fringe, cross-talk; sky background tracking and homogenisation during image stacking and mosaicing – possible extras may include removal of other 2D systematic effects from imperfect multi-sector operation of detectors; assessing and dealing with image persistence from preceding exposures if necessary; combining frames if part of an observed dither sequence or tile pattern; producing a consistent internal photometric calibration to put observations on an approximately uniform system; standard catalogue generation including astrometric, photometric, shape and Data Quality Control (DQC) information; final astrometric calibration based on the catalogue with an appropriate World Coordinate System (WCS) placed in all FITS headers; photometric calibration for each generated catalogue augmented by monitoring of suitable pre-selected standard areas covering the entire field-of-view to measure and control systematics; frames and catalogue supplied with provisional calibration information and overall morphological classification embedded in FITS files; propagation of error arrays and use of confidence maps; realistic errors on selected derived parameters; nightly extinction measurements in relevant passbands; pipeline software version control – version used recorded in FITS header; processing history including calibration files recorded in FITS headers.

6.3.2 Science archiving

The concept of the science archive (SA, Hambly et al. 2004 and references therein) is key to the successful exploitation of wide field imaging survey datasets. The SA ingests the products of pipeline processing (instrumentally corrected images, derived source catalogues, and all associated metadata) into a database and then goes on to curate them to produce enhanced database-driven products. In the VDFS science archive, the curation process includes, but is not limited to, the following: individual passband frame association; source association to provide multi-colour, multi-epoch source lists; global photometric calibration; enhanced astrometry including derivation of stellar proper motions; consistent list-driven photometry across sets of frames in the same area; cross-association with external catalogues; and generation of new image products, e.g. stacks, mosaics and difference images etc., all according to prescriptions set up for a given survey programme. Archive curation includes quality control procedures, as required and led by the public survey consortium, and supported by archive team members. All these features are available in the context of a continually updating survey dataset from which periodic releases (as required by the community) can be made.

Moreover, end-user interfaces were catered for from the beginning in the VDFS design process, and the philosophy has always been to provide both simple and sophisticated interfaces for the data. The former is achieved via simple point-and-click web forms, while the latter is achieved via exposing the full power of the DBMS back-end to the user. To that end, full access to Structure Query Language and the relational organisation of all data are given to the user.

We have developed a generalised relational model for survey catalogue data in the VDFS. The key features to note are the normalised design with merged multi-waveband catalogue data (the table of most use for scientific queries) being part of a related set of tables that allow the user to track right back to the individual source images if they require to do so; and also that the merged source tables (as derived either from individually analysed images, or consistently across the full passband set available in any one field) are seamless, and present the user with a generally applicable science-ready dataset. Similar relational models describe the organisation of all data in the science archive (image, catalogue, calibration metadata, etc.) - see Hambly et al. (2004) and references therein. The science archive has a high-speed query interface, links to analysis tools such as TopCat, and advanced new VO services such as MySpace. Data products are being successfully ingested into the WFCAM Science Archive (WSA) in Edinburgh, with the EDR in Feb 2006, and the WSA concept was also demonstrated on the SuperCOSMOS Science Archive (SSA). *http://surveys.roe.ac.uk/ssa/*

The WSA is designed to enable links to external datasets (e.g., Alma, IRAC) enhancing the value of the data for the community.

6.4 Expected data products:

Instrumentally corrected frames (pawprints) along with header descriptors propagated from the instrument and processing steps (science frames and calibration frames)
Statistical confidence maps for each frame
Stacked image data for the dithered observations
Derived catalogues (source detections from science frames with standard isophotal parameters, model profile fitted parameters, image classification, etc.)
Data Quality Control database
Database-driven image products (stacks, mosaics, difference images, image cut-outs)
Multi-colour, multi-epoch source catalogues with global photometric calibration, proper motions (where appropriate)
Source remeasurement parameters from consistent list-driven photometry across all available bands in the field

6.5 General schedule of the project:

If the first six months of data starts to arrive at time T0, the VDFS can deliver the reduced data to the ULTRA-team at T0+8months. This will then require further Quality Control and checking. Hence we expect that we can release these data to the public at T0+12months. It is foreseen that this schedule can be speeded up later, as we are re-observing the same field over and over again.

It is foreseen that the data will be re-processed several times, as our knowledge of the field improves. This is our experience with the very deep "FIRE-Survey".

References:
Dye S. et al, 2006 in prep, The UKIDSS Early Data Release
Emerson J.P. et al., 2004, "VISTA data flow system: overview", in Optimizing scientific return for astronomy through information technologies, eds. P.J. Quinn & A. Bridger, Proc. SPIE, vol. 5493, 401
Hambly N.C. et al., 2004, "VISTA data flow system: survey access and curation; the WFCAM science archive", in Optimizing scientific return for astronomy through information technologies, eds. P.J. Quinn & A. Bridger, Proc. SPIE, vol. 5493, 423
Irwin M.J. et al., 2004, "VISTA data flow system: pipeline processing from WFCAM and VISTA", in Optimizing scientific return for astronomy through information technologies, eds. P.J. Quinn & A. Bridger, Proc. SPIE, vol. 5493, 411
Lawrence A. et al, 2006 in prep, The UKIRT Infrared Deep Sky Survey

7 Envisaged follow-up: (1 page max)

There are four rather obvious, but different directions of follow-up to this proposal: 1) very deep IRAC imaging of the same field, to obtain rest-frame optical flux points for the galaxies z 6-10; 2) ALMA imaging of at least a sub-set of this area to obtain rest-frame 60-200μm thermal-IR fluxes from which to derive (obscured) star-formation rates; 3) WFC3 imaging on HST a very small subset of selected areas within this field; 4) IR spectroscopy at large ground-based telescopes, both to verify redshifts and do extract astrophysical parameters.

With IRAC, a flux limit of AB$\approx$ 26 can be reached at 3.6/4.5μm in very deep integrations (20 hours) which should allow the detection of z$\sim 6-9$ galaxies, if they have any evolved population. The search for a "Balmer-break", to constrain the stellar population ages will be a crucial piece of information. As these exposures would be 10-times deeper than the currently approved IRAC COSMOS observations, they can be requested for Spitzer. Our experience with deep IRAC data (Damen et al, in prep.) shows that source confusion will be a tractable problem, even at these depths.

Detecting the thermal-IR flux that corresponds to 10 $M_\odot$/year in (absorbed and re-radiated) star-formation at z$\sim$ 8 is in principle staightforward with ALMA: it takes 30 minutes. Given ALMA's field-of-view, or $\sim 1' \times 1'$, a complete coverage of 2000 sqarcmin is probably too expensive, especially as the role of dust-obscured star-formation at these early epochs is not known. It may play much less role. A pilot study (10% of the area), or a selected targetted subset (e.g. 100 z$\sim$ 8 galaxies), is presumably the right first step in this direction.

While we have illustrated that WFC3 is much slower than VISTA as a wide-field, deep survey instrument, it can go much deeper over a modest number (a few ?) of WFC3 pointings. Providing limits on the faint end of the luminosity function in a small area with WFC3 and getting some structural information (sizes, close pairs, etc..) on the overall source population (z> 6.5) in the course of it, is clearly an effective use of HST time.

Ultimately, spectroscopy is needed for a complete understanding of the earliest galaxy population. From the ground, with instruments such as KMOS on the VLT or, presumably starting earlier LUCIFER on the LBT, only emission-line objects will be accessible to such spectroscopy studies at z> 6.5. However, if the fraction of emission-line objects (esp. Lyα) is comparable to z=3-6, there should be plenty of targets available. Foremost, in this follow-up is the verification at least for a small subset, that the photometric redshift estimates are correct. Later follow-up can determine the statistics of Lyα emittors as a function of redshift, one of the indicators of re-ionization (but not easy to interpret). *ULTRA-VISTA will provide targets for KMOS (just) in time before* *it comes on the telescope. Delaying this survey will significantly impact the availability of such targets.*

8 Other remarks, if any: (1 page max)

Experience of the team

The team has experience with all aspects of the data reduction and data-analysis. Specifically, members of the team have led the reduction and analysis of the deepest groundbased Near-IR imaging performed until now (Labbe, Franx, Rix). Franx has extensive experience in the analysis of deep photometry to find $z \geq 6$ galaxies (see the paper by Bouwens et al. and references therein). Rix, and Bell have extensive experience in the analysis of large datasets (COMBO-17, GEMS). van der Werf, Smail, Dunlop have extensive experience in the analysis of sub-mm data. Rix, Bell, Franx, and van Dokkum have extensive experience with the reduction and analysis of large IRAC imaging datasets.

We note that one possible concern for this type of program might be whether the signal-to-noise ratio keeps increasing with exposure time. It is been our experience that this is indeed the case. As a matter of fact, all our experience with very deep imaging with groundbased and space based Near-IR imagers (ISAAC, NICMOS) has been that the only problems are due to patterns or artefacts which are clearly seen in 1 hour exposures. We note that ISAAC suffered from horizontal and vertical patterns, and NICMOS from artefacts likely due to cosmic ray persistence. Once the reduction process has been adapted to take this "noise" out, the extensive dithering (over 100x100 pixels !) makes sure that the noise keeps decreasing with exposure time. We note these artefacts will need to be cleaned up for all programs using VISTA. Hence we see no reason to suspect particular problems for this program. Obviously, extensive testing of the pipeline will be necessary to optimize the quality of the output stack.

Towards a cohesive program for VISTA

We understand that we may have to negotiate with other proposers if similar proposals are submitted by others. We wish to stress, however, that the depth requested is critical to reach the science goal of studying a large number of galaxies beyond $z = 6.5$. Hence it is important that the committee makes a recommendation concerning the desirability of this depth before discussions start.