1 Title: The VISTA 5k WIDE Survey

PI: F.J. Castander, Institut d'Estudis Espacials de Catalunya/CSIC, Spain

Co-Is: F. Abdalla (UCL, UK), J. Annis (FNAL, USA), N. Benítez (IAA, Spain), S. Bridle (UCL, UK), J. Carlstrom (Chicago, USA), A. Collister (Cambridge, UK), R. Crittenden (Portsmouth, UK), P. Doel (UCL, UK), E. Fernández (IFAE, Spain), P. Fosalba (IEEC, Spain), J. Frieman (Chicago, USA), J. García-Bellido (UAM, Spain), E. Gaztañaga (IEEC, Spain), L. Infante (PUC, Chile), O. Lahav (UCL, UK), J. Lalli (?), A. Liddle (Sussex, UK), H. Lin (FNAL, USA), M. Martínez (IFAE, Spain), J. Maza (U. Chile, Chile), R.G. McMahon (Cambridge, UK), R. Miquel (IFAE, Spain), J. Miralda-Escudé (IEEC, Spain), J. Mohr (UIUC, USA), R.C. Nichol (Portsmouth, UK), W. Percival (Portsmouth, UK), K. Romer (Sussex, UK), W.J. Sutherland (Cambridge, UK), J. Weller (UCL, UK)

1.1 Abstract: (10 lines max)

The existence of Dark Energy poses one of the greatest challenges to our understanding of Fundamental Physics and the Universe. We propose to carry out a 5 thousand square degree survey with VISTA in the J and Ks bands. The main goal of the survey is to measure the dark energy and dark matter densities and the dark energy equation of state by probing the largest scales of the universe sampling large volumes at different epochs. For this purpose, we will measure the baryonic acoustic oscillations , we will cross-correlate the observed galaxies with the Cosmic Microwave Background radiation and we will measure the abundance of clusters of galaxies up to $z \sim 1.5$. We have chosen to target the Dark Energy Survey area to take advantage of the deep $griz$ to be obtained by this survey which will be publicly available. This will provide the necessary multi-colour information to compute accurate photometric redshifts. Finally, the VISTA 5k WIDE survey intends to deliver a legacy survey to the public covering a large fraction of the South Galactic Cap to unprecedented depths (~ 5 magnitudes deeper than 2MASS) with which many project could be undertaken (e.g., high redshift QSOs, cool stars, EROs) as well as provide targets for second generation VLT instruments like KMOS.

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

2.1 Scientific rationale:

One of the main goals of Cosmology is to understand what the universe is made of and how it evolves. Both questions are related as the universe composition, its matter and energy content, drives its evolution. In the last years, we have made tremendous progress in knowing the constituents of the Universe. based on several observational probes: type Ia supernovae, the large scale structure of the universe studied through galaxy clustering, the power spectrum, gravitational lensing, baryon acoustic oscillations, clusters of galaxies, the Lyman alpha forest and the Cosmic Microwave Background radiation and its cross-correlation with the galaxy distribution. Based on these data we now believe that the universe is made of approximately 5% ordinary baryonic matter, 25% dark matter and 70% dark energy (e.g., Spergel et al 2003, Tegmark et al 2004b). However, we have no idea of what these two last components are. So, we are faced with the embarrassment of ignoring what constitutes 95% of the universe.

Another fundamental question in Cosmology and Extragalactic Astronomy is how the structures we observe (stars, galaxies, clusters) have formed and evolved. According to the standard lore, known as the gravitational instability picture, structures originate from primordial (inflationary) fluctuations of the density field which later grow due to gravity (Peebles 1980,1993; Padmanabhan 1993 and references therein). In order to fully characterize the structures we actually observe we need to understand how visible matter, such as baryons, interact with invisible (or dark) forms of matter and energy that drive the gravitational dynamics of the Universe. Thus, the energy-density content of the universe also plays a key role in the growth of the large scale structure of the universe.

In order to tackle these questions, we need to study the overall structure and evolution of the universe and its constituents. We thus need to sample large enough volumes as to be meaningful and representative. We also need to probe different epochs to learn about its evolution. Observationally, these requirements are only met by large surveys that probe wide areas and a large fraction of the local Universe and that are deep enough to sample structures at high redshift.

The realization that the expansion of the universe is accelerating has come as a great challenge to our knowledge of the structure and composition of the universe. Little is known about the explanation for this acceleration, but whatever the actual mechanism that drives the currently accelerated expansion is, commonly known as "dark energy", it will have profound implications for our understanding of the universe. According to General Relativity, if the Universe is filled with ordinary matter, the expansion should be slowing down due to gravity. Since the expansion is speeding up, we are faced with two logical possibilities, either of which would have profound implications for our understanding of the fundamental laws of physics: (i) the Universe is filled with a completely new kind of stress-energy with non-standard properties (in particular, negative effective pressure), or (ii) General Relativity breaks down on cosmological scales and must be replaced with a new theory, perhaps associated with extra dimensions.

2.2 Immediate objective:

In order to advance in our understanding of the universe, we propose a large area survey of 5000 deg^2 to be carried out with VISTA in the near infrared. We have designed the survey as to coincide with the Dark Energy Survey (DES; Annis et al 2005) area. In this way we can combine the deep *griz* DES photometry with the J and Ks VISTA photometry to provide a good sampling of the spectral energy distribution of the objects studied and therefore deliver accurate photometric redshifts.

The precision to which we can study the expansion rate of the universe and the growth of structure, and therefore determine the underlying cosmology, depends strongly on the accuracy to which we can determine the redshift of galaxies. Mock galaxy and galaxy cluster simulations show that the DES optical *griz* photometry will achieve robust photo-z's with errors $\sigma_z \sim 0.02$ for whole galaxy clusters and $\sigma_z \sim 0.1$ for individual galaxies, out to redshifts $z \sim 1.3$. However, because the $4000\mathring{A}$ break feature, which provides the primary photo-z signal, redshifts out of the DES z filter beyond redshifts $z \sim 1.3$, DES photo-z's will necessarily be less accurate and more susceptible to catastrophic errors at such higher redshifts. It is precisely at these higher redshifts where the addition of VISTA near-IR data will enable a significant enhancement to the photometric redshift quality. The photo-z benefits of the proposed VISTA JK data are readily illustrated in Figure 1. Here we have constructed a mock galaxy catalogue simulating a magnitude-limited DES galaxy sample with $20 < i < 24$ and redshifts $0 < z < 2$ (the mock catalogue is based on evolving galaxy luminosity functions and type distributions from the data of Lin et al. 1999, Poli et al. 2003, Capak et al. 2004, Wirth et al. 2004). Photometric redshifts have been computed using a neural-network technique, using either the DES *griz* filters only (left panel), or in combination with the proposed VISTA JK data (right panel). The improvement made possible by the near-IR data at higher redshifts is clearly evident, as the addition of JK data now allows us to constrain the crucial $4000\mathring{A}$ break beyond the $z \sim 1.3$ limit of the *griz*-only optical data. The combined $griz+JK$ data result in a 1σ (68% confidence) photo-z error $\sigma_z < 0.1$ for all redshifts out to $z \approx 1.7$. Overall, the combination of DES *griz* and VISTA JK data results in a significant reduction in the photometric redshift errors and will consequently help to improve dark energy cosmological parameter constraints.

There are several techniques to constrain the expansion rate of the universe and the growth of structure and thus dark energy and dark matter. In this proposal we focused in those requiring large area and volume, which include:

1) Galaxy clustering: On large scales, galaxy clustering and its evolution reflect the gravitational clustering and dynamics of the underlying dark matter distribution. In the linear regime, we can write the galaxy power spectrum as: $P_{gal}(k) \propto k^n \, T^2(k;p_i) \, g^2(z;p_j) \, b^2(k;z)$, where the initial dark matter power spectrum from the early universe is $\propto k^n$, $T(k)$ is the scale dependent transfer function for dark matter perturbations, $g(z)$ is the scale-independent linear perturbation growth, b is the galaxy bias and p_i indicates that these quantities

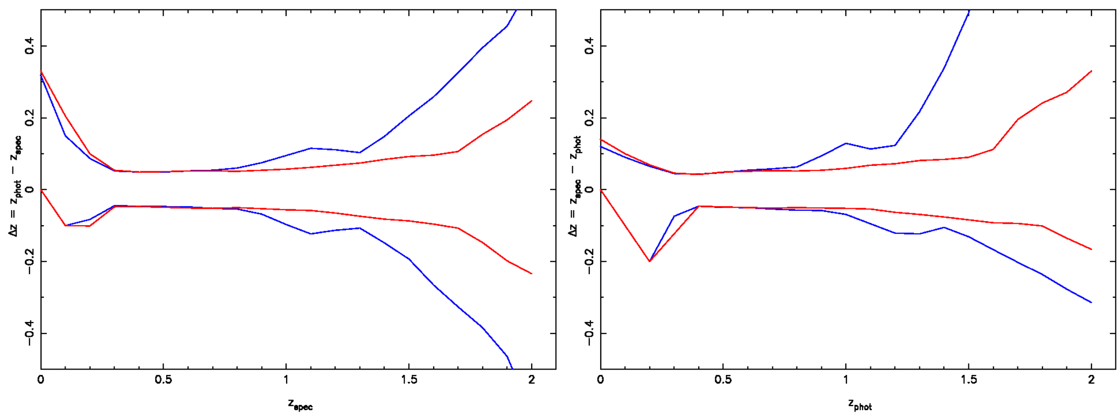

Figure 1: Difference between photometric redshifts and spectroscopic redshift for our simulations for DES $griz$ photometry (blue) and DES+VISTA $grizJK$ photometry (red) as a function of spectroscopic redshift (left) and photometric redshift (right).

depend on cosmological parameters. With the VISTA 5k survey we will measure approximately 300 million galaxies with will we compute the galaxy power spectrum to constrain these quantities. Moreover, there are characteristics scales in the galaxy power spectrum like the horizon scale at matter-radiation equality and the baryon oscillations, which provide standard rods that can be used to measure the angular diameter distance as a function of redshift. Because, these features depend primarily on the shape of of the galaxy power spectrum, they are relative insensitive to uncertainties in galaxy biasing.

In particular, baryon acoustic oscillations (BAOs) are imprinted in the galaxy distribution in the early universe due to oscillations of the coupled photon-baryon fluid and provide a standard ruler. These BAOs can be detected on the largest scales where gravity non-linear effect have not erased them. They have been recently significantly detected by the 2dFGRS and the SDSS teams.

In Figure 2, we show the expected constraints on the $\Omega_{Lambda} - w$ parameter space that we expect to obtain from DES alone and DES+VISTA, based on large cosmological simulations with the photometric redshift expected error distributions. With the measurement of the BAOs by themselves we expect to constrain w to 30% with the VISTA 5k survey.

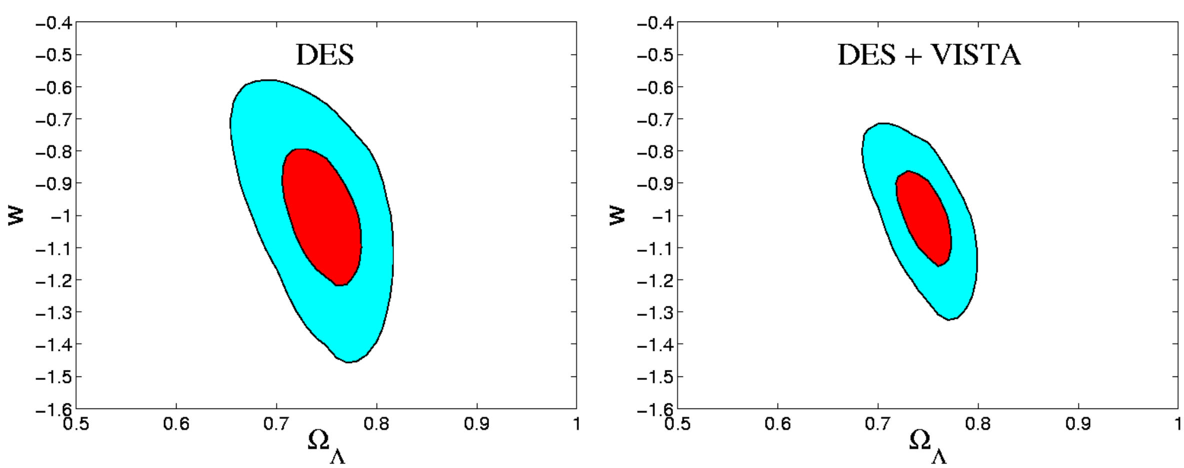

Figure 2: Expected 1σ (red) and 2σ (blue) likelihood contours in the w .vs. Ω_Λ parameter space in our simulations. Left: DES survey, Right: DES+VISTA.

Higher order statistics provide additional information about cosmological parameters and biasing. This additional information can also be used to determine physical properties of the tracers in a way that is independent of cosmological parameters or the linear growth of structure. We will use the VISTA 5k to compute the higher order statistics of its galaxy distribution.

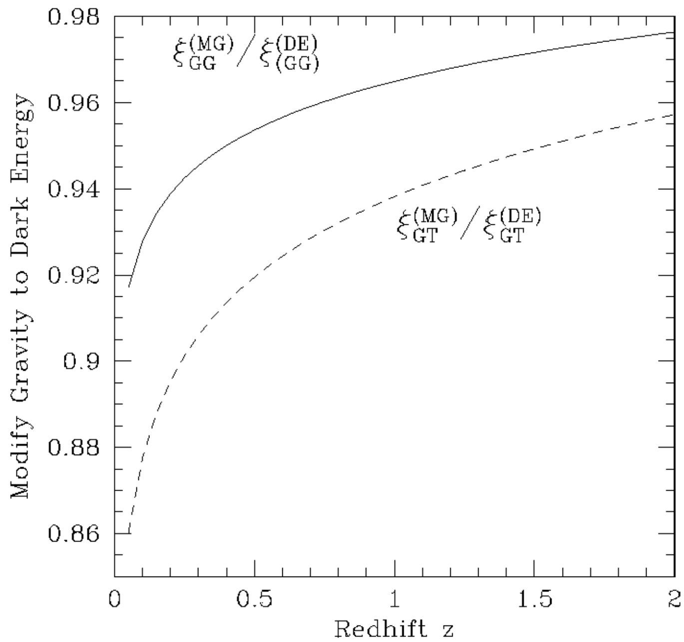

Figure 3: Galaxy-Galaxy (GG) and Galaxy-Temperature (GT) cross-correlations ratio for modified gravity (MG) and dark energy (DE) models.

2) Integrated Sachs-Wolfe Effect (ISW): The CMB photons get redshifted gravitationally as they traverse potential wells. If the universe accelerates (or decelerates) the net effect of the energy gained when falling into the potential wells and lost when coming out is different from zero. If we cross-correlate the CMB with a tracer of these changing potential wells we obtain a non-zero signal. The ISW is then a direct and powerful measurement of dark energy via the growth of cosmic structure.

Various groups have reported the detection of the ISW in a variety of cross-correlations with the WMAP 1st year data. We will cross-correlate the Planck CMB data with the galaxy distribution of the VISTA 5k survey in redshift slices with which we will be able to probe the evolution of the growth of structure.

3) Gravitational Lensing: The gravitational bending of light by mass in the universe offers the possibility to study the distribution of dark mater. Furthermore, measurements of how this distribution changes with time tell us about the nature of dark energy. There are several regimes/measurements that can be used to constrain dark energy and dark matter. Large scale structures generate correlated shear of low amplitude. It can be measured with shear-shear correlations (also called cosmic shear). One can also measure the angular correlation between foreground galaxies with galaxy shear (galaxy-shear correlations or galaxy-galaxy lensing).

Cosmic shear depends on dark energy through the distance-redshift relation, space curvature and the evolution of the growth factor. Its cosmological sensitivity depends on several factors, mainly the fraction of sky surveyed, the effective number of galaxies per unit area and the variance on the intrinsic ellipticities of galaxies. The effective number of galaxies depends on the survey depth and the image quality given by the observed seeing.

what are we going to do with VISTA 5k? Comment that we do not set a seeing requirement but large area is powerful anyway

4) Galaxy clusters: The redshift distribution of clusters of galaxies provides an exquisite probe of the growth of structures in the Universe. Since the growth of structure depends sensitively on the amount of matter in the Universe, clusters of galaxies provide a window on the make up of the Universe at different redshifts. The South Pole Telescope (SPT) will select on the order of 20.000 clusters in a 4.000 $\deg^2$ area via their Sunyaev-Zel'dovich decrement. The Sunyaev-Zel'dovich decrement has the advantage to be independent of redshift, hence allowing the detection of clusters at very large redshift. This is in particular the case for a one arcminute beam instrument, such as SPT.

While the abundance of clusters depends sensitively on the amount of matter ($\Omega_{\rm m}$) and the normalization of the primordial fluctuations (σ_8), the distribution with redshift is influenced by the equation of state of dark energy (w). In order to measure w precisely with the redshift distribution of clusters, it is necessary to obtain their redshifts, if possible, out to $z = 1.5$. To push redshift estimates so far with photometric methods requires J and K band observations. VISTA provides a unique opportunity to achieve this goal.

One of the main difficulties to use the evolution of galaxy clusters as cosmological probes comes from cluster selection which has to performed in some observable different from the cluster mass, although it can be somehow circumvented with the so called self-calibrating techniques. In that respect, the J and K photometry will allow us to measure the stellar mass of galaxy cluster members, which correlates well with cluster mass *Bob you can expand here and link these ideas together*. We will also use weak gravitational lensing in clusters to calibrate the observable-mass relations.

5) Dark Energy models: In the simplest cosmological models with dark energy, the universe is spatially flat, and the dark energy parameters to be considered are $\Omega_{DE} = 1 - \Omega_m$ and the equation of state parameter $w = p/\rho$. One way to extend the simplest models is to assume that w depends on redshift. The best way to study its dependence is via tomography studies of dark energy probes. In that respect the improvement in the photo-z's with VISTA help to improve the constrains on the possible evolution of w.

The use of different approach to study dark energy can also shed light on the very nature of dark energy:is dark energy a new energy source or just reflects a change in the laws of gravity? The Friedman equation alone cannot be used to separate these two possibilities as it is possible to recast a change in the laws of gravity as a change in the energy content. However, for a given Friedman equation, the growth of perturbations can be different for these two scenarios. The combination of DES probes of rate of expansion and the growth of structure can then be very useful to study the nature of dark energy. One possibility is shown in figure 3, where we show the ratio of the Galaxy-Galaxy (GG) and Galaxy-Temperature (GT) cross-correlations for modified gravity (MG) and dark energy (DE) models.

Besides the cosmological studies described before, the VISTA 5k survey is intended to be a Legacy Survey for the community in the same manner as 2dF, SDSS or 2MASS. It will cover a large fraction of the South Galactic Hemisphere to unprecedented depths for large area, approximately 5 magnitudes deeper than 2MASS. Many project could be undertaken in the general areas of galaxy evolution, quasar studies, stellar studies, galactic structure, nearby objects. The VISTA 5k survey will also provide targets for second generation VLT instruments like KMOS.

Amongst the broad science topic that can be addressed wit the VISTA 5k data, our group is interested in the following:

Quasar science. The combination of the DES and VISTA 5k will be very powerful to select QSO samples. We expect to detect approximately 1 million QSOs with $0 < z < 3$, 50000 with $3 < z < 4$, 10000 with $4 < z < 5$, 1000-2000 $5 < z < 6$, 50-100 $6 < z < 7$ and 10-20 $7 < z < 8$. Apart from the statistical studies of quasars, follow up spectroscopy can yield interesting results studying the QSO absorption lines. The highest QSO will also be very valuable to study the epoch of reionization. The clustering of QSOs will also provide cosmological information. QSO pairs can help study, for example, the quasar influence in the intergalactic medium. The study of lensed quasars would also be interesting g for cosmological studies and halo modeling. The combination of DES+VISTA will also provide information on quasar variability.

Low mass stars. Brown dwarfs and very low mass stars are the more numerous objects in the galaxy but they are difficult to detect due to their faintness. They emit the bulk of their radiation in the near infrared where, however, they are characterised by rich molecular and atomic absorption spectra. Very low mass stars studies are important in the fields of cosmology (are they primordial?, galaxy dynamics (how much mass they contribute to the galaxy?), star formation (how is the IMF?) as well as in extrasolar planets (how to distinguish between star and extrasolar planet?).

Photometric searches for these low mass objects usually concentrate in the near infrared. However, due to the absorption bands present in their near infrared spectra, optical data are also necessary for a efficient selection of these low mass stars.

Proper motion objects. The multiple imaging of a large area of the sky during a long time baseline will make the combined DES and VISTA 5k database very valuable to detect moving objects. On the largest time scales, we can probably detect halo white dwarfs proper motions that should move > 0.5" in 5 years. Halo WD can be very useful to date the age of the halo and study its contribution to the halo mass. On the shortest time scales, we can study nearby objects and/or Kuiper belt objects. The luminosity of these bodies in the J and Ks bands are dominated by reflected solar radiation although with the peculiarity that both bands can provide useful information on the presence of silicates in the surfaces of these bodies. In fact, the interval located between 0.7 and 2.5 μm gives information on pyroxenes, olivines and plagioclase because of the absorption bands of these minerals are centered near 1 and 2 μm. Additional windows in the visible can provide valuable spectral information to identify the nature of the imaged objects.

PROVIDE TARGETS for FUTURE SPECTROGRAPHS

CLOSING WRAP-UP SENTENCE

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

Richard (or somebody else), can you deal whit this point?

KIDS

VHS

All other surveys in VISTA Twiki page

Talk about DES as well.

4 Observing strategy: (1 page max)

The VISTA 5k survey science goals require probing large volumes at different epochs thus needing large area coverage and depth. In terms of area, we have decided to cover the Dark Energy Survey (DES) area to take advantage of its deep *griz* public photometry. In order to improve the photometric redshifts beyond the redshift where DES starts to be inaccurate as the 4000 Å break moves out of the z band, we need redder bands that reach deep enough to detect higher redshift galaxies. We have carried out simulations to estimate the optimum sensible depth and the number of near infrared bands needed to allow us reach our scientific goals. We find a good compromise using just two bands J and Ks down to a depth of $Ks \sim 19$ and $J \sim 21$.

The area we want to cover is the DES area which is divided in 3 patches in the South Galactic Hemisphere (see figure). One large patch of 4000 deg^2, the DES/SPT area (to be covered by both surveys) with coordinates RA: 20h-7h and DEC:-65deg to -30deg and RA:19h-20h and DEC:-65deg to -45deg. Another patch, the DES/Connecting area of 800 deg^2 with coordinates: RA:1h20m-3h20m and DEC:-1deg to -30deg. A third patch, the DES/SDSS equatorial area of 200 deg^2 with coordinates: RA:20h40m-3m20m and DEC:-1deg to 1deg.

Since our science priority is area and depth, our strategy is to cover the whole area with the Ks filter first and then repeat the area with the J filter. The standard tiling strategy of 6 pawprints to cover a tile of approximately 1.5 deg x 1.0 deg is adequate for our purposes. In fact, it is well-suited to match the width of the DES equatorial strip with two adjacent VISTA tiles.

According to the ETC v1.1, we need 320 s on source to reach $Ks = 19.0$ and 300 s to get to J. This value is for point sources in 1.0" seeing conditions measured in a 2" diameter aperture. The actual depth will depend on the seeing. But according to Paranal seeing statistics a seeing equal or better than 1.0" is reach approximately 75-80% of the photometric time.

The expected bad pixels in the detectors is approximately 1.5%, and therefore any position of the sky needs to be observed with more than one pixel to be on the safe side. In the standard 6 pawprints to cover a tile, every position of the sky is imaged by at least two pixels. To avoid problems with bad pixels, we have decided to take three jitters in each tile so that every position of the sky is at least image with six different pixels. More jitters, although desirable, would make the overheads too expensive.

We have thus settled for a strategy in which each tile is taken with 6 pawprints in the standard VISTA fashion. For the Ks band each pawprint pointing is composed of 6 coadds of 9 second detector on-chip integration (DIT) exposure, and we take 3 jitters. This accounts for 324 s on source per tile, with total exposure time per tile 972 s and total elapsed time per tile 1310 s, with an observing efficiency per tile of 74.2%. For the J band the maximum DIT time can be longer, and therefore we have settled for 2 coadds of 25 s and 3 jitters. This combination gives 300s on source, 900 s exposures per tile and a total elapsed time of 1166 s with an observing efficiency of 77.2%. See table 4

Part of our VISTA 5k area is also included in the KIDS survey. See figure 4. The overlap between both surveys is approximately 480 deg^2. We do not intend to re-image the KIDS area as we can use their near-infrared imaging for our survey as they intend to go slightly deeper in our chosen bands and, in fact, image in more bands.

In order to improve our efficiency, we will finish tiles before slewing to the next tile. Covering for example the area 3 times with a third of the exposure time each time (no jitters but visit the tile 3 different times e.g., once a year) would incur in too onerous overheads. Our strategy would also be to obtain contiguous tiles in each observing night, so as to minimise the slew time.

Although desirable, we do not require good seeing conditions and can observe with up to 1.2" seeing conditions. Photometric conditions are not required. We intend to calibrate our photometry with the 2MASS survey and therefore we have no photometric calibration overheads.

Table 4. Tiling exposure strategy

	Ks	J
Detector on-chip integration (DIT)	9.0 s	25.0 s
Exposure coadds (Ndit)	6	2
Exposure loops (Nexp)	1	1
Jitter pattern (Njitter)	3	3
Number of pointing (Npaw)	6	6
Time on source	324 s	300 s
Total exposure time per tile	972 s	900 s
Total elapsed time per tile	1310 s	1166 s
Observing efficiency per tile	74.2%	77.2%

5 Estimated observing time:

Table 5. Estimated observing time for a 5 year survey completion

Period	Time (h)	Mean RA	Moon	Seeing	Transparency
P80	270	19h30m-07h00m	bright/grey/dark	1.2	clear
P81	165	19h30m-07h00m	bright/grey/dark	1.2	clear
P82	270	19h30m-07h00m	bright/grey/dark	1.2	clear
P83	165	19h30m-07h00m	bright/grey/dark	1.2	clear
P84	270	19h30m-07h00m	bright/grey/dark	1.2	clear
P85	165	19h30m-07h00m	bright/grey/dark	1.2	clear
P86	270	19h30m-07h00m	bright/grey/dark	1.2	clear
P87	165	19h30m-07h00m	bright/grey/dark	1.2	clear
P88	270	19h30m-07h00m	bright/grey/dark	1.2	clear
P89	165	19h30m-07h00m	bright/grey/dark	1.2	clear

5.1 Time justification: (1 page max)

Our survey simulations indicate that we can reach our science goals if we reach limiting magnitudes of $Ks \sim 19$ and $J \sim 21$. According to the ETC v1.1, we need 320 s in Ks and 300 s in J, to reach this depth for point sources in 1.0" seeing conditions measured in a 2" diameter aperture. In fact our simulation are done taking into account these exposure times.

Taking into account the overhead implied by our observing strategy (see previous section) and adding an extra 5% time overhead for slews to acquire tiles, we estimate using the ETC that we need 1310 + 5% = 1375 s in Ks and 1166 + 5% = 1224 s in J per tile. To cover the DES area without including the KIDS overlap (approx 4500 deg^2) we need 3000 tiles of 1.5 deg $\times$ 1.0 deg. The whole area thus requires 1146 hours in Ks and 1020 hours in J. The DES area distribution in the sky implies that 38% of the area is best accessed in the April-September semester and 62% in the October-March semester. Assuming ten hours per night in April-September semesters and 8 hours per night in the October-March semesters, the total time amounts to 133 nights for the Ks band and 118 nights in the J band.

In table **??** we specify a possible time distribution per semester required by the survey for a completion time of 5 years. Approximately the first 2.5 years are devoted to the Ks band and the last 2.5 to the J band and assuming a 38% – 62% distribution into the April-Sept and Oct-march semesters with a start in semester P80.

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 XXX Survey specific tasks		
x	Data Quality Control-III	x	x
	etc.		

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.

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 XXX, 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/*

XXX is intrinsically a multi-wavelength project and most science will come from the linking of VISTA data with YY data and the WSA is designed to enable such links.

6.4 Expected data products:

Instrumentally corrected frames (pawprints, tiles etc) 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 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)
Frame associations yielding a survey field system; seamless, merged, 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 any one field
+ DON'T FORGET XXX survey specific products

6.5 General schedule of the project:

Note *(to be deleted - not part of the form)* **The times below refer to VDFS-related times only and aren't intended to suggest specific timescales of public release. There is, for example, no time included here for the Survey Consortium to do their own quality control on the VDFS products. The draft document wasn't intended to suggest specific timescales of public release, merely to indicate to the PIs that they should probably say something about this and adapt the template accordingly.For VST ESO expect to have data products delivered to the ESO archive within the semester following the one in which the raw data was sent out. By default they will likely expect the same for VISTA - so you may wish to factor this into the times on the proposed schedule for your project. On the other hand as selected proposals will later need to submit a separate full Data Management Plan spending a lot of effort on this now is probably not too important.**

T0: Start of observations
T0+4months; Release of science products from first month of survey observations
T0+8month; Release of science products from first 6 months of survey observations
Thereafter we anticipate that standard science products can be released to the PIs within 1-2 months of raw data arriving in the UK.
Optional reprocessing of data based on improved knowledge of the instrument would also be considered

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

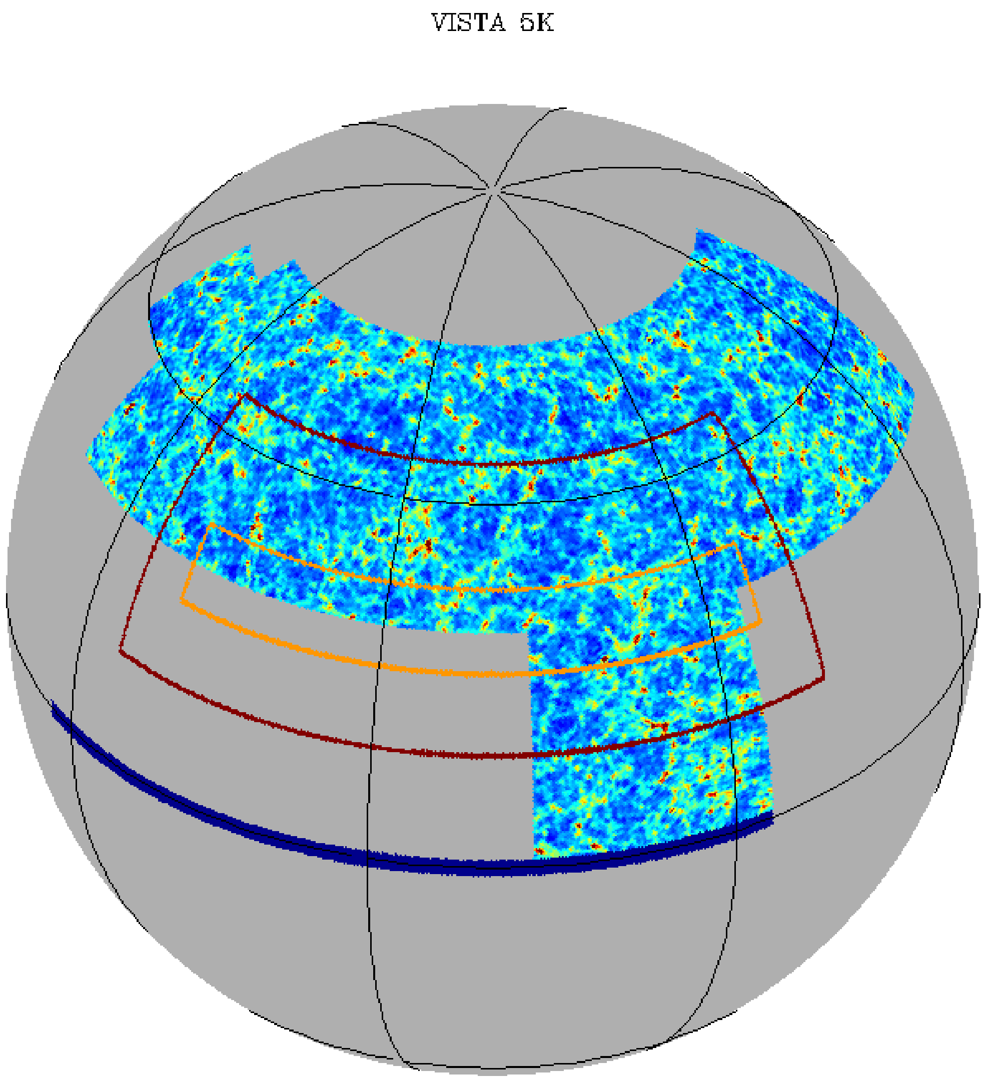

Figure 4: Area coverage of the DES survey with simulated galaxy density map (DES/SPT and DES/Connecting areas) and in blue the DES/Equatorial strip. In red the area coverage of VST Atlas and in yellow KIDS.