Hubble Deep Fields Initiative 2012
Science Working Group Report

Members:
James Bullock (Chair, UC Irvine), Mark Dickinson (NOAO), Steven Finkelstein (UT Austin), Adriano Fontana (INAF, Rome), Ann Hornschemeier Cardiff (NASA/GSFC), Jennifer Lotz (STScI), Priya Natarajan (Yale), Alexandra Pope (U. Mass), Brant Robertson (University of Arizona), Brian Siana (UC Riverside), Jason Tumlinson (STScI), Michael Wood-Vasey (University of Pittsburgh)

Ex-officio:
Neill Reid (STScI), Ken Sembach (STScI)

Table of Contents

1. Executive Summary

The first billion years of cosmic time saw primordial galaxies emerge from a near-featureless early universe and begin assembling into the chemically complex and morphological diverse population of galaxies we see strung across the cosmos today. The nature of this process – how stars and galaxies have emerged and evolved from the Big Bang to today – is one of the most compelling questions in science. From the original HDF through several iterations of the HUDF, the Hubble Space Telescope (HST) has been at the forefront of discovering and characterizing the earliest galaxies to emerge in the early universe.

Inspired by this legacy and with an eye to the future with JWST, the STScI Director asked the Hubble Deep Fields Initiative (HDFI) committee to identify opportunities for Hubble's next great advance in our knowledge of early galaxy formation. The Committee weighed valuable advice from the astronomical community and deliberated on a set of major science goals to motivate future deep Hubble observations. After considering a broad range of open questions in galaxy evolution, the committee has unanimously recommended a program of six deep fields centered on strong lensing galaxy clusters in parallel with six deep "blank fields". Our recommended program devotes 140 orbits to each field, achieving AB ≈ 28.7-29 mag in 7 optical/NIR photometric bands for a total of 840 orbits. The key science aims of these new deep fields are: (1) to reveal hitherto inaccessible populations of $z = 5 - 10$ galaxies that are 10 - 50 times fainter intrinsically than any presently known; (2) to solidify our understanding of the stellar masses and star formation histories of sub-L_* galaxies at the earliest times; (3) to provide the first statistically meaningful morphological characterization of star forming galaxies at $z > 5$; and (4) to find $z > 8$ galaxies stretched out enough by cluster lensing to discern internal structure and/or magnified enough by cluster lensing for spectroscopic follow-up.

This proposed Hubble Deep Fields Initiative (HDFI) optimally combines proven techniques for studying high-redshift galaxies in blank fields with the potentially revolutionary use of natural gravitational telescopes to exploit their magnification of the faintest galaxies in the distant universe. The blank fields will increase three-fold the area covered at comparable depth by the HUDF09 and its parallel fields, tracing the history of star formation and the growth of stellar mass with improved statistics and reduced cosmic variance. In the cluster fields, **HST can reach high redshift galaxies as faint intrinsically as those that JWST can detect in blank fields**, even down to the dwarf galaxies thought to be the progenitors of typical L_* galaxies in the modern universe. We specify scientific criteria for the choice of cluster fields based on their lensing strengths as characterized by magnification maps as well as their accessibility to existing and future observatories across the spectrum in space and on the ground. These fields will greatly expand the deep sky available to JWST in its first years, and refine the cluster lensing technique to extend its reach out to the first galaxies. The HDFI will extend Hubble's contributions to early galaxy evolution as far possible for a reasonable allocation and prepare the astronomical community to make the best use of JWST on some of the most compelling problems in astrophysics.

It is important to recognize that cluster lens fields do not necessarily enable the detection of *more* galaxies at a given redshift than a blank field survey to the same depth. However, they do enable the detection of *intrinsically fainter galaxies* than would otherwise be possible. These intrinsically faint galaxies are crucial high-z progenitors of normal Milky-Way size galaxies today and may provide key information about the sources that reionized the universe.

The use of galaxy clusters as gravitational telescopes will rely critically on understanding cluster magnification maps and associated volumes probed. The committee critically evaluated the robustness of currently available maps, and we are now unanimously convinced that the best lens maps are suitably accurate for the purpose of exploiting these natural telescopes, and can be honed even further with data from this program itself and related support for community development of improved maps. As part of this evaluation process, we asked three independent lensing groups to provide magnification maps to compare for two massive galaxy clusters Abell 1689 and the Bullet Cluster. As far as we are aware, this was the first method comparison of its kind. Even though the maps relied on slightly different input data sets (multiple image systems and their adopted redshifts), the approaches provided good agreement for typical magnifications (factors of ~5-10 in luminosity), only disagreeing significantly for the highest magnifications in the tails of the distributions. Such agreement will only improve once a more concerted effort is made to use exactly the same input (best) data for the lens model comparisons. Moreover, the lens models will become iteratively more accurate as the cluster images themselves improve.

As part of this recommendation, we are urging Space Telescope to issue a data challenge to select several independent groups of cartographers to produce magnification maps for the selected clusters. Maps should be part of the data products and will be crucial to enable non-lensing specialists to take full advantage of the proposed observations. As more data are accumulated as part of the HDFI project, the maps should be updated and refined as part of planned data releases. Crucially, *we recommend that the data challenge involve different lens codes/techniques and enlist multiple groups*. The associated magnification maps can then be built and associated model/uncertainty errors provided based on comprehensive model comparisons performed during the calibration stage of the exercise for the specific clusters observed.

This committee recognizes that Hubble has been a workhorse for cluster imaging over the years, but predominantly for programs aimed at studying clusters themselves. Of particular interest is CLASH, a 524-orbit Multi-Cycle Treasury program. The cluster-lensing component of the HDFI explores an entirely different parameter space than CLASH, some 1.5 magnitudes deeper per cluster. Unlike the shallower CLASH lensing fields, which are well suited for finding intrinsically bright high-z sources, the deep HDFI cluster lensing exposures will reveal significant numbers of high-redshift objects that are intrinsically fainter than anything possible with HST blank field imaging. Importantly, the depth of HDFI cluster lens fields is such that the parallel blank field science is compelling, filling a gap in area and depth that complements HUDF and CANDELS-Deep science. By going deeper than CLASH, we enable a complementary set of parallel blank field exposures that alone provide a factor of ~3-5 gain in faint high-redshift galaxies compared to current blank-field counts, and that will roughly double the number of z~9 candidates compared to existing blank fields.

The proposed combination of cluster fields and accompanying blank fields offers a two-pronged approach that will increase high-redshift galaxy counts and discover fainter high-redshift galaxies than will otherwise be possible without JWST. This program will also provide targets for JWST spectroscopy and for future AO-assisted, IFU spectrographs on 30m-class telescopes that could for the first time study the internal kinematics and properties of high-redshift galaxies.

2. Process

The Director of STScI convened the *HDFI* Committee to explore the science case for devoting a substantial fraction of his Director's Discretionary time in the next three cycles (21-23) to new deep observations of the universe. The primary motivation for this initiative is to extend our knowledge of the cosmic frontier at high redshifts through the collection of data that only HST can obtain. A secondary motivation is to lay the groundwork for future observations of the early universe with JWST.

The work of the HDFI committee occurred from July to November 2012 in a series of weekly telecons and two in-person meetings held in Irvine (August 3) and Baltimore (September 17).

The committee solicited input from the broader astronomical community in the form of short white papers that were due August 31, 2012. We received 35 responses to this call contributing a wide range of scientific and programmatic ideas; the titles and submitting authors of these white papers are listed here in an Appendix. These white papers were considered non-binding advice to the committee and all of them were carefully read and discussed by the committee. We evaluated all science cases that came to us, narrowed down the field and investigated four possible science cases in detail. These science cases are discussed in more detail in Section 3 of this document. We also solicited informal community input where necessary to clarify technical details about the science cases we considered.

The committee arrived at a consensus recommendation for the key science aims and program design and presented this to the STScI Director on October 24, 2012.

2.1 Charter

Given a practical limit of **~1000 orbits**, the Charter of the HDFI Committee charged the working group with the following tasks:

- Define the science case and a set of science goals for a new set of ultra-deep imaging fields with sensitivity comparable to those of the HUDF and HUDF-09 infrared follow-up. Provide an assessment of the urgency of pursuing this science.
- Assess the prospects for near-field science that can be achieved with these deep-field observations.
- Recommend the locations and number of fields that should be obtained to meet the science goals defined for the HDFI.
- Recommend the suite of filters and exposure times necessary to accomplish the science goals defined for the HDFI.
- Solicit input from the astronomical community in defining the science goals and recommendations described in the above tasks.
- Produce a report describing the results of the above tasks.

By forming this Working Group, STScI sought to ensure that many voices are heard in the formulation of the science case for the HDFI and its eventual data products, with a goal to maximize the science return and legacy value of the observations.

The Working Group took into account both the archival research value of the planned observations and the coordination of these observations with other observatories.

3. Science Cases Considered for the HDFI

The majority of received white papers advocated for additional deep optical/IR imaging of blank fields, with many papers focused on specific aspects of the implementation including field selection, coordination with other telescopes, and E/PO opportunities. Multiple white papers also suggested new science topics and approaches, including deep grism surveys, deep UV imaging/spectroscopy, and deep surveys of lensing clusters. Motivated by the white papers we received, the HDFI committee investigated four approaches in detail: 1) Deep blank fields; 2) Cluster lensing deep fields; 3) Grism deep fields; 4) COS deep fields. Here we summarize the key science for each of these strategies.

3.1 Deep HUDF-like Blank Fields

This committee was charged with specific goals to understand the process of galaxy formation and evolution at very high redshifts. Specifically, our charter instructs us: "The primary goal of this initiative is to extend our knowledge of the cosmic frontier at high redshifts through the collection of data that only HST is capable of obtaining" and to "identify the most effective strategy for complementing and supplementing the HUDF". Clearly, the Hubble Ultra Deep Field (both the pioneering optical and near-infrared observations with ACS and NICMOS in 2003-2004, and the deeper near-infrared observations with WFC3 in 2009 and 2012) has changed the field of extragalactic astronomy. Our charge was to specify a set of observations that can further transform the way we view our universe.

The strategy of "blank field" imaging, where one images a high Galactic latitude field with no bright foreground objects, has resulted in many great successes in the HST era. In addition to the Director's Discretionary blank-field programs of the HDF and HUDF, TAC approved programs like GOODS, COSMOS, and CANDELS have proven irreplaceable in our analyses of the universe. Thus, we began our examinations in a logical place, exploring whether further blank field observations could enable the desired transformative leap in our understanding of galaxy evolution.

We examined three blank-field imaging scenarios: 1) A pair of fields at roughly the depth of the HUDF (i.e., HUDF "clones") in a different region of the sky; 2) Multiple pairs of deep fields at ~0.5-1 mag shallower depth than the HUDF (similar to the HUDF09 parallel fields); 3) More fields of the area and depth of the two CANDELS-Deep observations (~1.5-1.7 mag shallower than the HUDF). Given STScI guidance about the practical size of the entire program, the required depths allow for one pair of HUDF-clones (~600-800 orbits), five pairs of HUDF-parallel-like fields (~1000 orbits), or two new CANDELS-Deep-like fields (~1100 orbits, as they would need both ACS and WFC3). Only Scenario #2 allows efficient parallelization by beam-switching ACS and WFC3. Out of these blank field options, Scenario 2 was carried forward for further investigation as the best possibility for significant scientific advances, as it would bring the number of fields at a depth of the HUDF09 parallels (or greater) from 3 to 13; a factor of >4X gain in area, compared to only 3X and 1X for the HUDF-clone and CANDELS Deep scenarios.

The HUDF has opened the door for a wide range of high-redshift discoveries, but the volume probed by the small number of fields limits further advances. We have identified several science topics that would gain substantially from several new deep field observations.

The faint end of the luminosity function: Estimates of the integrated ultraviolet (UV) luminosity density are essential for understanding the nature of sources that affect cosmic reionization (e.g., Robertson et al. 2010). Faint galaxies make a large and possibly dominant portion of the UV luminosity budget, but the current uncertainties on the faint-end of the high-redshift rest-frame UV luminosity functions are large, with currently published uncertainties on α at $z = 7$-8 ranging from 0.2 – 0.4 (e.g., Bouwens et al. 2011a; Oesch et al. 2012; Bradley et al. 2012; Grazian et al. 2012). The wide range of plausible faint-end slopes translates into unacceptably large uncertainties on the integrated UV luminosity density. There are presently only ~30 galaxies in the range $L < 0.5L^*$ known at $z > 7$; while this number will increase with the UDF12 results, a large addition of new blank deep fields would better constrain the measured UV luminosity density.

Stellar mass function*:* Stellar mass functions at $z > 6$ are presently ill-constrained, with the current state-of-the-art relying on extending relations between the stellar mass to UV light measured at $z=4$ to higher redshift (e.g., Gonzalez et al. 2011, Stark et al. 2012). Stellar mass functions measured directly with improved statistics at higher redshifts are key to understanding star-formation histories and efficiencies in the early universe.

Stellar Populations: Much of the interesting phase of galaxy evolution occurs in faint, low mass galaxies as they offer a probe of the most primordial progenitors of typical galaxies today. One goal of high-redshift studies is to push towards detection of the first galaxies, which by definition will be very metal poor, and almost certainly the faintest. A number of papers have examined this with the existing data, but the uncertainties are still high due to the small number of faint galaxies known at $z > 7$ (e.g., Bouwens et al. 2010, 2012; Dunlop et al. 2012, Finkelstein et al. 2012).

Galaxies at z ≳9*:* Images of the highest redshift galaxies provide one of the most direct clues to first light phenomena in the early universe: How quickly and efficiently do small dark matter halos begin to turn their baryons into stars? What number density (and associated halo mass) is associated with these first objects? The use of the F140W and F160W infrared bands allows the selection of galaxies to $z \gtrsim 9$, where only a few galaxy candidates are known (e.g., Zheng et al. 2012; Bouwens et al. 2011b; Coe et al. 2012). Distant objects at $z\sim10$ selected with only F125W and F160W have more uncertain redshifts (e.g., Bouwens et al. 2011b). While the UDF12 program adds F140W to the HUDF, increasing the area probed by both of these filters should yield > 10 galaxies at these redshifts, allowing the first robust construction of a luminosity function, providing key insight into the build up of galaxies at very early times.

Morphology and Structure: The shapes and internal structures of galaxies reveal the dynamical and gas processes that formed them. The presence of star-forming disks, red central bulges, clumpy substructure, and faint tidal tails are all tracers of a galaxy's recent assembly and evolution. Quantitative measurements of morphology require S/N > 3 per pixel, and spatially resolved stellar population studies require even higher S/N per resolution element. Only the HUDF and its parallel fields provide the depth necessary to quantify structures fainter than ~25-26 mag per sq. arcsec. Such faint limits are needed to study the majority of galaxies at $z > 5$, sub-L_* galaxies at $z\sim2$, and tidal features at $z>1$. Additional deep blank fields would substantially increase the number of $z > 1$ galaxies with detailed morphology and stellar population measurements.

***Preparing for JWST**:* One of the breakthroughs of JWST will be its spectroscopic multiplexing capability coupled with space-based image quality. At present, only the HUDF and its parallel fields are available to select very high redshift galaxies down to the faintest limits and high surface densities that NIRSpec can probe spectroscopically, AB ~ 29 mag for 10-30 hour integrations at R ~ 100. JWST will be more effective and efficient, particularly in its early cycles, if more fields are imaged to this depth in widely separated regions of the sky. NIRSpec will be able to confirm the redshifts of galaxies using continuum shape and breaks even during the epoch of reionization, when Lyman-alpha emission may be scattered by a modestly neutral intergalactic medium.

Cosmic variance: There are presently only three fields imaged to depths fainter than AB = 28.5; the HUDF and its two parallel fields. Everything we have deduced about the faint end of the high-redshift luminosity and mass functions are derived from these fields, which do not represent independent lines of sight, as they are very close on the sky. It is estimated that the 1-σ cosmic variance uncertainty in the faint luminosity function is ~30%. This is a tolerable level of error for some investigations, but it remains possible that these fields cover local density perturbations in the large-scale structure that are > 1-σ or more away from the cosmic mean (Robertson 2010). An increase in the area probed at this depth over multiple *independent* lines-of-sight would substantially increase the robustness of our results against cosmic variance.

Our consensus recommendation discussed in Section 4 provides a “light” version of Scenario 2 in the form of six blank fields imaged in parallel with six cluster fields. This plan will bring the number of fields at a depth of the HUDF09 parallels from 3 to 9 (compared to 3 to 13 for a full Scenario 2) and an associated factor of 3X gain in area (rather than 4.3X for Scenario 2). The lensing field science made possible by this tradeoff is significant.

3.2 Deep Fields Strongly Lensed by Massive Foreground Clusters

Clusters of galaxies are the most recently assembled, massive, bound structures in the universe. As predicted by general relativity (Einstein 1936), clusters deform space-time in their vicinity and thus deflect light rays traversing through them from distant sources in the background. The images of distant objects therefore appear distorted and magnified. Strong gravitational lensing causes giant arcs, multiple images, and arclets that are unique features of massive foreground objects such as galaxy clusters. The detection of multiple images of the same background object can be used to constrain the detailed mass distribution of clusters, and the accuracy of these models depends crucially on the number of multiple images identified. This requires high resolution imaging from space, and HST and ACS in particular has revolutionized the field of cluster lensing. New observations of strongly lensed cluster fields can address the following science topics:

Intrinsically Faint Galaxies at High Redshift: Given our charge to explore the high-redshift universe, cluster lenses are most interesting for their power as gravitational telescopes to unveil hitherto undetected faint galaxy populations at high redshift. Clusters as natural telescopes provide an avenue for detecting galaxies that are 10-50 times fainter than would otherwise be possible at $z > 5$. By probing far into the sub-L* regime, clusters enable the detection of the direct progenitors of Milky-Way size galaxies at $z \sim 9$ (at least as estimated in hierarchical merging scenarios) and provide constraints on the faint end slope of the luminosity function, with the associated assessment of the role these faint galaxies play in re-ionization. This enhanced access to faint galaxies is unique to gravitational lensing fields and cannot be practically achieved otherwise.

Magnified Internal Galactic Structure*:* Foreground clusters stretch and distend the images of background galaxies in addition to making them brighter. Since surface brightness is conserved in gravitational lensing, morphological features in these most distant galaxies come into sharper view. In some cases extreme magnifications allow us to probe internal structures on ~100 parsec scales - down to the size of individual giant molecular clouds and star-forming regions. Deep cluster images can unveil the structural properties of lower redshift galaxy populations (out to $z \sim 5$) for not just the brightest sources, as in blank fields, but over a much larger luminosity range owing to the additional magnification.

Cluster Structure and Cosmology: Deep images of galaxy clusters reveal the underlying substructure of the dark matter that dominates their mass. Such observations can address the very foundations of cosmology by constraining the formation and dynamical evolution of clusters and the nature of the dark matter. In concert with fields of background lensed objects, cluster images allow mapping of the clustering of dark matter within clusters; among the most compelling of recent results is the clear demonstration of the existence of dark matter, distinct from the hot X-ray emitting intracluster medium, in the Bullet Cluster. Clusters also constrain the geometry of the universe, as the strength of lensing depends on the ratios of angular diameter distances between the lens, source and observer, such that lens deflections are sensitive to the value of cosmological parameters and offer a powerful geometric tool to probe Dark Energy (Jullo et al. 2010). Finally, intracluster light at faint levels between cluster galaxies provides a means for studying the dynamical evolution and processing of galactic stars in interactions between galaxies.

Our consensus recommendation provides for six new deep fields with strong-lensing clusters in the foreground. These potentially revolutionary gravitational telescopes complement the blank fields, together providing the best balance of proven techniques and new opportunities to observe the highest-redshift galaxies.

3.3 Deep Fields with Grism Spectroscopy

As important as HST WFC3 and ACS imaging programs have been, they require redshift information in order to put faint galaxies onto a cosmic timeline and to study their evolution. Intensive campaigns of ground-based spectroscopy have measured thousands of redshifts in fields like GOODS, CANDELS and COSMOS, but are often limited to relatively bright magnitudes (AB ~ 25) compared to the depths of the HST imaging data, and are subject to significant redshift gaps and “deserts”, particularly when common optical emission lines shift into the near-infrared. Near-infrared multiobject spectrographs on 8-10m telescopes are now starting to rectify this, but will still have their limitations, particularly for continuum and emission line spectroscopy for faint galaxies without strong emission lines at accessible wavelengths. The limitations of ground-based spectroscopy require an unfortunately heavy reliance on photometric redshifts for deep field science. JWST will revolutionize this with deep infrared spectroscopy from NIRSpec and NIRISS, but HST can make a significant contribution now thanks to the slitless spectroscopy capabilities of the WFC3-IR camera.

WFC3 and ACS both have grisms that enable slitless spectroscopy. In this observing mode, every object within the imaging field of view is dispersed into a low-resolution spectrum ($R \approx 100$-200 for point sources; lower for extended objects), making it possible to measure accurate redshifts, emission line fluxes, and spectral continuum breaks and features. With the WFC3-IR channel, in particular, this is a powerful capability. Compared to ground-based multislit spectroscopy, the WFC3 grisms have several advantages, including (1) high multiplex, obtaining spectra for nearly every object within the field of view; (2) no target pre-selection – all objects are observed; (3) no atmospheric limits on wavelength coverage; (4) no slit losses – the full flux from all objects is measured; (5) very faint flux limits for continuum spectroscopy, much fainter than has been achieved with ground-based observations; (6) rather faint emission line flux limits, competitive with very deep ground-based observations; and (7) HST angular resolution, capable of spatially resolving galaxy structure – in particular, grism spectra effectively produce monochromatic emission line images (e.g., for Hα).

Several recent HST observing programs have obtained large amounts of relatively shallow WFC3 grism spectroscopy. Notably, 3D-HST (PI: van Dokkum) and AGHAST (PI: Weiner), considered together, have observed most of the area in the five CANDELS fields with 2-orbit depth G141 spectroscopy (1.1 to 1.65 ⌠m), while WISPS (PI: Malkan) is a pure parallel program that has obtained observations with both the G102 (0.8 to 1.1 ⌠m) and G141 grisms in many fields. These data address a wide variety of science topics:

Direct measures of SFR: Star formation rates can be measured from Hα at $z < 1.5$, as well as nebular reddening from the Balmer decrement. Emission line mapping can resolve the distribution of star formation within galaxies, while line ratios can measure nebular excitation properties to distinguish AGN from star-forming galaxies and to constrain metal abundances.

Galaxy Properties vs. Environment: Accurate redshifts enable more accurate studies of galaxy evolution, and in particular make it possible to correlate galaxy properties with local environmental density in a manner that is simply impossible using photometric redshifts.

Passive Galaxies: Continuum breaks and absorption lines (particularly the 4000A and Balmer breaks) can constrain stellar population properties such as ages and mass-to-light ratios, and measure redshifts for quiescent or "passive" galaxies out to z=3.

The grism observations of the CANDELS fields from the combined 3D-HST and AGHAST programs are expected to yield useful infrared spectra for up to 10000 galaxies, with a total investment of about 300 HST orbits. However, with only 2 orbit exposure times in G141, the useful data are limited to relatively bright galaxies. Specifically, for typical faint (extended) galaxies, these data reach S/N=5 per resolution element in the continuum S/N=5 at $H \approx 23.1$ (AB), and S/N=5 for an emission line flux of 5×10^{-17} erg s^{-1} cm^{-2} (e.g., Brammer et al. 2012). This continuum flux limit restricts the detection of spectral breaks (e.g., the 4000A break at z = 2 to 3) to relatively rare, very luminous galaxies (significantly brighter than L*). In principle, the WFC3 grisms could detect Lyα emission lines or the Lyα forest continuum break for galaxies at $6 < z < 12$. The current flux limits for emission lines are enabling good science now, but still restrict the analysis to rather strong line emitters. E.g., for Hα at z=1.5, this flux limit corresponds to a star formation rate of 5.7 M_o yr^{-1}, and considerably larger when the large extinction expected for dusty star-forming galaxies at these redshifts are taken into account. Detection of weaker lines, Hβ for example, are comparatively rare, limiting opportunities for measuring diagnostic ratios. Very few observing programs with substantially deeper grism observations have been carried out to date, although the first results from a few deep programs are starting to appear. Indeed, these demonstrate the possibilities for new science using deep grism spectra, e.g., measuring stellar population parameters using continuum breaks out to $z \approx 3$ (Gobat et al. 2012).

Three community white papers suggested deeper grism observations. Their proposals spanned a range of area-depth parameter space, from 8 to 100+ orbits per grism per pointing. All three white papers advocated using both the G102 and G141 grisms in order to expand the spectral range covered; in particular this would significantly increase the redshift ranges over which multiple emission lines can be detected, increasing redshift reliability and enhancing the opportunities for physical measurements based on line ratios. Some argued for observations a factor of ~2 or so deeper (in flux) compared to AGHAST/3D-HST, covering a relatively wide, GOODS-like area, and extending the wavelength coverage using G102. This would largely be aimed at science at $1 < z < 3$, where rest-frame optical lines fall in the grism spectral ranges, for a galaxy sample large enough to be statistically informative. There were also arguments for much deeper observations over correspondingly smaller areas, deep enough to push into a new redshift regime and detect the Lyα emission line or Lyα forest continuum breaks for galaxies at $z > 8$, a realm where current ground-based spectroscopy has yet to penetrate. Further developing the technical tools and hands-on scientific expertise for deep HST spectroscopy could also have programmatic benefits in preparing the astronomical community to use JWST, where faint object and/or slitless spectroscopy is expected to make up a significant component of the science program.

The HDFI committee appreciates the great scientific potential for deep grism observations. There appears to be more open parameter space for HST grism spectroscopy than there is for new blank field deep imaging programs, where nearly all parts of achievable area-depth-wavelength parameter space already have *some* data, often with very large past investments of observing time. However, in the end, the working group judged that most of the new information from such a grism program would be gained in the $1 < z < 3$ redshift regime, with only limited impact on the high redshift science that we were charged with pursuing. Even very deep observations would likely yield only a few Lyα emission line or break detections, particularly at genuinely new redshifts $z \gg 7$, given the small solid angles that could be surveyed. Ground-based multi-object near-infrared spectroscopy seems to be coming into its own now with instruments like Subaru/MOIRCS and FMOS, Keck/MOSFIRE, Magellan/MMIRS, LBT/LUCIFER and VLT/KMOS, and while HST slitless spectroscopy has certain advantages, some of this science will also be done with terrestrial facilities. We felt that the results now coming from WFC3 grism spectroscopy, and the enthusiasm that they are generating in the community, make it quite likely that significant new programs can be proposed successfully through the normal time allocation process. Because the main interesting science from such programs was at redshifts below our charge, we did not favor carrying out a grism program with this new Hubble deep field initiative.

3.4 Deep COS Spectroscopy with Imaging Parallels.

Imaging in parallel with spectroscopic observations using COS or STIS has become a common mode of using Hubble. Imaging in parallel can readily detect bright galaxies at $z \sim 7$-8 galaxies even in relatively short spectroscopic visits (Trenti et al. 2012), suggesting that deep spectroscopy would combine well with the science goals of deep blank imaging fields as presented above. Recognizing this opportunity, the committee evaluated the science case for deep UV spectroscopic observations of AGN and their outflows, Lyman continuum radiation escaping galaxies, and of hot gas and metals in the low-density IGM.

AGN Structure and Feedback: Two white papers advocated for measurements of the kinetic luminosity and reverberation-mapped velocity fields of local AGN using long-term monitoring campaigns with COS. These observations would sharply constrain the energy imparted by the AGN to its host galaxy and greatly refine the scaling factors that enable SMBH mass measurements at $z > 0.1$. Because they rely on the response of blue-shifted absorption to the varying outflow rate and radiation, these measurements require 1-5 orbits of COS observation over periods of 50-150 days, with details depending on exact choice of target and S/N requirements.

Lyman Continuum Escape and Cosmic Reionization: The fraction of ionizing photons that escape star-forming galaxies determines how effective they are at reionizing the IGM, while the fraction they retain influences the Lyman alpha “signpost” often used to detect them. Direct measurements of the Lyman continuum escape are practically impossible at $z > 3.5$ because of the increasing opacity of the IGM but is more feasible at low redshift with deep spectroscopy in the FUV. LyC escape fractions of 1-10% could be measured in the COS FUV modes on star-forming galaxies at $z \sim 0.1 - 0.3$ in 100 - 200 orbits per galaxy

Hot gas and Metals in the Low-Density IGM: The baryon budget at low redshift leaves approximately 1/3 of all baryons still unaccounted for (Shull et al. 2012). These missing baryons may lie hidden in gas at $T = 10^{5-6}$ K in IGM gas at 1-50 times the cosmic mean density. This gas far from galaxies may also hold metals ejected from star-forming galaxies over all of cosmic time; and if so, it would be a sensitive tracer of the ejective feedback from galaxies. Deep spectroscopic observations of a bright $z \sim 1$ QSO could achieve S/N = 80 - 100 (per pixel) in 150-200 orbits. These data would be unprecedented in their combination of depth and path-length for IGM science. They would enable the detection of hot baryons traced by high-ionized species such as O VI, Ne VIII, and Mg X if the WHIM gas is metal-enriched and by broad Lyα systems even if it is not. Detection of gas metallicities down to ~1% even in the under-dense IGM could sharply constrain the energy and mass lost in galactic feedback, a key uncertainty in current models of galaxy formation.

These “near-field” programs with COS visit single targets for 100-200 orbits and so are essentially compatible with the requirements of deep “blank-field” imaging observations as described above (blank field Scenarios 1 or 2). The scientific aims of these spectroscopic programs were well regarded and extensively discussed by the committee. However, because they require specific COS targets they offer a limited degree of flexibility in the choice of blank field location and are not at all compatible with using massive clusters as gravitational telescopes. The AGN observations also require specific observing cadences that place additional limitations on the depth and dithering of imaging observations. While these are not insurmountable issues, the committee judged that the number of strong lensing clusters required to effectively carry out the recommended science program would not allow for an additional program component to image a blank field in parallel with COS.

4. Vetting a Deep Lensing Component for the HDFI

4.1 Uncertainties in the Lensing Models

The use of galaxy clusters as gravitational telescopes is a particularly exciting possibility, but the interpretation of any results will depend crucially on the reliability of the lensing maps available to characterize individual targets chosen. As we discuss below, magnification uncertainties are high only when the magnification is high, but when the magnification is just a few, as is typical for the bulk of the objects that will be detected at high-redshift, the uncertainties are minimal and can be well characterized.

At the present time, there are many lensing groups actively working throughout the world and the technique is fairly mature, however as part of our deliberations we looked into how models generated for a given cluster by diverse techniques compared. We set about understanding the uncertainties and the implications thereof for achieving our science goals. Below we outline the process by which we compared and contrasted the map-making process by groups employing a range of tools. First, we compared lens maps of real clusters with *HST* data produced by different groups (independent groups using their respective codes) in order to gauge differences. Second, we ensured that lensing magnification maps made available had the accuracy necessary to use them to probe the luminosity function.

There are essentially 3 kinds of techniques for generating maps: parametric, non-parametric, and hybrid methods. Different groups have adopted one or the other of these and have developed their own tools to construct mass maps. As part of this evaluation process, this group coordinated magnification map comparisons among 3 different groups, first for the cluster Abell 1689 and second for the Bullet Cluster. The groups we brought in for comparison were a) Jullo et al., who use a hybrid approach including strong + weak lensing; b) Zitrin et al., who support the CLASH project and use a parametric tool including only the strong lensing constraints; and c) Bradac et al., who use a non-parametric method that combines observed strong + weak lensing as inputs. Figure 1 shows estimated errors for Abell 1689 (left) and the Bullet Cluster (right).

Errors were estimated by comparing pairs of lensing maps. Two maps were subtracted to derive an estimate for an error in magnification for the magnified background lensed source that would be observed behind the cluster. The y-axis in the top panel is roughly the fractional difference in linear magnification μ that would be inferred from the 2 maps. For example at $\mu = 5$, the two independent maps would differ by 20% in flux -- comparable to typical uncertainties in photometry and distance modulus for dropouts. The lower panel shows the cumulative distribution function of the magnification distribution of the average magnification map. Over 70% of the solid angle, and therefore volume, the magnification factor is ~5 or less, where fractional uncertainties in magnification are below 0.2. *Only 20% of the area has magnification >25 where*

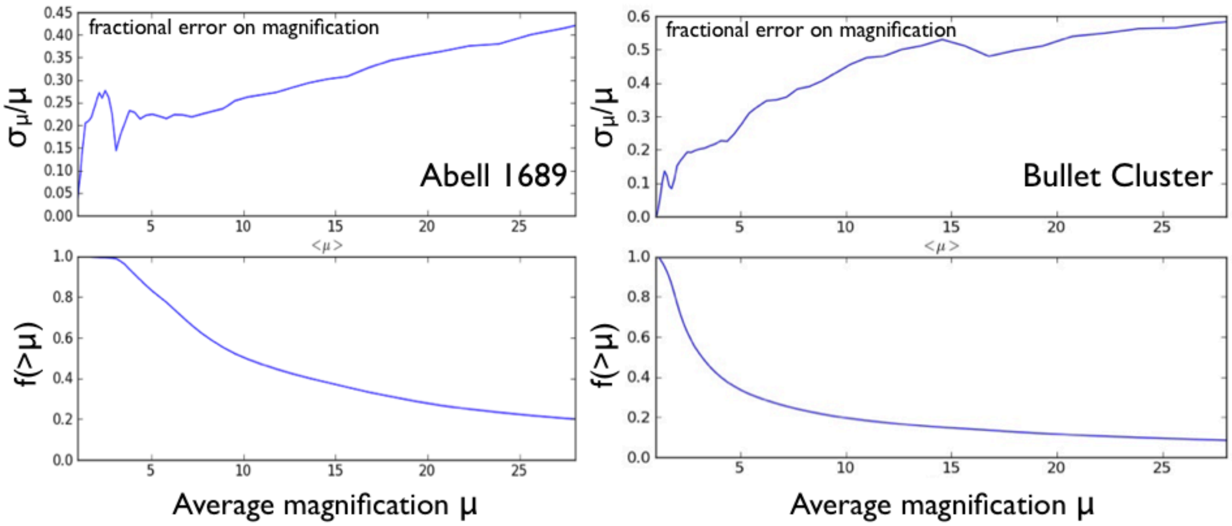

Figure 1: **Vetting lensing as a well-understood tool**. Shown in the upper panels are fractional errors on magnifications as estimated for real galaxy clusters with HST imaging by comparing two different lensing techniques from two different groups. Lower panels show the cumulative distribution of magnifications for each cluster (Abell 1689, left; the Bullet Cluster, right). Note that the lensing maps agree well for typical average magnifications. For example, at magnifications of ~5, each of the two independent maps would differ by ~20% in flux, comparable to typical uncertainties in photometry and distance modulus for dropouts. Uncertainties grow for the highest magnifications.

uncertainties are of order 40%. Note also that in case a multiply imaged dropout is found near a critical line (technically the region with infinite magnification) their positions will constrain locally the mass model, thus reducing the uncertainty on its magnification to below 10-20%. The identification of and verification of a handful of multiply imaged systems will significantly decrease the errors and improve the quality of the magnification maps.

4.2 Complementarities with and Differentiation from CLASH

CLASH (Postman et al. 2012) is an on-going shallow survey of 25 massive clusters that exploits the panchromatic imaging capabilities of HST. The CLASH cluster sample was selected with three key considerations: (1) clusters are dynamically relaxed, (2) clusters are massive (T_x >5 keV), and (3) clusters are sufficiently numerous to overcome scatter in mass profile determinations. Five were selected for high magnification.

The HDFI cluster-lensing program specifically selects targets that are the most efficient lenses for high-redshift background targets, with critical curves for sources at z ~ 10 that lie within a single field of view of WFC3 and ACS. The median redshift of the HDFI candidate clusters is z = 0.5, significantly higher than that for CLASH (median z = 0.35). The HDFI program is also significantly deeper than CLASH, which allows for the detection of intrinsically fainter systems (typical magnifications for these sources are factors of ~10). As the volume for the very high magnification tail falls off fast, it is hard to find intrinsically faint high-z galaxies by looking at a lot of shallow clusters, the strategy adopted by CLASH. Of course, CLASH will uncover the rare, extremely bright high redshift sources, for instance the recently reported detection of a multiply imaged galaxy with photometric redshift 10.7 (Coe et al. 2012). Deeper imaging should also uncover many more multiply imaged sources, which can be used to improve the accuracy of the lensing models and hence the calibration of these gravitational telescopes.

Table 1 - Comparison between HDFI and CLASH depth per filter. Shown are the number of orbits and integration time associated with each filter for both projects. The right most columns show the integration time ratio and magnitude differences for each filter. Note that magnitude differences are maximal (1.82 and 1.73) for the “dropout filters”, which are typically the limiting factor in deep Lyman-break searches.

Filter	HDFI		CLASH			
	orbits	**ksec**	**orbits**	**ksec**	**ratio**	**delta(mag)**
F160W	23	66.7	2	4.9	13.6	1.42
F140W	11	31.9	1	2.3	13.6	1.42
F125W	12	34.8	1	2.5	14.4	1.45
F105W	24	69.6	1	2.4	28.8	1.82
F814W	43	100.0	2	4.1	24.4	1.73
F606W	10	25.0	1	2.0	12.7	1.38
F435W	17	42.5	1	2.0	21.4	1.66

4.3 Effects of Clusters on Blank Parallel Fields

It is important to estimate the effects of magnification due to the clusters at the location of the accompanying blank fields. Ideally, these effects should be small, in order to ensure that galaxy magnitudes and number counts in the blank parallel fields can be easily compared with those in other surveys like the HUDF and CANDELS.

The separation of the WFC3 and ACS fields in the HST focal plane is 6.0 arcminutes. This corresponds to 2 to 3 virial radii at the redshift range of the clusters considered here. From detailed studies of cluster lensing done so far, given the inferred profiles of dark matter, it is found that the magnification drops off rapidly (as 1/r). Typical magnification values for candidate clusters on our list 6 arcmin from the center are approximately 1.01 - 1.03, i.e., a 1 to 3% boost in flux. The weak lensing distortion in shapes of galaxies, shear, induced by the cluster in the blank field is expected to be 1% to 3% depending of course on the source redshift. This only modestly exceeds the mean cosmic value of ~1%.

Galaxies in the outskirts of the clusters, approximately 2 Mpc from the cluster center at typical redshifts z=0.5, will be present in the parallel fields. It seems quite likely that these can easily be distinguished from high redshift background galaxies based on their colors and photometric redshifts using the deep 6- to 7-band HST imaging, and are unlikely to cause significant confusion for most deep field studies.

5. Consensus Recommendations

After weighing the main alternatives, the HDFI Committee is offering a unanimous recommendation to the STScI Director. The recommended program combines 6 new deep fields centered on strong-lensing clusters, in parallel with six HUDF-like blank fields. We describe a nominal observing program of 140 orbits per cluster in 6-7 photometric bands with ACS and WFC3, for a total of 840 orbits.

Our combined recommended program addresses the following key science aims:

- Reveal populations z=5-10 galaxies that are 10-50 times fainter than any presently known, the key building blocks of ~L_* galaxies in the local universe.
- Characterize the stellar populations of faint galaxies at high redshift and solidify our understanding of the stellar mass function at the earliest times.
- Provide, for the first time, a statistical morphological characterization of star forming galaxies at $z > 5$.
- Find $z > 8$ galaxies stretched out enough by foreground clusters to measure sizes and internal structure and/or magnified enough for spectroscopic follow up.

5.1 The Dual Power of Gravitational Telescopes and Blank Fields

Through a considerable investment of resources and talent, HST has pushed the high redshift frontier beyond z=8, finding galaxies that existed just ~500Myr after the Big Bang. While remarkable in their own right, these detections may be of limited value for understanding how typical galaxies form. For example, the current detections represent the brightest galaxies that existed at z>8. Taking into account the mergers expected in hierarchical cosmologies, these systems are too rare to be the main progenitors of typical ~L* galaxies today. Detections some two magnitudes deeper than the current limits are likely required to find galaxies that are reasonable progenitor candidates for a galaxy like the Milky Way.

Even at slightly lower redshifts, where detections of fainter galaxies are somewhat more plentiful, there is a relative paucity of systems that are bright enough to obtain information on stellar populations, sizes, and morphologies. Moreover, precise redshifts from spectroscopy are only available for the brightest of the bright z>7 galaxies. The next great advances in galaxy evolution at early times require wider survey areas at deep limits for improved statistical samples, and a reach to galaxies that are intrinsically much fainter and thus more like the progenitors of modern normal galaxies. **Going deeper and going wider are usually in tension with one another in any competition for limited resources; a program that combines both approaches can significantly increase our understanding of how galaxies came to be**. By combining optical/infrared observations of two parallel fields with the natural gravitational telescopes first envisioned by Einstein in 1936, Hubble is poised to take the next great leap in depth and area in the service of these important astrophysical goals. As shown in Figure 2, strong cluster lensing fields will reach galaxies more than two magnitudes fainter than otherwise possible, while adjoining blank fields will solidify

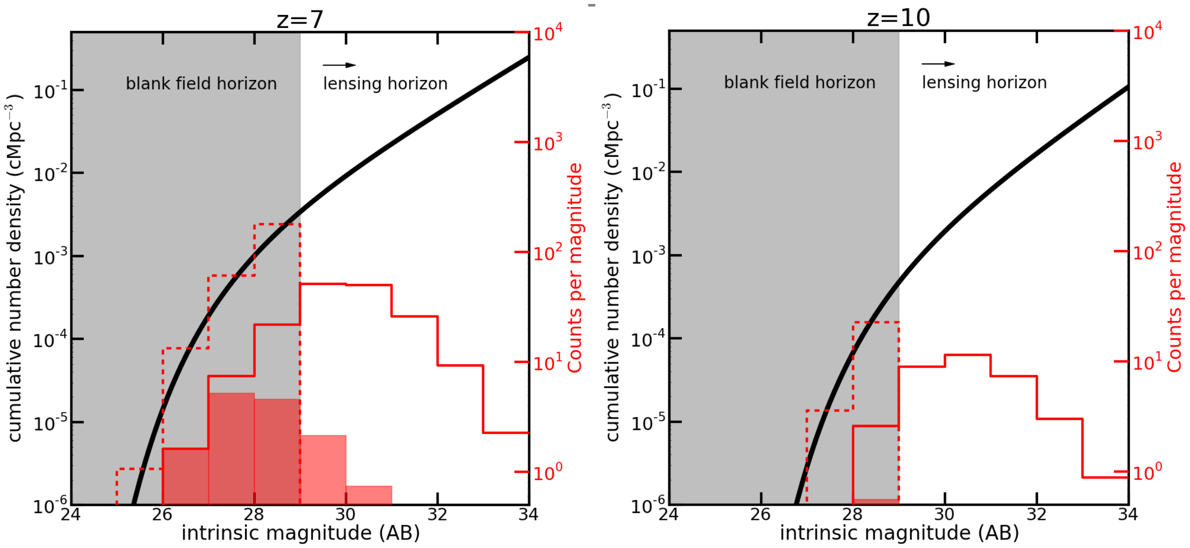

Figure 2: **Dual power of cluster telescopes plus blank fields:** The black solid line (corresponding to the left vertical axis) shows the underlying cumulative UV luminosity function we assume at z=7 (left) and z=10 (right), based on Schechter parameters that smoothly evolve with redshift in a manner consistent with recent measurements at $4 < z < 8$ (e.g., Bouwens et al. 2012). The dashed red histogram gives counts per magnitude for blank fields (right vertical axis), which cut off at the "blank field horizon" for six WFC3 areas integrated down to an example AB = 29 (the right axis minimum is set so that the dashed red line tracks the solid black line at the bright end). **The solid red histogram shows the expected counts per magnitude for six lensed fields to the same depth**, as determined from existing lensing maps for one of our typical candidate clusters. The blank fields provide high counts while the cluster fields allow for detections more than two magnitudes deeper. The shaded red histogram marks those lensed galaxies that are bright enough in apparent magnitude for possible spectroscopic followup (AB < 26.5). The practical limit for JWST/NIRCam blank fields is AB ~ 30-31.

uncertain statistics on galaxy luminosity functions, mass functions, and structural properties in the high-redshift universe over a larger area than has been already surveyed.

5.2 The Plan: Twelve New Fields to Unprecedented Area and Depth

Our recommended program comprises twelve new "deep fields": six pointed at massive clusters and six nearby "blank fields" (centered 6 arcmin apart). Relying on a beam-switching technique with ACS and WFC3, the combination of new blank field data and lensed galaxy data will solidify our understanding of the faint, high-z luminosity function. Our nominal observing program is specified in Tables 2-4.

Table 2: ACS Imaging

Filter	# Orbits	Exptime (ksec)	Total Mag (AB)	SNR	Fraction of PSF	Aper Corr (mag)	Total mag_{corr} (AB)
F435W	17	42.5	29.10	4.98	0.85	-0.34	28.76
F606W	10	25.0	29.12	5.02	0.85	-0.34	28.78
F814W	43	100	29.46	5.01	0.85	-0.34	29.12

Table 3: WFC3 Imaging

Filter	# Orbits	Exptime (ksec)	Total Mag (AB)	SNR	Fraction of PSF	Aper Corr (mag)	Total mag_{corr} (AB)
F105W	29	84.1	29.30	4.98	0.75	-0.21	29.09
F125W	14	40.6	28.79	5.01	0.70	-0.13	28.66
F160W	27	78.3	28.79	5.02	0.62	---	28.79

Table 4: WFC3 Imaging (cluster field option)

Filter	# Orbits	Exptime (ksec)	Total Mag (AB)	SNR	Fraction of PSF	Aper Corr (mag)	Total mag_{corr} (AB)
F105W	24	69.6	29.19	5.02	0.75	-0.21	28.98
F125W	12	34.8	28.70	5.04	0.70	-0.13	28.57
F140W	11	31.9	28.72	5.03	0.66	-0.07	28.65
F160W	23	66.7	28.70	5.03	0.62	---	28.70

Tables 2-4 Nominal Observing Program: These sensitivity estimates were obtained via the ACS and WFC3 exposure time calculators (ETCs). We assumed that a faint object near the detection limit is a point source with a spectrum that is flat in f_nu, scaled to the magnitude in the 4th column. We assumed low zodiacal light and average Earthshine. We assumed that each target will be split into 2-orbit visits composed of four total exposures, of t_{exp}=2500 sec/orbit for ACS, and 2900 sec/orbit for WFC3. The orbit numbers yield the required exposure time to detect a galaxy with a total magnitude as listed in column 4 at ~5σ significance, with the signal-to-noise measured in a 0.4” diameter aperture (optimal for z > 6 galaxies). The last column gives the corrected total magnitude limit in each band, accounting for the varying fraction of the PSF flux that is included within the measurement aperture. This is effectively the 5-sigma total magnitude limit for each band accounting for PSF-matching to the resolution of the F160W image.

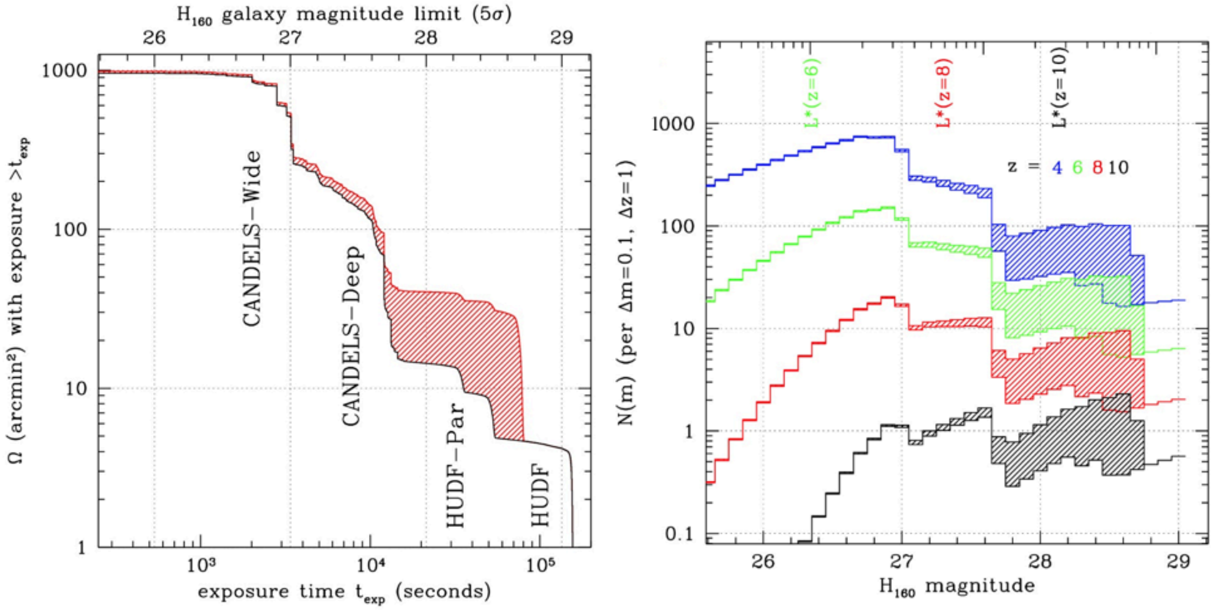

Figure 3: **Blank Field Gain** – (Left) The black line shows the cumulative HST area surveyed vs exposure depth in F160W for all of CANDLES and HUDF09, including parallels. Our proposed 6 HDFI blank fields program is indicated in shaded red, filling the relative void in HUDF-Parallel depth. (Right) Number of galaxies that could in principle be detected from the same set of fields, with the shaded regions showing the gain expected from our 6 HDFI blank fields program, providing a factor of ~3-5 gain in faint galaxies compared to current counts, and roughly doubling the number of total z~8-10 candidates. The calculations assume UV luminosity function Schechter parameters that smoothly evolve with redshift in a manner consistent with recent measurements at $4 < z < 8$ (e.g., Bouwens et al. 2012) and extrapolated to z=10, and 100% selection efficiency to the 5-sigma detection limit. Note: the lensing fields (Figure 1) will push detections more than two magnitudes fainter than the edge of this figure.

Our choice of depth and number of fields complements naturally the existing area and depth coverage for blank fields *and* cluster lensing fields. The blank fields are deeper /narrower than CANDLES-Deep and wider/shallower than the HUDF (see Figure 3). Our cluster fields are deeper but less numerous than CLASH. This will allow for the detection and characterization of intrinsically faint high-z galaxies that, importantly, have moderate (and therefore well-understood) magnifications. Because they are shallower, CLASH data enable confident explorations of only the brightest high-z objects (comparable in reach to the HUDF, but with somewhat larger errors because of lensing uncertainties). The six deep cluster fields promise to deliver ~10 sub-L_* galaxies at $z \sim 10$ and up to 100 at $z \sim 8$. The six blank fields will increase the number of sub-L_* $z > 7$ galaxies by a factor of 3-5 and will reduce the effect of cosmic variance on luminosity and mass functions by roughly a factor of two. Given natural statistical fluctuations, these samples cannot be achieved with a significantly smaller number of fields.

As our proposed program has a goal to image 6 clusters over < 1000 orbits, we will not be able to substantially increase the depth at $m > 29$, which is the domain of the HUDF.

However, examining Figure 3, the current state of near-infrared blank-field imaging shows a substantial drop-off at $m > 27.6$ (the depth of the CANDELS Deep surveys). There is a clear parameter space at $28 < m < 29$ magnitudes that would gain substantially from additional surveyed area. L_* galaxies at z=9-10 are expected to fall in this magnitude range. The present contribution in this regime comes from the HUDF09 parallels. These two fields have a total magnitude H-band depth of ~28.3 mag (28.8 in a 0.4" diameter aperture). As shown in Table 2, in 27 orbits per field, we can achieve a total H-band depth of ~28.8 mag, ~0.5 mag deeper than the parallel fields. As the HUDF is the only current field with imaging of this depth, we will be increasing the total area surveyed to this depth by a factor of 7. Additionally, at the depth of the HUDF09 parallels, we will be increasing the total area surveyed by a factor of 3 (3 HUDF09 fields + 6 new HDFI fields = 9 total fields). The added 0.5 mag in depth is crucial to achieve our primary science goals, which center around the faint-end of the luminosity and stellar mass functions. The HUDF09 parallel depth is not sufficient to meet these goals.

Given the total number of orbits (< 1000), as well as the necessary depth in the H-band (27 orbits), we have designed a starting-point plan for the implementation team as shown in Table 2. We set the WFC3 exposure times such that the depths in the F125W and F160W were roughly equivalent, and that the depth in the F105W band was 0.4-0.5 magnitudes deeper. As a substantial goal for our program is the identification and characterization of galaxies at $z > 8$, very deep F105W imaging is necessary for a robust detection of the Lyman break, as a galaxy at z~8 will have its Lyman break between the F105W and F125W filters. Typical Lyman-break galaxy (LBG) selection planes required a Lyman break of at least 0.5 magnitudes (commonly > 1 magnitude) - if the F105W depth was equivalent to F125W/F160W, it would thus render the deepest 0.5 mag in those redder filters unusable. The total number of orbits used is 70. We estimated that we would require another 70 orbits with ACS, for a total of 140 orbits per pointing. We note that while one could imagine using fewer orbits in the ACS bands, it would result in fewer WFC3 orbits in the cluster, which would compromise our primary science goals. In the ACS bands, we do a similar split where we spend much of the time imaging in F814W, which will be the dropout filter for z~7 galaxies, and we split the remaining depth between F435W and F606W (which will both allow the discovery of z=4 - 6 galaxies, as well as helping to promote the fidelity of the $z > 7$ candidates via multiple non-detections).

We feel that 140 orbits is the minimum that can be allocated per cluster and associated parallel field. Shallower observations would narrow the difference between the HDFI cluster program and the existing CLASH data, weakening the primary scientific objective of our science program. Shallower blank fields would also be more similar to the existing HUDF09 parallel fields, substantially reducing the impact of this new program in the area-depth plane (Figure 3), and compromising our ability to improve the luminosity and stellar mass functions.

5.3 JWST with HST

Using clusters as telescopes is tantamount to a 10-50x increase in HST's aperture. This major stride in sensitivity will allow us not only to discover but also to examine in detail the magnified background populations of galaxies – objects that are hitherto little understood or entirely unknown because they are out of reach in even the deepest practical blank fields.

Figure 2 compares the "blank field horizon" to the "lensing horizon". The magnification from the foreground cluster reaches galaxies that are 10 to 50 times fainter intrinsically than the limits of blank fields. HST/CLASH has already detected two lensed galaxies with photometric redshifts indicative of $z = 10\text{-}11$, proving the basic technique (though CLASH is much shallower than our proposed program, thus probing intrinsically brighter galaxies than the ones that will be detected in the HDFI data). The 6 blank fields we contemplate will be commensurate in depth and two times larger in area than the existing HUDF and its parallel fields - a big gain in area and nearly as deep as HST can go. Using clusters to magnify half of the fields will reach galaxies that are intrinsically fainter than those in the HUDF by 2-4 magnitudes - even beyond the AB = 30-31 limits that JWST/NIRCam can achieve in its own blank fields. That is, **Hubble and gravitational telescopes may very well reach further down into the small building blocks of galaxies than JWST can by itself in a blank field.** Thus Hubble will get an unprecedented glimpse into the first Gyr; we need not await JWST to probe the end of the cosmic Dark Ages.

At high redshifts ($z > 7$) cluster lenses bring into view galaxies that are inaccessible to any other current facilities in space or on the ground. And at lower redshifts ($z < 7$) this technique offers a major stride step beyond merely counting galaxies. These data will push back the faint frontier by enabling studies of the structural properties (sizes, colors, stellar masses) of these highly magnified sources that are intrinsically 2 - 4 fainter than L_* - a regime that is also currently unexplored at $z > 6$. Lensing magnification will significantly increase the number of galaxies detected at AB < 26, the limits of ground-based and HST grism spectroscopy, which will measure the precise redshifts, star formation rates, metallicities, and internal structure of these galaxies. By pushing back all three high-redshift frontiers, this program will yield an unprecedented data set that will inevitably lead to major strides in our understanding of early galaxy formation.

5.4 Legacy Extragalactic Fields

The proposed program invests considerable HST resources in twelve relatively new 'legacy' fields in six locations on the sky. This committee discussed in detail the pros and cons of defining new fields vs. revisiting the existing legacy HST fields (e.g. GOODS-N/HDF). The existing legacy fields have the advantage of many years of study, and a great deal of additional data from other space observatories (Spitzer, Herschel, XMM, Chandra, GALEX) and many ground-based observing campaigns. Such supporting data are important for both the study of the high redshift universe as well as ancillary science studies, and many white papers suggested strategies for deep follow-up with Chandra, XMM, Herschel, ALMA, and SKA/ASKAP.

Therefore STScI is strongly encouraged to engage other observatories in coordinated observations of these new fields in order to increase their scientific productivity. Deep Spitzer IRAC observations were identified as crucial for understanding the high-redshift galaxy populations. ALMA imaging and spectroscopy may also provide powerful constraints on the properties and redshifts of high-redshift galaxies. The most distant galaxies seen by HST in these new deep fields are unlikely to be detected in deep X-ray, UV and far-IR observations. However, multiwavelength data have proven invaluable for the study of the $z<5$ universe and will greatly increase the legacy value of these new fields. In particular, the HDFI cluster fields will become the best-calibrated and most intensively studied gravitational telescopes anywhere in the sky, and will enable new science to be done with multiwavelength data on lensed galaxies at all background redshifts.

Appendix 1 includes additional discussion of the potential importance of multiwavelength observations of these fields from other facilities, as well as considerations for issues of Galactic and zodiacal foregrounds that may be relevant for multiwavelength observations (e.g., Galactic cirrus, neutral hydrogen column density, nearby bright radio sources, etc.).

6. Suggestions for Implementation

Our committee was tasked with identifying a science case for new deep fields, for specifying their number and location and the basic parameters of the deep exposures that can accomplish the science goals. The recommended number of new fields and their depth are critical to the overall success of the science. However, there are many practical details of implementation, including the exact choice of cluster fields that depend on practical factors beyond the scope of our work. A program of this scope and size will require years of complex effort to plan and execute, so we expect that STScI will convene a group of technical experts to implement the science plan and guide it to completion. This group should assume final responsibility for the exact choice of clusters and the detailed definition of exposure times, accounting for dithering and scheduling issues while adhering as closely as practical to the recommendations above.

6.1 Lensing considerations for optimizing new HST cluster deep fields

Clusters that HDFI targets need to be already known as strong and efficient lenses with extant robust lensing models with current data in hand. We have arrived at a list of ~ 16 optimal cluster candidates that are the optimal targets from which to down-select. This list was developed with the following critical constraints in mind. The most massive clusters are the most efficient lenses in general, therefore the bulk of the candidates are drawn from the MAssive Cluster Survey catalog (MACS; Ebeling et al. 2001). Given our specific science goal to uncover the z > 7 galaxy populations that lensing offers privileged and singular access to, the list comprises clusters at z > 0.2 due to the dependence of the lensing strength on the angular diameter distance ratios from the observer to the lens to the source. The following scientific criteria were used to determine inclusion in our target list:

(i) Clusters are massive and among the most efficient lenses known,

(ii) Clusters have several sets of known multiple image systems confirmed with spectroscopic redshift measurements,

(iii) The majority of these clusters have high-quality current magnification maps available with data in hand that have been published by several independent groups, although these independent maps have yet to be cross-checked and calibrated against one another (see Section 6.4),

(iv) Based on the preliminary magnification maps, measured within the WFC3-IR field of view, the distribution of magnification values versus source-plane area at high redshifts (z>4) should be weighted toward large magnification values. This will ensure a maximal effect on the number counts of lensed, high-redshift sources. Excellent lensing clusters like RXJ 1346.5-1144 (which the working group analyzed in some detail) produce average magnifications >1.4 mag over the WFC3-IR field of view at z>4,

(v) The area that is enclosed interior to the critical curve for z~4-10 sources can be fully captured within a single pointing with WFC3 and ACS,

(vi) The combination of the constraints above implies an optimal redshift range typically above z > 0.37 for the candidate clusters except for one case, a cluster at z = 0.3. An exception is made for this cluster as the Einstein radius is particularly compact and is only 30 arcseconds (i.e. small so it will fit on within the WFC3-IR field of view). Higher redshift clusters also have the advantage of having somewhat less likelihood of blocking background sources by cluster galaxies themselves and have fainter associated intraculster light.

The positions, brightnesses and shapes of identified and confirmed multiple images are used as key input constraints to build mass models for lensing cluster. For cluster candidates to be considered as targets, we require several sets of multiple image systems to be already known as it confirms the efficacy of the lens and the accuracy of the magnification map and the mass model depend entirely on the availability of these as input constraints. Previous HST mapping, albeit with shallow exposures, is needed to determine the potency of the cluster lens. With the proposed survey strategy as more data is taken the lens model can and will be iteratively improved. We require that the entire critical region for sources at z ~ 10 must fit within a single WFC3-IR field of view, in order to ensure identification of all multiply imaged systems. Given the depth of the images proposed here, we expect to find 30 to 50 new multiple image systems for each cluster target, which will in turn lead to significant improvements in the accuracy of the lens models. Counter-images in multiple systems typically lie at the diametrically opposite end of the critical curve, so it is essential to image the entire critical region with HST.

The target list that is optimized solely on the basis of lensing criteria includes 16 clusters. Figure 4 shows the expected relative number counts expected for z = 10 sources behind a representative set of 11 out of 16 clusters. This figure is included to illustrate example-ranking criteria. The numbers are computed relative to the most efficient cluster in the sample. The strongest lens in our candidate list (MACS J0717.5+3745) is too far north to be observed with ALMA. However, given its excellent lensing characteristics, we recommend considering its inclusion despite this handicap.

We did not rank order the target list as this can only be done after incorporating the non-lensing constraints that will inform the choice and even to rank order solely on the basis on lensing strength, models will first need to be calibrated. We strongly recommend that the model calibration part of the data challenge be undertaken with the highest priority in order to appropriately rank order the target list before the final selection is made.

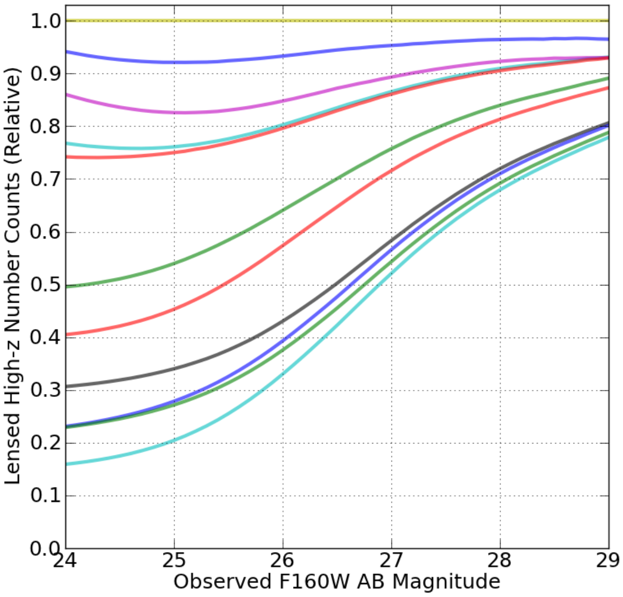

Figure 4: **The relative number versus magnitude of lensed high redshift galaxy number** counts behind the cluster lenses for illustration: these are a proxy for their lensing strengths, shown here for 11 out of the 16 candidate clusters on the list. The number counts are calibrated relative to the strongest lens on the list. The magnification maps for these 11 clusters was provided by Johan Richard and were derived combining strong and weak lensing HST observations using a hybrid modeling technique. This plot was generated with Johan Richard's maps independently by Dan Coe.

6.2 Non-lensing considerations for optimizing new HST cluster deep fields

In addition to the magnification potential for lensing clusters, many other considerations should be taken into account when selecting fields for the new HST deep field initiative. Here, we review the most important of these considerations, and make specific recommendations, prioritizations, and suggestions for further work by the implementation team. We highlight five general categories of issues in order of *decreasing importance* as judged by the working group. Detailed remarks and specific recommendations regarding these issues, as well as others not summarized here, are presented in Appendix 1.

It should be noted that it is difficult or impossible to meet all of these constraints simultaneously. In particular, the simultaneous preferences for high ecliptic and Galactic latitudes and low declination are nearly mutually exclusive, and it is inevitable that these will have to be relaxed, or not met for all of the targeted fields.

Ecliptic latitude and zodiacal foregrounds: Zodiacal foreground brightness can easily have a larger effect on the achieved signal-to-noise ratio for faint galaxy observations than other factors such as Galactic extinction. Ecliptic latitudes > 45° (absolute value) ensure low zodiacal foregrounds and have significantly enhanced visibility and schedulability with HST and JWST. Losses in sensitivity at low ecliptic latitude can be recovered with longer exposure times and more careful, restrictive observation scheduling, but it is important to take ecliptic considerations very seriously when choosing cluster targets.

Declination and visibility from ground-based facilities: Equatorial targets are visible from ground-based facilities in both hemispheres. Fields with declinations from $-30° < \delta < +25°$ would be accessible with telescopes from Cerro Paranal and Mauna Kea. 8-10m class OIR telescopes are available in both hemispheres to observe these fields, and at present the 20-40m class telescope projects are also ongoing in both hemispheres, but some facilities like ALMA and the VLA offer unique capabilities in single hemispheres. Among these, ALMA is arguably most important for this initiative, as it may offer the best opportunity for spectroscopic confirmation and study of deep field galaxies at $z > 7$, particularly through observations of far-infrared atomic lines such as [CII] 158 μm. There are also uniquely interesting scientific programs enabled by the combination of deep HST and ALMA observations of lensed galaxies at less extreme redshifts. For this reason, fields in the south should be preferred over northern fields that are inaccessible to ALMA. Additionally, two out of three planned 20-40m telescope projects are located in the southern hemisphere.

Galactic latitude, extinction, and foreground emission: Although zodiacal foreground emission can have a larger impact on signal-to-noise ratios for faint galaxies than Galactic extinction, it is still highly preferable to observe fields at high Galactic latitudes. Foreground extinction cannot be easily measured on scales of arcminutes, and if the expected Galactic extinction is large then the net uncertainty on faint galaxy luminosities and reddening corrections to galaxy colors can be significant. Far- and mid-infrared emission from Galactic dust is also an important consideration for observations from Herschel, SPICA and JWST, and the Galactic hydrogen column density is also a factor for deep X-ray observations of these cluster fields. Generically, Galactic latitudes $|b| > 40°$ would be preferred, with foreground extinction $E(B-V) < 0.025$, IRAS 100 micron foreground intensity < 1.3 MJy/sr, and neutral hydrogen foreground column $N(HI) < 2 \times 10^{20}\ \mathrm{cm}^{-2}$

Spitzer IRAC: Among current observing facilities, the Spitzer IRAC camera arguably offers the most immediately valuable data to support science from both the cluster and parallel blank fields, providing rest-frame optical flux measurements that are essential for evaluating stellar population masses and ages for galaxies at very high redshift. These will be quickly superseded by JWST NIRCAM observations, but for the next few years IRAC data will significantly enhance the scientific value of the HST observations from this initiative. An approved Spitzer Exploration Science Program will provide very deep IRAC observations of several of the top-priority HST cluster fields, and the Spitzer Science Center director has expressed willingness to invest discretionary time to support the HST deep field program. Very careful effort is required to ensure proper synchronization of the HST and Spitzer observations, particularly regarding telescope orientation for the cluster and parallel fields.

Other considerations: Far-infrared observations from Herschel can enhance the scientific output from these cluster fields, probing the bolometrically dominant star formation from sub-L* dusty galaxies at intermediate redshifts ($z \approx 2$) and measuring the currently unknown contribution from dusty star formation at $z > 2$. Most of the famous lensing clusters that the working group has evaluated as candidates for this deep field program already have moderately deep Herschel PACS and SPIRE observations, but in a few cases only relatively shallow SPIRE data are available. When the final cluster targets have been chosen the availability of Herschel data should be reconfirmed, and if necessary STScI should encourage ESA to schedule last-minute Herschel observations of the cluster fields. (The blank parallel fields are less important or less unique in this regard.) In turn, Herschel observations (or future far-infrared observations from SPICA) greatly benefit from the availability of very deep 24 micron observations from Spitzer.

Most of the cluster fields have deep X-ray observations from Chandra, XMM, or both. However, there may well be arguments for deeper observations to measure more accurate X-ray temperatures or emission substructure. Observers from the X-ray community may propose such observations; fields with low Galactic foreground neutral hydrogen column density would be preferred in this respect.

6.3 Community Support and Involvement

These observations are designed to provide a legacy dataset for the extra-galactic community for many years to come. We expect that many significant science discoveries will be made using these data, ranging from high redshift universe to understanding the nature of the clusters themselves. In order to encourage quick turnaround time for these discoveries and planning for follow-up observations with HST and other facilities, the raw data and high-quality data products, primarily the magnification maps should be released in a timely manner and regularly, perhaps via planned data released every 6 months.

It is imperative that accurate well-calibrated magnification maps are provided when the data is made public, as the majority of the extra-galactic community is not accustomed to using clusters as telescopes. In order to enable high-quality science by a diverse set of scientists, we recommend that additional data products and/or tools be provided simultaneously with the Hubble data. In particular, magnification maps for each selected cluster are needed in order to understand the likely locations, intrinsic luminosities and sizes of background-lensed objects. Additionally, coordinated deep Spitzer observations of the clusters will be challenging to match to high-spatial resolution HST images.

This recommendation presents unique data challenges that have not been encountered before. The primary challenge of course arises from the need to provide an early set of magnification maps produced for the selected cluster candidates from existing lensing data. The magnification maps for each of the cluster targets needs to be made publicly available to the community in order to enable non-lensing specialists to be able to take full advantage of this incredible data-set. We refer to this initial set of lens magnification maps as the first repository maps. As data is accumulated, these first repository maps will need to be updated and refined and this can be timed to occur during the set of data releases. The key challenge however is going to be the provision of maps that are derived from at least 3 different techniques adopted by modelers along with robust error bars. This is the reason that the calibration part of the data management has to be undertaken with enormous care.

6.4 Data Products and Tools for the Cluster Lensing Fields

Prior to setting up a data challenge, it is clear from the work of this committee and its deliberations that the different modeling techniques used need to be cross-calibrated. We feel strongly that this is the first step that needs to be undertaken as part of the data challenge. To our knowledge comparison beyond what we looked at for the Bullet Cluster and Abell 1689 during the course of our discussions have not been done prior. This is the first time that these disparate modeling techniques have been co-evaluated. There are essentially 3 kinds of techniques: parametric, non-parametric and hybrid methods. Different groups have adopted one or the other of these and have developed their own tools to construct mass maps.

Step 1: calibrating cluster modeling. We recommend the setting up a challenge inviting modeling groups that employ one of the above 3 techniques to reconstruct the following: 2 observed HST clusters + 1 simulated one with a training set for Abell 1689. This step needs to happen **first** to enable careful calibration of the methods. The HST cluster candidates with multiple images that have been spectroscopically followed up that we can choose from to present in this challenge are Abell 370, RXJ1347.5-1144 and MACS1149.5+2223 (incidentally all three of them appear in the top candidate target list for HDFI) + 1 simulated cluster. For the chosen clusters, the appropriate HST image + catalog of known multiple images with redshifts (both photometric redshifts and spectroscopic redshifts) and their spatial positions needs to be made available. Selecting the specific candidates for the data can rest with the implementation committee. The entire data for the training set also needs to made available.

Once the maps have been cross calibrated across techniques, groups can be chosen/selected to provide the first repository maps for all 6 clusters that will make the final target list. To produce the first repository maps – the people tasked with setting up the challenge need to provide the existing HST images + multiple image catalogs + spectroscopic & photometric redshifts for these images for the selected map-makers. We recommend a calibration workshop to first compare and contrast the various techniques and the errors arising in each of them. Convening all the groups that successfully meet the first data challenge for more careful calibration is recommended strongly in order to select the official map-making group.

Step 2: Producing the first repository maps. The selected group of map-makers will be required to provide the magnification (Kappa) and shear (gamma) maps for each cluster employing their chosen technique. These need to be the unweighted (not scaled) by DLS/DOS (where DLS = angular diameter distance D between the lens L and the source S; DOS = D between the observer O and the source). This appears as a ratio in the lensing magnification, deflection angle and other quantities of interest and is the only term that depends on the source redshift. Because it is separable from all other terms, maps presented in this way will allow the freedom for the community to generate magnification maps at any source redshift desired. It would be useful to make available to the community (i) all the first repository map data that is already in hand; (ii) unweighted magnification and shear maps derived from all three techniques with the data in hand prior to the commencement of the first observations.

***Step 3: Refining the first maps**.* After sufficient new HST data on each cluster are obtained, the repository of magnification maps should be updated periodically, taking into account new information from the HDFI observations. The new maps would then be provided as part of regularly scheduled data releases from STScI. Meanwhile, any spectroscopic follow-up for any new and additional multiple images and new multiwavelength data acquisition (beyond what is available at hand) will be left to individual PIs to do as they please.

Step 4: Final maps. Once all the data are taken, at the final data release the final honed maps should be provided.

Acknowledgements

We would like to thank the following people for their contributions to this effort: Marusa Bradac, Dan Coe, Jean-Paul Kneib, Johan Richard, Tommaso Treu and Adi Zitrin for helping with lens models and their evaluation; Harald Ebeling, Jens Hjorth, Eric Jullo, and Ian Smail for their aid in compiling ancillary data for clusters; Rob Ivison and Chris Carilli for useful information regarding JVLA performance issues; Alex Fullerton for providing information on JWST NIRISS spectroscopic sensitivity; Gabe Brammer, Pieter van Dokkum, Raphael Gobat, Sangeeta Malhotra, Norbert Pirzkal, James Rhoads, and Ben Weiner for valuable discussions on grism science with HST; Todd Tripp, Christopher Howk, Nicolas Lehner, John O'Meara, Xavier Prochaska, Jerry Kriss, and Stephan McCandliss for help evaluating deep COS spectroscopy; Tom Brown, Julianne Dalcanton, Marla Geha, Raja Guhathakurta, and Evan Kirby for their thoughts on near-field science; and finally Richard Ellis, Garth Illingworth, and George Rieke for general advice.

References

Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al 2010, ApJL, 709, L133
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2011a, ApJ, 737, 90
Bouwens, R. J., Illingworth, G. D., Labbe, I., et al. 2011b, Nature, 469, 504
Bouwens, R. J., Illingworth, G. D., Oesch, P. A., et al. 2012, ApJL, 752, L5
Bradley, L. D., Trenti, M., Oesch, P. A., et al. 2012, ApJ, 760, 108
Brammer, G. B., van Dokkum, P. G., Franx, M., et al. 2012, ApJS, 200, 13
Coe, D., Zitrin, A., Carrasco, M., et al. 2012, arXiv:1211.3663
Dunlop, J. S., McLure, R. J., Robertson, B. E., et al. 2012, MNRAS, 420, 901
Ebeling, H., Edge, A. C., & Henry, J. P. 2001, ApJ, 553, 668
Finkelstein, S. L., Papovich, C., Salmon, B., et al. 2012, ApJ, 756, 164
Gobat, R., Strazzullo, V., Daddi, E., et al. 2012, ApJL, 759, L44
Gonzalez, V., Labbe I., Bouwens, R. J., et al. 2011, ApJL, 735, L34
Grazian, A., Castellano, M., Fontana, A., et al. 2012, A&A, 547, A51
Jullo, E., Natarajan, P., Kneib, J. P., et al. 2010, Science, 329, 924
Oesch, P. A., Bouwens, R. J., Illingworth, G. D., et al. 2012, ApJ, 759, 135
Postman, M., Coe, D., Benitez, N., et al. 2012, ApJS, 199, 25
Robertson, B. E., Ellis, R. S., Dunlop, et al. 2010, Nature, 468, 49
Robertson, B. E. 2010, ApJ, 713, 1266
Shull, J. M., Smith, B. D., & Danforth, C. W. 2012, ApJ, 759, 23
Stark, D. P., Schenker, M. A., Ellis, R. S., et al. 2012, arXiv:1208.3529
Trenti, M., Bradley, L. D., Stiavelli, M., et al 2012, ApJ, 746, 55
Zheng, W., Postman, M., Zitrin, A., et al. 2012, Nature, 489, 406

Appendix 1: Recommendations on Field Selection

Section 6.2 of this report discusses various considerations for field selection that are unrelated to the properties of the lensing clusters themselves, including location on the sky, foreground extinction and emission, availability of supporting data, and accessibility from various ground-based telescopes. In this appendix, we elaborate on these points and make specific observations and recommendations for consideration by the implementation team.

Ecliptic latitude and zodiacal foreground emission:

Fields with high ecliptic latitude have two major advantages: low zodiacal foreground emission, which can limit sensitivities for HST and JWST observations, and improved visibility and schedulability for HST and especially JWST (indeed for most space-based facilities).

For HST and JWST observations in which other sources of scattered light (starlight or earthlight) are not an issue, zodiacal emission is the dominant source of sky background, and thus a main limitation for achievable sensitivity. The zodiacal background is always larger near the ecliptic plane. For abs(lat) > 45°**,** the zodiacal background is >0.75 mag fainter than at the ecliptic, corresponding to an effective difference in exposure time for background-limited observations of a factor of >2. Also, at low ecliptic latitudes the background varies strongly during the year, with sun angle or ecliptic longitude, and can be >1 mag brighter than the minimum value.

The effect of high zodiacal foregrounds on the signal-to-noise ratio for HST, Spitzer, or JWST measurements of faint galaxies can easily be more important than the effects of Galactic extinction, at least at levels that are normally considered for extragalactic survey fields. As described above, in the F814W band, a factor of 2 in zodiacal brightness has equivalent an impact on faint galaxy signal-to-noise ratios equivalent to a Galactic foreground extinction E(B-V) = 0.20, nearly ten times larger than the extinction in any of the CANDELS fields.

Ecliptic latitudes > 45° (absolute value) have significantly greater visibility for JWST (>200 days per year, compared to 100 days per year on the ecliptic plane). The JWST continuous viewing zone is the region within 5 degrees of the ecliptic poles.

The working group defers a full evaluation of the consideration of visibility and zodiacal foregrounds to the implementation group, but notes the following general rules of thumb:

- Ecliptic latitudes > 45° significantly improve JWST schedulability.

- Ecliptic latitudes > 45° are at least 0.5 mag darker than latitudes < 15°.
- Low latitude fields can be scheduled when backgrounds are lowest, but this can be significantly restrictive on field visibility.
- Zodiacal foreground emission can have a larger impact on exposure times and achieved signal-to-noise ratios for faint galaxies than would Galactic foreground extinction over ranges that are typical for extragalactic survey fields.

We note that although three of the five CANDELS fields have ecliptic latitude abs(lat) > 45° (GOODS-N, GOODS-S, and the EGS), two others, the UDS and COSMOS, have latitude = -17.9° and -9.2°, respectively, and are well known for having higher and more strongly time-variable zodiacal foregrounds.

Galactic latitude, extinction, and foreground emission:

It is important to minimize foreground Galactic extinction, and as much as possible to also minimize foreground emission from Galactic dust and cirrus, which can limit the sensitivity of far-infrared observations (e.g., with Herschel, or in the future with SPICA), and potentially mid-infrared observations with JWST as well. Low Galactic latitude fields will also have a higher density of Galactic foreground stars; bright stars may introduce problems with scattered light, particularly for JWST, while faint, cool dwarfs can occasionally mimic the colors of very high redshift galaxies. Generically, Galactic latitudes $|b| > 40°$ would be preferred, with foreground extinction E(B-V) < 0.025. (For example, all five CANDELS fields meet these criteria.) However, the implementation team should examine IRAS foreground maps for all fields. As a rule of thumb, the five CANDELS fields have IRAS 100 micron foreground brightness < 1.3 MJy/sr (and four out of five have foregrounds in the range 0.4-0.8 MJy/sr).

The foreground column density of neutral hydrogen is also an important factor for X-ray observations of these fields. Most of the clusters already have X-ray data, but it is conceivable that new Chandra or XMM observations might be proposed to improve the measurements of X-ray temperatures or substructure. For reference, the largest foreground hydrogen column among the CANDELS fields is $N(H) = 2.04 \times 10^{20}$ cm^{-2} for the UDS.

On a practical level, for HST, Spitzer and JWST observations, the zodiacal foreground brightness has a larger impact on achieved signal-to-noise ratios for faint galaxies than does foreground Galactic extinction at the levels normally considered for high latitude extragalactic fields. As described above, a field with low ecliptic latitude can easily have zodiacal foregrounds that are 2x brighter (0.75 mag) than a field with high ecliptic latitude, reducing signal-to-noise ratios in a fixed observing time by $\sqrt{2}$. The equivalent in terms of foreground extinction would be 0.375 mag. For the critical HST ACS and WFC3 bandpasses from F814W through F160W, this would correspond to E(B-V) = 0.20 to 0.60, far more extinction than any typical high latitude survey field. Even for U-band observations this would correspond to E(B-V) = 0.08, several times larger than the

CANDELS fields. Therefore, from the point of view of observing efficiency we prioritize ecliptic and zodiacal considerations over foreground Galactic extinction. That said, the *uncertainty* in any single measurement line-of-sight foreground extinction must be considered in terms of its potential effect on high-z galaxy luminosities (e.g., for luminosity function measurements) or colors (for UV spectral slope measurements). It is difficult or impossible to precisely calibrate the foreground extinction on scales of arcminutes.

The working group reviewed one white paper that suggested placing fields in regions with constellations of relatively bright foreground stars suitable for use with adaptive optics instrumentation. While interesting, this is impractical, given the goal of targeting particular lensing clusters in this program. However, if some of the candidate clusters do have bright stars nearby it might conceivably be an advantage for ground-based AO facilities. The implementation team may want to investigate whether these stars will introduce significant scattered light problems for JWST or HST.

Declination and visibility from ground-based facilities

Fields near the celestial equator can be observed with telescopes in either hemisphere, maximizing the opportunities to obtain supporting or follow-up observations of these new deep fields. Currently, OIR telescopes with apertures 8-10m are available in both hemispheres, but other facilities such as ALMA and the VLA offer unique capabilities from one hemisphere alone. Below, we discuss in more detail the potential importance of ALMA and VLA observations for these new deep fields; both can be valuable, but ALMA is arguably the most important for this initiative, providing opportunities for spectral line detections of galaxies out to extremely high redshifts. For this reason, fields in the south should be preferred over northern fields inaccessible to ALMA. Additionally, two out of three 20-40m telescope projects, GMT and EELT, are located in the south.

Declinations in the range $-30° < \delta < +25°$ would be reasonably accessible (sec(z) < 1.55) by telescopes on Cerro Paranal (latitude = -24.6°) and Mauna Kea (latitude = +19.8°), as well as by ALMA. The VLA is farther north, at latitude = +34.1°, but can still observe as far south as $\delta = -30°$, albeit with reduced sensitivity.

The working group received white papers arguing for southern circumpolar fields, which would be ideal for future Antarctic telescope facilities. While potentially interesting, we did not feel that these arguments trumped the benefits of easier accessibility at less extreme declinations.

Combined coordinate considerations:

Unfortunately, fields that have both high ecliptic latitude and high Galactic latitude tend to be far from the equator; it is very difficult (or impossible, depending on the adopted requirements) to satisfy all three conditions. Even neglecting considerations of declination, the simultaneous preference for both high Galactic and ecliptic latitudes push observations into rather narrow sky regions, which are more or less where many of the existing deep fields are. Conversely, equatorial fields at high Galactic latitude tend to have low ecliptic latitude (e.g., COSMOS, SXDS/UDS).

It therefore seems inevitable that most of the cluster fields chosen for the HDFI will not simultaneously meet the desired ecliptic, Galactic and equatorial criteria, particularly when lensing considerations are prioritized. The working group offers the following comments as guidelines:

- Higher zodiacal background can be minimized with optimal HST or JWST scheduling, but this will have a significant effect on scheduling observations with those telescopes.
- It seems very unlikely that most of the clusters can be selected to have equatorial declinations, particularly when lensing considerations are weighed heavily. The working group recommends prioritizing southern fields to ensure ALMA accessibility.
- Galactic latitude and foreground absorption and emission considerations could be relaxed somewhat with smaller impact on achieved sensitivity than the effects of high zodiacal emission. However, there is no certain solution for the potential uncertainty in foreground Galactic extinction on the scale of arcminutes, which could compromise the scientific fidelity of precise measurements of galaxy colors and magnitudes.

Other foreground considerations:

Bright radio sources close to the field of interest can severely limit the achievable dynamic range and sensitivity for VLA observations. This is particularly true at low frequencies, where the primary beam of the VLA is large (approx. 30 arcmin diameter at half power at 1.4 GHz). Radio astronomers have complained that this is a significant limiting factor for GOODS-North, even though foreground radio sources were a consideration when selecting that field. The EGS is far worse, with 3C 295 located nearby. The working group did not include experts in radio interferometry, nor try to evaluate potential fields with regard to their radio properties; the implementation team should consult some experts in this area.

The new EVLA antennae and wide-bandwidth correlator have significantly reduced some of the factors that could significantly limit dynamic range with the old VLA, e.g., bandwidth and time smearing. Bright sources within the primary beam are no longer as much of a limiting factor, but bright sources out in the wings of the primary beam can still limit sensitivity, in part because of uncertainties in the asymmetric shape of the primary beam.

The VLA also offers important capabilities for molecular spectroscopy at frequencies below those covered by ALMA. Because these observations are generally carried out at frequencies > 20 GHz, the VLA primary beam is much smaller, and potentially contaminating radio sources are generally much fainter. Therefore, interference may not be as much of an issue for spectral line measurements, although transitions of interest (e.g., CO 1-0) are often extremely faint.

Considerations for ground-based facilities:

As discussed above, ALMA and the VLA are unique facilities that have no counterparts in the opposite hemispheres. 8-10m class OIR telescopes are available in both hemispheres, and there are also planned 20-40m telescope projects for both hemispheres (although 2 out of 3 of those are in the south).

ALMA offers enormous potential for studying high redshift galaxies in these lensing fields, including measurements of dust continuum emission and atomic and molecular spectral lines out to very large redshifts. [CII] 158 μm can be measured with the current ALMA receivers at $4.1 < z < 8.0$ (and also at $1.64 < z < 2.16$), and the forthcoming ALMA band 5 receivers will extend that to $z < 10.66$. In fact, [CII] and other atomic lines ([OI] 63 μm, [OIII] 88 μm, [NII] 122 μm and 205 μm) may be the best way to measure redshifts for the highest redshift lensed galaxies in these deep fields, and can yield valuable information about ISM conditions in those objects. ALMA may be the only facility in the foreseeable future that can directly constrain the contribution of dusty galaxies to the star formation history of the universe at high redshift ($z >> 2$), or to directly measure the bulk of energy emitted from star formation at the peak era of star formation at $z \approx 2$. The lensing clusters surveyed in this HST program will be the most accurately calibrated and most powerful natural telescopes in the universe, and will be the target of intensive survey at all wavelengths into the JWST era. For this reason, it is important to ensure that most (if not all) of these clusters are observable by ALMA. The **PdBI NOEMA** upgrade will significantly enhance northern-hemisphere capabilities for millimeter interferometry, but ALMA will still be superior.

The **VLA** offers complementary spectral line capabilities to those of ALMA, and if Herschel or other far-IR/submm observations are available for these clusters (see below) then deep VLA continuum maps can help to localize and identify lensed far-IR sources (although in that case so can ALMA). The **Large Millimeter Telescope** will also be a powerful tool for millimeter spectroscopy, and is at latitude +19° (similar to Hawaii).

8-10m OIR telescopes are available for both hemispheres, and projects for 20-40m telescopes are ongoing in both hemispheres, albeit with 2 out of 3 planned for Chilean mountaintops.

Considerations for other space-based facilities and supporting data:

Spitzer IRAC seems most immediately relevant to the core high-z science, and we've discussed this extensively. Coordination with the Bradac et al. GO-9 Exploration Science Program is essential, especially regarding telescope orientations in order to optimize joint HST and Spitzer observations of the parallel cluster + blank fields.

Most or all of the candidate clusters have **Chandra and/or XMM** data, since they are famous massive clusters, often X-ray-selected. But it is quite possible that astronomers may propose for deeper data to measure better X-ray temperatures, or to achieve higher S/N with Chandra for measuring X-ray substructure to improve cluster mass models. From this point of view, low foreground N(HI) is a good thing.

Herschel: Although the Herschel far-infrared observatory is unlikely to detect the $z > 6$ galaxies which are the primary motives for this HST observing program, there is significant science on dusty star-forming galaxies at $2 < z < 6$ which can benefit from the lensing boost from massive clusters with ultradeep OIR measurements from HST. Lensing can let Herschel probe regions of the infrared luminosity function at $2 < z < 6$ that are currently inaccessible and almost unknown. At present, the deepest blank-field Herschel observations reach only to ~L*(IR) at z=2; most of the infrared emission from dusty star forming galaxies at $z > 2$ is unresolved by Herschel, and in fact we have almost no direct constraint on how much energy from star formation at $z > 2$ is absorbed by dust, leading to considerable uncertainty in the cosmic star formation history (and many other related topics) at high redshift. We have essentially no idea what the slope of the infrared luminosity function is at $z >> 1$, and lensing can help to address this. So can ALMA, but very deep blank field mapping with ALMA will be expensive - 4 arcmin^2 is a very wide area by ALMA standards! Moreover, such maps will mainly be done at longer ALMA wavelengths (870 μm and longer) where the bolometric corrections to total IR luminosity are large and have large and uncertain temperature dependence. It will always be valuable to have measurements from Herschel near the peak of the far-infrared emission. Deep HST + Spitzer data will identify the counterparts to the Herschel sources (which are frequently optically invisible, and can even be extremely faint in the near-IR), and to measure photometric redshifts, morphologies, etc.; the HST data will be particularly important for strongly lensed or multiply imaged infrared sources.

Many of the candidate clusters for the HST deep field initiative already have Herschel data from open and guaranteed time observing programs. Specifically, we have scanned the Herschel Science Archive for observations of the 16 clusters in the prioritized “long list” for this program. Most of them have relatively deep PACS and SPIRE observations from the Herschel Lensing Survey (PI: Egami, programs KPOT_eegami_1, OT2_eegami_5) or the PEP and HerMES guaranteed time programs (KPGT_soliver_1, KPGT_dlutz_1). One cluster (MACSJ2214.9-1359) only has shallower SPIRE observations from the SPIRE Snapshot Survey (OT1_eegami_4). It is of course possible that other clusters beyond the list of 16 considered here will also be reviewed for implementation. If any of the selected clusters do not have Herschel imaging, and in particular deep PACS observations, STScI would do well to encourage ESA to make last-minute observations with PACS or SPIRE or both, as needed. Whichever clusters are chosen for this deep field program will inevitably become most intensively studied and best-understood examples, and in the long run the community would regret not having good far-infrared data for these fields. The Herschel data are really only essential for the cluster fields, not the parallel blank fields. Existing Herschel data in GOODS and CANDELS, etc., are more valuable than new observations in these new blank parallel fields, thanks to the other supporting data available within CANDELS. Target sensitivities should be at least 2.5 mJy (5σ) at 100 microns with PACS (160 microns is obtained simultaneously), and an instrumental (only, without confusion) 5σ limit of at least 8 mJy.

In turn, deep Herschel observations (or future far-infrared observations from SPICA) will greatly benefit in turn from the availability of deep **Spitzer MIPS 24 micron data.** It is too late to make new observations with MIPS, so fields with existing 24 micron data sensitive to 60 μJy (5σ) or better would be preferred.

Appendix 2: List of HDFI White Papers

Lead author and title are provided.

Andreani	Deep Cosmological surveys to test alternative Cosmologies
Atek	Deep Fields: Beyond the UDF
Baker	White Paper on HST/SKA Survey Synergy
Burgarella	A White Paper on Behalf of the Herschel Team
Caputi	Maximising the HST Legacy for the JWST: the Importance of Multiple HUDF-like Fields for the JWST Early Planning
Coe	A Preview of JWST: Measuring Cosmic Evolution from z ~ 8 to z ~ 11 with Gravitational Lenses
Damjanov	Adaptive Optics Deep Field
Elmegreen	Understanding Galaxy Evolution through another HUDF-like Field
Franx	HST Deep Fields will greatly enhance the effectiveness of NIRSpec in the first two years of the JWST mission
Haberzetti	Star Forming Galaxies at z>1 in the extended HDF-S
Illingworth	JDF - the JWST Deep Fields Survey Treasury Program: A HST Legacy for Galaxy Assembly in the First 2 Billion Years
Illingworth	White Paper on the use of strong lensing clusters to study distant galaxies
Kirshner	A Handful of RAISINS: A very large science program for HST
Kriss	A Deep-Field Parallel Program to Measure the Kinetic Luminosity in an AGN Outflow
Kriss	Deep Ultraviolet Spectroscopy in Parallel with Hubble Deep Fields
Liu	The Next Hubble Deep Field In COSMOS – An Education and Outreach Motivation
Matsuo	Deep Fields in the South for Hubble Survey
McCandliss	Deep Field Parallel Program: COS Search for Lyman Continuum Escape from Star Forming Galaxies
Peterson	AGN Reverberation Mapping: A Deep-Field Parallel Program
Rhoads	Deep Slitless Spectroscopy in the Near Infrared
Riechers	The Need for an Equatorial Hubble Deep Field
Rieke	Additional HST Deep Survey Fields Are Needed in Support of JWST
Robotham	Hubble Deep Fields Placed on Local Group Analogues
Ryan	The Case for Broad Wavelength Coverage in the Next Generation of Deep HST Imaging

Schartel	On the possibility of X-ray follow-up observations with XMM-Newton
Serjeant	A new HST/Herschel deep field at the North Ecliptic Pole: preparing the way for JWST, SPICA and Euclid
Slane	Chandra Visibility for Deep Field Observations
Sheth	A White Paper Towards the Choice of ANother Hubble Ultra Deep Field
Teplitz	The Need for More UV Ultradeep Fields: Building on the UVUDF
van Dokkum	From census to astrophysics: a case for deep WFC3 grism observations
Weiner	Deep slitless infrared spectroscopic surveys with HST/WFC3
Zabludoff	Targeting the First Galaxies with HST and a New Kind of Gravitational Lens