FILLING IN THE GAPS: ILLUMINATING (A) CLEARING MECHANISMS IN TRANSITIONAL PROTOPLANETARY DISKS, AND (B) OLIANITITATIVE ILLITERACY AMONG LINDERGRADILATE SCIENCE

(B) QUANTITATIVE ILLITERACY AMONG UNDERGRADUATE SCIENCE STUDENTS

by
Katherine Brutlag Follette

A Dissertation Submitted to the Faculty of the DEPARTMENT OF ASTRONOMY
In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY
In the Graduate College
THE UNIVERSITY OF ARIZONA

THE UNIVERSITY OF ARIZONA GRADUATE COLLEGE

As members of the Dissertation Committee, we certify that we have read the dissertation prepared by Katherine Brutlag Follette entitled "Filling in the Gaps: Illuminating (a) Clearing Mechanisms in Transitional Protoplanetary Disks and (b) Quantitative Illiteracy among Undergraduate Science Students" and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

	Date: 23 October 2014
Laird Close	
	Date: 23 October 2014
Donald McCarthy	
DI 1111	Date: 23 October 2014
Phil Hinz	
Edward Prather	Date: 23 October 2014
Edward Framer	
Daniel Apai	Date: 23 October 2014
Final approval and acceptance of this dissertation date's submission of the final copies of the disse	O I
I hereby certify that I have read this dissertation and recommend that it be accepted as fulfilling t	1 1
	Date: 23 October 2014
Dissertation Director: Laird Close	

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED: Katherine Brutlag Follette

ACKNOWLEDGMENTS

I would like to thank my science research advisor Laird Close for always being generous with his time and providing critical guidance and support throughout all of the projects included in this thesis, even those for which he was not directly involved. I'd also like to thank my educational research advisor Don McCarthy for being a tireless advocate for me in addition to a co-crusader against quantitative illiteracy.

I would like to thank Jared Males for gently prodding me to become a better programmer and patiently answering all of the stupid questions that I was too embarrassed to ask Laird, and Vanessa Bailey for answering all of the even stupider questions that I was too embarrassed to ask Jared.

I would like to thank Erin Dokter, Sanlyn Buxner and Ed Prather for providing endless support and encouragement as I learned the ropes of educational research.

On the science end, there are almost too many people to thank, so to name just a few: Colette Salyk, Katie Morzinski, Carol Grady, Glenn Schneider, Barb Whitney, Motohide Tamura and Jun Hashimoto all provided critical guidance on various projects.

I would like to acknowledge the University of Arizona College of Science Fellowship for funding year 1 of my graduate career, the National Science Foundation Graduate Student Fellowship Program for funding years 2-4, the Japan Society for the Promotion of Science for funding my summer at the National Astronomical Observatory of Japan in Mitaka in 2011, The University of Arizona Technology Research Initiative Fund Scholarship for supporting me during year 5, and a National Science Foundation Transforming Undergraduate Education in STEM grant for funding year 6. I was also supported in conference travel and research supplies by various NSF and NASA grants supporting Magellan AO.

This thesis uses astronomical data collected with the Magellan Clay telescope at Las Campanas Observatory and the Subaru telescope on Mauna Kea.

DEDICATION

This thesis is dedicated to Dave and Alex, the most important people in my life and cheerful sources of much-needed emotional support and baby giggles, respectively, during the dissertation process.

Also for my dad, who took me out to see Halley's comet in 1989, and gave me my first memorable experience with astronomy when he told me how old I would be when it came back. For my mom, who demonstrated throughout my childhood that women could have families and love their jobs as well. For my grandmother Gertrude, who instilled those values in my mother in the first place.

For my advisor Laird Close, who has never been afraid to tell me when I'm wrong, and for Don McCarthy who supported not just my interest in education research and quantitative literacy, but also provided the much-needed outlet of Astronomy Camp during my graduate experience.

For a few other notable mentors over the years, especially my 8th grade science teacher Russ Nahorniak, my undergraduate advisor, Frank Winkler, and Katy Garmany, who directed Columbia University's Universe Semester program, one of the highlights of my life and astronomy career thus far.

And for my friends, colleagues and mentors. Thank you all!

TABLE OF CONTENTS

	LIST	OF FIG	GURES	10
	List	OF TA	BLES	25
	ABS	TRACT		26
1	Inti	RODUC	TION	28
	1.1	Exopl	anet Discovery Methods	28
	1.2		mstellar Disk İmaging	
		1.2.1		31
		1.2.2	Disk Classification	33
		1.2.3	Disk Structure	36
		1.2.4	Disk Evolution	37
		1.2.5	Clearing Mechanisms	38
	1.3	High-	Contrast Direct Imaging of Young Exoplanets and Circum-	
			Disks with Modern Adaptive Optics Systems	44
		1.3.1	Adaptive Optics	46
		1.3.2	Deformable Mirrors and Wavefront Sensing	46
		1.3.3	MagAO and Subaru AO Systems	50
	1.4	Differ	ential Imaging Techniques	53
		1.4.1	Simultaneous Differential Imaging	
		1.4.2	Polarized Differential Imaging	56
		1.4.3	Angular Differential Imaging	58
2	Н-в	AND S	CATTERED LIGHT EMISSION FROM THE MYSTERIOUS SR21 TRA	N-
_			DISK	
	2.1		luction	
		2.1.1	Observational History of SR21	
	2.2		vations and Data Reduction	65
		2.2.1		66
		2.2.2	Intensity Images	
	2.3	Resul	ts and Analysis	
		2.3.1	Polarized Intensity Isophotes	74
		2.3.2	Radial Brightness Profiles and Asymmetries	77
		2.3.3	Intensity Images and Compatibility with the Existence of a	
			Substellar Companion	81
	2.4	Mode	ling	85
		2.4.1	The Whitney Code and Input Assumptions	86
		2.4.2	Sub-mm Image Modeling	88
		2.4.3	SED Modeling	91
		2.4.4		101

	2.5 2.6	Modeling Discussion	
2			
3		DS ADAPTIVE OPTICS IMAGING OF THE ASYMMETRIC TRANSITIO	
		K OPH IRS 48 IN SCATTERED LIGHT	
	3.1	Introduction	
	3.2	Observations and Data Reduction	
		3.2.1 Subaru/HiCIAO	
	2.2	3.2.2 IRTF/SpeX	
	3.3	Results	
		J	
		3.3.2 Cavity Morphology	
		3.3.4 Oph IRS 48 Age and Spectral Type	
	2.4	3.3.5 Accretion Rate Estimate	
	3.4	Discussion	
		3.4.1 Monochromatic Radiative Transfer Modeling wth Sprout .	
		3.4.2 SED modeling	
		3.4.3 Modeling of Scattered Light Images	
		3.4.4 Brightness Asymmetries in Real and Modeled Data	
		3.4.5 Offset of Cavity Center from Stellar Residual	
	2.5	3.4.6 On the Nature of the Western Extension	
	3.5	Conclusion	. 148
4	Тне	FIRST CIRCUMSTELLAR DISK IMAGED IN SILHOUETTE WITH ADAI	P_
	TIVE	E OPTICS: MAGAO IMAGING OF ORION 218-354	. 150
	4.1	Introduction	. 150
	4.2	Observations and Data Reduction	
	4.3	Results and Analysis	
	4.4	Conclusion	
5	CON	ICLUSIONS AND FUTURE WORK	166
0	5.1	Summary	
	5.2	Resolving Photoevaporative Disk Surfaces	
	J.Z	5.2.1 Modeling and Test Cases	
		5.2.2 Case Study: HD100546	
		5.2.3 Future Work	
	5.3	MagAO Giant Accreting Protoplanet Survey	
	0.0	5.3.1 Case Study: HD142527	
		5.3.2 2014A Candidates and Future Work	
	5.4	Future Work	
	J.T	5.4.1 Multiwavelength Polarized Intensity Disk Imagery	
		c. I.	

		5.4.2	Total Intensity Disk Imagery	. 185
		5.4.3		
6	AN	Infori	MED APPROACH TO IMPROVING QUANTITATIVE LITERACY A	ND
	MIT	IGATIN	ig Math Anxiety in Undergraduates through Intr	0-
	DUC	CTORY S	SCIENCE COURSES	. 189
	6.1	Backg	ground and Literature Review	. 189
		6.1.1	Cross-Curricular and Continuous Application	. 193
		6.1.2	Faculty Development and Training	. 195
	6.2	Inves	tigatory Study Component	. 196
		6.2.1	Study Motivation	. 196
		6.2.2	Regarding Important At Risk and Influential populations	. 199
	6.3	Facul	ty Development Component	. 205
	6.4	Concl	usion	. 207
7	Тнв	e Quan	ITITATIVE REASONING FOR COLLEGE SCIENCE (QUARCS) A	S-
			I: DEVELOPMENT AND VALIDATION	
	7.1		act	
	7.2	Introd	luction	. 208
		7.2.1	Why Should Science Educators Join the Numeracy Crusade	e? 209
		7.2.2	Why General Education College Science Courses?	. 213
		7.2.3	Why Astronomy 101?	. 216
	7.3	Asses	sment Development	
		7.3.1	Skill Selection	. 218
		7.3.2	Question Wording	. 223
		7.3.3	Development of Demographic and Attitudinal Questions	. 224
		7.3.4	Inclusion of Confidence Rankings	. 228
		7.3.5	Format	. 228
		7.3.6	Self-Reporting of Effort	. 233
		7.3.7	Assessment Length	
		7.3.8	Matched Data	. 239
	7.4	Item 1	Analysis	. 240
		7.4.1	Free Response Items	. 240
		7.4.2	Expert Analysis	. 242
		7.4.3	Item Statistics	
	7.5	Imple	ementation and Lessons Learned	. 255
		7.5.1	Attrition and Apathy	. 255
		7.5.2	Effects of Effort Filter on Distribution	
		7.5.3	Additional Filters	
		7.5.4	Demographics of Low Effort Sample	. 262
	76	Concl	usions	. 263

8	EDUCATION CONCLUSIONS
A	SR21 DISK MODELING SUPPLEMENT
В	10μ M Feature
C	RADIAL PROFILES OF OPH IRS 48
D	SKILL SURVEY RESPONSES
E	QUARCS DEMOGRAPHIC AND ATTITUDINAL QUESTIONS
	References

LIST OF FIGURES

1.1	Schematic representation of a full disk with annular regions that emit as blackbodies at successively longer wavelengths toward the	
	cooler outer regions. If the innermost disk component is removed,	
	it manifests itself as a lack of excess in the NIR	35
1.2	Schematic representation for the three main candidates for the clearing of transitional disk cavities. Each makes its own predictions	
	about the degree of clearing inside of the disk cavity	38
1.3	Numerical simulations of gap opening by a single planet showing the differences in gap size and degree of clearing according to both	50
	the size of the planet and the size of the grains in question. Larger	
	planets open wider, more cleared cavities. Larger grains are most	
	efficiently cleared, while small grains are less affected. These sim-	
	ulations extend only to grain sizes of 0.03mm. For grains an order	
	of magnitude smaller - the μ m sized grains probed by NIR scat-	
	tered light imaging - the cavity all but disappears. Reproduced	
	from Zhu et al. (2012)	43
1.4	A selection of observations, model images and solar system analogs	
	of features that may be indicative of the influence of planetary per-	
	turbers in circumstellar disks, many of which are non-axisymmetric.	45
1.5	A sample MagAO H α PSF, showing a clear dark hole inside of the	
	system's control radius at 0.25". This is the region of Fourier space	
	corresponding to well-sampled and corrected wavefront aberrations	51
1.6	The mechanical design of the Magellan VisAO and PWFS system,	
	with the key features for visible AO science indicated by blue ar-	
	rows. The red loops indicate the two control loops for wavefront	
	correction. Light from the ASM enters from the bottom	52
1.7	Top: Schematic of the Subaru + AO188 + HiCIAO optical system.	
	Bottom: More detailed schematic of the HiCIAO optics	53
1.8	A schematic of MagAOs SDI imaging mode. The visible light beam	
	is passed through a Wollaston prism, which splits it, allowing half	
	of the light to pass through a narrowband filter centered on a spec-	
	tral line of interest (H α , [OI] or [SII]), and half through a filter cen-	
	tered on the neighboring continuum. Each spectral line probes a	
	different type of phenomenon, described in detail in the text. In the	
	case shown at right, $H\alpha$ emission is being used as a probe of accret-	
	ing material. In five of the six filters shown, the primary (lower)	
	star is brighter, however the additional accretion luminosity of the	
	secondary (upper) star makes it the brighter one in the H α image	
	(upper left panel)	55
	C 11 L	

1.9	Sample Polarized Intensity image with centrosymmetric polarization vectors overlain (left) and histogram of orientations relative to the radial direction (right). Image credit: Jun Hashimoto	56
2.1	Median polarized intensity (PI) image of SR21 at H-band. The gap radius inferred by A11 and the companion orbit inferred by E09 are indicated with white and red dashed ellipses respectively	67
2.2	Left: Polarization vectors overplotted on the median polarized intensity (PI) map of SR21. A general centrosymmetric geometry is evident in the SR21 vectors, suggesting that this emission does indeed trace scattered light from a circumstellar disk. Right: Polarization vectors for the PSF star HD 148212 overplotted on its median PI image, which has been scaled by a factor of 3 to compensate for the difference in exposure time. In this case, the vectors do not show a coherent pattern, and PI emission is weak and compact. This is interpreted as photon noise residuals and is used to estimate the error in the polarized intensity profiles shown in figure 3. In both cases, vectors are scaled by the strength of the polarized intensity signal in the region and are not plotted when they fall at or below noise level. Vectors whose lengths correspond to 500, 1000, 1500 and 2000mJy/arcsec² are shown in the bottom right of the SR21 panel as a scale reference	69
2.3	SR21 polarized intensity image (background) and polarization vectors (overplotted) after first order correction for a seeing limited halo component. This method reveals the polarization vectors to be more centrosymmetric, particularly along the disk minor axis) than originally inferred. Local deviations from centrosymmetry persist in the northeast and southwest quadrants of the disk, where unresolved disk structures and/or companions could induce local polarization structure that deviates from the larger centrosymmetric pattern. The process of subtracting this halo component has introduced additional PSF artifacts into the background PI map, most evident in the negative astigmatic pattern centered on the star. We have thus chosen to model the uncorrected polarized intensity map, which we believe more closely resembles the true structure of the SR21 disk. Vectors whose lengths correspond to 500, 1000, 1500 and 2000mJy/arcsec² are shown in the bottom right of the SR21 panel as a scale reference.	73

2.4	Intensity images of SR21(left) and HD 148212 (right). HD 148212 has been scaled according to differences in exposure time and intrinsic H-band magnitude (based on published 2MASS values) to match SR21 for subtraction. Some uncorrected astigmatism is evident in the four lobes of the PSF in both stars	74
2.5	Elliptical isophotal fits to the SR21 Polarized Intensity image. The fits are shown overplotted on the PI image in white starting at a semimajor axis of 10 pixels (well outside the saturation radius) to a semimajor axis of 50 pixels (where the PI flux drops to noise level). The major axis of each isophote is overplotted as a straight white line. The average value of the major axis position angle is clocked by $\sim \! 15^{\circ}$ from the inner to outer disk, suggesting unresolved disk	, 1
	structure. The ellipticity of the isophotes was also used to estimate the inclination value of 14° (where 0° would be face-on)	75
2.6	Radial polarized intensity profiles for SR21 binned to the H-band diffraction limit (0.5") and shown as triangles. The black filled triangles represent the azimuthally-averaged radial profile and the	
	red open triangles represent the profile as measured in an 11 pixel wide region centered on the major axis of the disk. The blue open triangles represent the profile of the PSE subtracted image.	
	triangles represent the profile of the PSF-subtracted image. All three profiles follow a very similar slope, with the major axis pro-	
	file systematically higher in flux per pixel by 20% relative to the	
	azimuthally-averaged profile, and the PSF subtracted image sys-	
	tematically lower, as is to be expected. Because the images were	
	saturated interior to 0.07", radial profiles are only shown for bins	
	that fall completely outside of the saturated region (from 0.1" out-	
	ward). Also shown is a canonical r^{-2} profile in green and a r^{-3} pro-	
	file in blue. The r^{-3} profile is a better fit to the data, although there	
	is some evidence that it becomes shallower in the outer disk be-	
	yond 0.45". The veracity of this shallower profile in the outer disk	
	is brought into question by the steepness of the PSF subtracted pro- file, suggesting that the flattening in the outer profile may be due	
	to halo effects. Errors were estimated from the photon noise in the	
	four individual channels used to create the PI image summed in	
	quadrature (black error bars). The location of the sub-mm cavity	
	wall at 36AU and the noise floor are indicated with vertical and	
	horizontal dashed lines, respectively	78

2.7	H-band PSF-subtracted intensity image with the 18AU orbit of the companion proposed by E09 overplotted. There is a distinct excess of emission in the Northeast quadrant (A) at a location that is consistent with their hypothesized companion, although its proximity to PSF artifacts makes robust identification impossible with this data. There is also a northward extension of the disk that is nominally present in the PI image (B). The image has been smoothed with a FWHM=6 pixel Gaussian in order to suppress artificial struc-	
	tures induced by speckle noise	83
2.8	Left: The $880\mu m$ SMA image of A11. Right: A large grain disk	
	model with A11's best-fit input assumptions produced by the Whit-	00
2.0	ney code	88
2.9	Highly depleted (δ =10 ⁻⁸ -10 ⁻⁶) large grain disk models produced by the Whitney Code. From left to right, the panels are (a) the ob-	
	served SR21 SMA image of A11 (b) the final Whitney model image,	
	where the raw model output was sampled with the same spatial	
	frequencies observed with the SMA (c) the residuals and (d) the	
	model fit (in red) to the SMA visibilities. Input parameters were	
	identical to those of A11, except that the depletion factor (δ) for	
	large grains inside of 36AU was varied. The depletion factor is	
	given above each model row. A scale factor was also applied to	
	each model image to equalize the model and SMA fluxes before subtraction from the observed image. This factor is given in panel	
	(b) of each model. The minor discrepancies between modeled and	
	observed fluxes can be easily explained by slight differences in in-	
	put opacities	92
2.10	Moderately depleted (δ =10 ⁻⁵ -10 ⁻³) large grain disk models pro-	
	duced by the Whitney Code. See Fig. 9 caption for details	93
2.11	Minimally depleted (δ =10 ⁻² -10 ⁰) large grain disk models produced	
0.10	by the Whitney Code. See Fig. 9 caption for details	94
2.12	Two model SEDs produced by otherwise identical models with	
	(top) a small grain disk that extends inward to 0.07AU and (bottom) a more distant 7AU truncation in the small grain disk and a	
	warm companion with E09 parameters. Both models provide good	
	fits to the λ <12 μ m NIR excess	97

2.13	Top: model SEDs of varying r<36AU depletion factors overplotted on the observed SR21 SED. The nine "Gapped" model runs of dif-	
	ferent depletion factors are shown with solid colored lines, ranging from 10^{-8} depletion in red to zero depletion in purple. Bottom: H-	
	band radial polarized intensity profiles for the same models over-	
	plotted on the observed profile. The flux of the undepleted model	
	was fixed to the data at $r=36AU$, and all other models scaled by	
	the same amount, in order to highlight changes in the radial pro-	
	file at the cavity wall. None of the models provide a good fit to the data, revealing that the small grain disk is not only undepleted, its	
014	geometry is different from that of the large grain disk	102
2.14	Sample model outputs for a representative optically thin disk model. The lefthand panels represent the density of large grains (upper)	
	and small grains (lower). The righthand panels provide model fits	
	to the observed SED (upper) and H-band radial polarized intensity	
		104
2.15	Sample model outputs for a representative three component disk	
	model. In this case, the underlying optically thick small grain disk	
	is of the 'Steeply Curved' variety, and the scattered light contribu-	
	tion from the optically thin component is thirty times that of the optically thick component The lefthand panels represent the den-	
	sity of large grains (upper) and small grains (lower). The righthand	
	panels provide model fits to the observed SED (upper) and H-band	
	radial polarized intensity profile (lower)	106
2.16	A schematic view of the three component disk model that provides	
	the best fit to our data. The vertical scale has been inflated by a factor of two relative to the horizontal. The origins of sub-mm	
	light, H-band scattered light and the steep 12-20 μ m SED rise are	
	· · · · · · · · · · · · · · · · · · ·	107
3.1	Onb IPS 48 HiCIAO H band (loft) and Ke band (right) PL images	
3.1	Oph IRS 48 HiCIAO H-band (left) and Ks-band (right) PI images. In all cases, an r=20 (0.19") pixel mask has been applied to cover	
	the stellar residual. Polarization vectors are shown overplotted in	
	bins equivalent to half of the FWHM of the star at each wavelength,	
	and the length of each vector corresponds to the strength of the	
	signal in that bin.	116
3.2	(a) $0.8-2.4\mu m$ SpeX spectrum, with the spectrum of another A0 star,	
	SAO206463 overplotted.; (b) Pa β line fit used together with a similar fit to the Breeline to derive a week apprecian rate of $\dot{M}=10^{-8.5}$	
	ilar fit to the Br γ line to derive a weak accretion rate of $M=10^{-8.5}$ M $_{\odot}$ /yr for Oph IRS 48.; (c) 2.3-4.3 μ m SpeX spectrum, with perti-	
	nent lines labeled	120

3.3	Top panels: Elliptical fits to the Oph IRS 48 scattered light cav-	
	ity at H (left) and Ks (right) bands. The locations of the peaks for	
	each 10° azimuthally binned radial profile are shown with white	
	crosses. The best fit ellipse to these points is shown overplotted as	
	a dashed white line in each case. The center of that ellipse is shown	
	with a large white cross. As discussed in the text, the Ks band cav-	
	ity fit is significantly more robust, as it is (i) higher resolution data,	
	(ii) brighter and (iii) the morphology is clear enough at this wave-	
	length that profiles dominated by the asymmetry can be excluded.	
	Bottom panels: The same images, but deprojected along PA=97.4°	
	to account for an inclination of 50° and with the deprojected circu-	
	lar cavity fit shown. The large northern offset between the stellar	
	centroid and the cavity center is discussed in detail in Section 3.4.5	121
3.4	Ks band polarized intensity contours overplotted on the H band	
	polarized intensity image of OPH IRS 48. The improvement in	
	image quality in the K band image clarifies that the excess emission	
	seen to the West of the star at H band is probably the Southern disk	100
2 -	rim, clearly visible in the K band image	123
3.5	Radial profiles taken through the H (left) and Ks (right) images.	
	Azimuthal asymmetry is especially apparent along the Western	
	major axis (dashed black lines), where profiles at both wavelengths	
	peak ~0.15" farther from the cavity center, though the asymmetry	
	is also hinted at in the Southwestern profiles (red dashed lines).	
	The blue scattering properties of the asymmetry are readily apparent, the H hand Western major axis profile neaks at 1850 of	
	parent - the H-band Western major axis profile peaks at \sim 85% of the Eastern major axis value, while at Ks-band this value is \sim 25%	
	lower (\sim 60%). Profiles were taken with respect to the apparent	
	cavity center rather than the location of the stellar residual, and	
	this choice is described in detail in Appendix C. The disk im-	
	ages were deprojected to account for an inclination of 50° along	
	PA=97.4°, and the Ks band image was convolved with a 5.65 pixel	
	(0.05") Gaussian in order to simulate similar Strehl-ratio imagery.	
	In all cases, solid lines represent profiles through the Eastern half	
	of the disk and dashed lines represent profiles through the West-	
	ern half of the disk. Black profiles were taken along the disk major	
	axis and blue and red profiles were taken 20° North and South of	
	the major axis, respectively.	126

The spectral energy distribution (SED) of Oph IRS 48. The lower set of data are observed values, while the upper sets represent the data de-reddened using two different reddening laws. Both reddening laws assume log(g)=4.0 and a distance of 121pc. The magenta curve is an A0 stellar template dereddened according to Weingartner & Draine (2001) Case B with $R_V=5.5$ and $A_V=11.5$, which yields L=14.3L $_{\odot}$. This was the assumption of Brown et al. (2012b). The blue line is an A0 template dereddened according to McClure (2009), which may be more appropriate for this high extinction region. This reddening assumption yields a factor of ~ 1.7 higher dereddened luminosity of $23.6L_{\odot}$. The upper curve tracks the stellar photosphere more closely out through the NIR, calling the presence of a hot dense inner disk component into question . . 128

3.7 An H-R Diagram showing the location of Oph IRS along Siess et al. (2000) PMS evolutionary tracks. The green square shows the location of Oph IRS 48 according to the Brown et al. (2012b) reddening assumption, while the red asterisk shows the location of Oph IRS 48 derived from our best-fit dereddening. The factor of \sim 1.7 difference in the luminosity derived from the two reddening assumptions results in different assumptions about the mass and age of Oph IRS 48. Brown et al. (2012b)'s Weingartner & Draine (2001) law puts Oph IRS 48 at \sim 15Myr along the 2.0M $_{\odot}$ evolutionary track. We find reasonably well-fitting SEDs with Luminosities as high as $45L_{\odot}$, corresponding to a $2.5M_{\odot}$ star at 5Myr, and our best fit McClure (2009) law puts it at \sim 8Myr along the 2.2M $_{\odot}$ evolutionary track. Whether or not these isochrones are the most appropriate for young A-stars is a matter of contention, however our higher derived luminosity puts Oph IRS48 along consistently higher mass isochrones and at a younger age than previous work regardless of chosen evolutionary model, as described in the text. . 131

3.8	Top: Simulated images of Oph IRS48 at H-band (left) and Ks-band	
	(right) generated by the monochromatic radiative transfer Sprout	
	code. The data are shown overplotted in each case as black con-	
	tours. The location of the illumination source is marked with a	
	white diamond, and the locations where profiles were extracted	
	are marked with white dashed lines, corresponding roughly to the	
	"major axis" relative to the apparent cavity center. The combi-	
	nation of an elevated flaring parameter (β =1.3) and a larger scale	
	height at the cavity rim (15AU) are able to reproduce the apparent	
	Northward offset of the scattered light cavity from the star, verify-	
	ing that it is likely a result of viewing geometry and is not a true	
	offset, as described in detail in the text. Bottom: Simulated radial	
	polarized intensity (blue) profiles for these models shown relative	
	to the data at H (left) and Ks (right).	135
3.9	The best fitting SED model generated with the Whitney code, shown	
	in black and following the parameters outlined in Table 3.2, over-	
	plotted on observational data from the literature. The sources for	
	the literature data are described in detail in Table 3.1. The poor	
	quality of the fit in the $10\mu m$ region is discussed in Appendix B	136
3.10	Simulated images of the Oph IRS 48 disk created with the Whit-	
	ney code at H band (left) and K band (right). In the bottom panel,	
	these images have been convolved with the observed PSF at each	
	wavelength.	140
3.11	Radial profiles taken through the deprojected H (left) and K (right)	
	model images, which are informative in contrast to the true ob-	
	served radial profiles shown in Figure 3.5a. The model profiles are	
	drastically different from the observations in their East/West sym-	
	metry. They do, however provide a reasonable reproduction of the	
	relative shape, location and relative intensity of the peaks outside	
	of the spiral arm.	141
3.12	Contour map of Geers et al. (2007) 18μ m VLT VISIR data in yellow	
	and van der Marel et al. (2013) 0.44mm continuum ALMA data	
	in green overplotted on our HiCIAO median halo-subtracted H	
	band PI image. Note that the 18μ m thermal emission is centered	
	on the star, and that its Eastern lobe is cospatial with ours, whereas	
	the $18\mu m$ Western lobe lies inside of our scattered light lobe. As	
	discussed in the text, this is evidence that the Western arc is fainter	
	simply because it is farther from the central star	142
	* ·	

Upper left: MagAO H α channel. Lower left: MagAO 643nm continuum channel. Lower right: MagAO difference image created by subtracting the scaled continuum (lower left) from H α (upper left). Upper right: HST ACS image of the same disk (Ricci et al. 2008). All images are North up, East left and have the same physical scale. Note that the disk is visible in all panels except the continuum image, where it shouldn't be visible. Subtraction of the simultaneous PSF provided by the continuum channel effectively removes both the primary star and a secondary star 2" to the SE, isolating H α emission from the background nebula. The disk stands out starkly in silhouette in this (lower right) image, and light blue contours reveal that the background flux from the nebula is attenuated to an increasing degree as column density increases towards the center of this disk. The extent of the contours all the way to the central star is suggestive of an optically-thin disk at $H\alpha$ 154

Radial intensity profile of the Orion 218-354 disk normalized to the background nebular intensity level. Profiles are shown for the major (black diamonds) and minor (blue asterisks) axes. An opaque optically-thick disk model with the best-fit parameters derived in McCaughrean & O'dell (1996) convolved with both 150x150 and 250x250 MagAO continuum PSFs are shown overplotted (purple triangles, yellow squares). In order to demonstrate that the character of the profiles is the same, the 250x250 pixel model has been normalized to the 150x150 model to remove the pedestal effect caused by "filling in" the inner disk with nebular light. The inconsistency in shape and character of these models with our data is striking, and the steadily decreasing profile we observe suggests an optically-thin disk. The performance of a modern AO system, as well as the simpler methodology of SDI-mode PSF subtraction, gives us a clear advantage over the HST data, in which the central star is heavily saturated. The best-fit to the MagAO profile, a powerlaw with exponent 0.46 and an exponential cutoff beyond \sim 0.75", is shown overplotted in red. Errors due to a 3% deviant scaling of the continuum PSF for subtraction are also shown.

4.3	Radial extinction/column density profile of the Orion 218-354 disk in linear and log (inset) space. Overplotted on the data (black aster-	
	isks) in both plots are the Figure 2 intensity fit translated to units of	
	extinction (red line) and a best-fit powerlaw to the extinction pro-	
	file in the inner disk (blue line), which has an exponent of -1.43.	
	This value is used to estimate the radial mass distribution of the	
	disk ($\rho \propto r^{-\alpha}$), under assumptions described in detail in the text.	
	This extinction should be considered to represent the column den-	
	sity of primitive, ISM-like material in the disk, and not of the disk	
	as whole	61
4.4	Multiwavelength photometry of Orion 218-354 (Muench et al. 2002;	
	Da Rio et al. 2009; Mann & Williams 2010; Terada et al. 2012) over-	
	plotted on our best-fitting Whitney model output. The stellar spec-	
	trum (dashed black line) and best-fit model (solid black line, Table	
	1) are shown. The dashed red line corresponds to the same disk	
	model with $r_{disk,in}=r_{sub}$. The poor fit suggests that Orion 218-354	
	contains an inner \sim 2AU clearing	62
5.1	Simulated MagAO [OI] SDI images of HD100546 with three plau-	
	sible origins: A continuous disk, A disk with a gap from 5-10AU	
	and a jet. In all cases, despite a conservative Strehl PSF and an	
	implausible PSF mismatch, the morphology of the system is well	
		.72
5.2	Model (top) and recovered (bottom) images of a disk+jet system,	
	where the jet is an order of magnitude brighter than the disk 1	
5.3	Predicted contrast between [OI] and continuum emission in HD100546	
- 1	as a function of radius in the disk (plotted in VisAO pixels) 1	1/5
5.4	Continuum subtracted MagAO [OI] image of HD100546. Aside	
	from various repeatable artifacts (labeled), no obvious [OI] emis-	70
E E		176
5.5	SNR Maps of MagAO images of the 86mas (~13AU) companion	
	HD142527B. Top panel: Improved ASDI processing of the 2013	
	data and 2014 followup observations (Follette et al. 2014). Or-	
	bital motion of the companion is apparent. Bottom panel: Various zoomed-in PCA reductions of the 2014 data. The location of the	
	ASDI source, which we assume is the Rodigas, Follette et al. 2014	
	total intensity source, is marked with a yellow circle. The red circle	
	shows the offset of the polarized intensity source in the GPI data.	
	An elongation in this direction is often present in PCA reduced im-	
	ages, indicating a possible extension (circumsecondary material?)	
	•	81

MagAO GAPlanetS PCA reduced images of HD100546 revealing an apparent H α point source at 0.13" and PA \sim 150 that is bright (\sim 8 sigma) at H α and \sim 20-50% fainter in the continuum. This location corresponds roughly with the April 2014 position of a planet candidate predicted by Brittain et al. 2014 based on high resolution CO spectroastrometry over a 10 year baseline (right). Reproduced from Follette et al. 2015 (in prep)
Distribution of scores (out of 25 total possible) on the 2013 Pre- Semester Assessment. The range of student abilities is very wide, and the average ability level is low. Mean = 14.8 (59%), Median=15 (60%), Standard Deviation=4.8 (19%)
The perceived importance of various quantitative skills to science literacy (circles) and being a savvy consumer/citizen ("real life", squares) among Astronomy Educators (red symbols) and Math/Numeracy educators (blue symbols). The skills are ranked in order of their perceived importance from left to right, with those skills included in the final assessment in bold black (left) and those that were excluded in grey (right). Three skills measurement, using numbers in writing and making graphs were ranked higher than several skills included in the assessment, but were not added because of the difficulty in assessing them in multiple choice format. Average ranking for all contexts are shown as horizontal bars. Histograms showing the full breakdown of educator responses by discipline
and situation are given in Appendix D
educators
2013 and Fall 2014 administrations

7.6	Histograms of student confidence in correct answers (right) and	
	in incorrect answers (left). More than half of students answering	
	questions incorrectly report that they are "confident" or "very con-	
	fident" in their answers.	230
7.7	A histogram of student responses from the Fall 2013 and Fall 2014	
	administrations reflecting the self-reported usage of calculators.	
	More than half of students report that they use a calculator on 25%	
	or fewer of the questions.	232
7.8	Distribution of student responses to the effort question described	
	in this section. These responses (N=944) include all three adminis-	
	trations (Fall 2013 and Fall 2014 v1 and v2) in which the question	
	was asked. Variations between administrations in students choos-	
	ing each effort ranking are <10% between administrations	234
7.9	Histogram of scores among "experts" taking the QuaRCS (N=34,	
	filled purple). A similar histogram for students (N=251, dashed	
	purple) taking the same version of the assessment. Only the stu-	
	dents who devoted effort to the assessment were selected here, be-	
	cause this group has the highest mean score among students. Even	
	so, the two distributions are discrepant at the 0.000 significance	0.40
- 40		243
7.10	Average correct response rates for students (blue squares, N=251)	
	and experts (red diamonds, N=34) on the Fall 2013 pre assessment.	
	Experts perform substantially better than students, indicating that	
	the test is effective at measuring quantitative ability. Three prob-	
	lematic questions were identified through this analysis. One (item	
	20, red diamond) was removed from the assessment, and the other	
	two (items 2 and 14, yellow diamonds) were substantially revised	0.45
711	for clarity, as described in the text	245
7.11	Pre/Mid semester changes in item difficulty (bottom panels) and	
	discrimination (top panels) for the Fall 2013 instrument. The mid administration was choen in this case both because more students	
	completed it and, as will be discussed in Section 7.5.1, a larger pro-	
	portion of them devoted effort to the assessment. The data are	
	shown both unfiltered (right panels) and filtered to exclude only	
	students whose effort waned midway through the assessment (left	254
	panels)	254

7.12 Visual representation of the distribution of students completing just one assessment (red, yellow, blue wedges), those completing two assessments (orange, green, purple wedges) and those completing all three (black wedge) in classes where all three were assigned for participation credit. Among the 371 unique students in this sample (represented by the outer circle), roughly one-third fall into each category. The dark wedges along the inner circle represent the students within each category who completed the assessment with effort. These students represent 50\% of the population in each lightly shaded wedge, except in the region representing students who completed all three assessments (black region),

7.13 Charts representing the movement of students among effort categories on pre-mid (top) and pre-post (bottom) matched assessments. Effort on the pre-semester assessment is along the x-axis in each case and effort on the mid/post is along the y-axis. The size of each bubble in the grid corresponds to the number of students falling into that category. Bubbles are colored according to whether students improve (green), maintain (blue) or reduce (red) their effort level. The background of the plot is divided according to regions where students have always devoted sufficient effort to persist into the final filtered sample (green), students who never devote enough effort to be included in a sample (red) and those who fall out of the sample at the pre (blue) or mid/post (purple) assessments. The number of students in the purple region increases throughout the semester, indicating increasing apathy.

7.14 Histograms of students scores on the three Fall 2013 assessments - pre (top), mid (middle) and post (bottom) semester. The lefthand column represents the raw samples (N=518, 343, 202), and the righthand column represents the effort and time filtered sample (N=282, 175, 91). The mean of the filtered samples (pre=16.7, mid=17.6, post=17.6) are systematically higher than those of the unfiltered sample (pre=14.8, mid=14.7, post=13.7). This filter becomes particularly important when comparing the pre and postsemester distributions. The raw post population appears to score significantly lower on the assessment than the pre semester sample, but the large population of students in the 5-10 score range that appears in the post-assessment histogram all but disappears when filtered for self-reported effort. We believe this is a reflection of increased apathy toward the end of the semester, as discussed in detail in the text. Note also that the number of students completing each assessment drops significantly as the semester progresses. . . 264

7.15	Histograms of time elapsed on the three Fall 2013 assessments - pre (top), mid (middle) and post (bottom) semester. The lefthand column represents the raw samples (N=518, 343, 202), and the right-hand column represents the effort and time filtered sample (N=282, 175, 91). The median of the filtered samples (pre=30.0, mid=26.0, post=20.0) are systematically higher than those of the unfiltered sample (pre=26.5, mid=20.0, post=16.0). Furthermore, for all samples, the median time to completion decreases by \sim 5 minutes with each administration. We believe that this is primarily a result of increased apathy, as discussed in detail in the text	265
A.1	Sample model outputs for a representative gapped disk model. The lefthand panels represent the density of large grains (upper) and small grains (lower). The righthand panels provide model fits to the observed SED (upper) and H-band radial polarized intensity profile (lower)	282
A.2	Sample model outputs for a representative undepleted small grain disk with a scale height discontinuity at the 36AU large grain dust wall. The lefthand panels represent the density of large grains (upper) and small grains (lower). The righthand panels provide model fits to the observed SED (upper) and H-band radial polarized intensity profile (lower)	283
A.3	Sample model outputs for a representative model in which the region of the sub-mm cavity is more steeply curved in the small grain disk. The lefthand panels represent the density of large grains (upper) and small grains (lower). The righthand panels provide model fits to the observed SED (upper) and H-band radial polarized intensity profile (lower).	284
B.1	Best-fitting Whitney model SED for Wood et al. (2002) Model 1 grains. This provides a better fit at 10μ m, but the IRS spectrum is poorly matched and the corresponding model imagery, shown in Figure B.2, is a much poorer fit to our observations	286
B.2	K-band model image for Wood et al. (2002)'s Model 1 grains convolved with our PSF (left) and radial profiles taken through the same image (right). The H-band model and profiles are qualitatively similar. This model, though a better fit to the SED, is too bright in the South relative to the North, as revealed by the relative brightness of the model profiles along the major axis and to the south and north.	

C.1	Radial profiles taken through disk images deprojected to account for an inclination of 50° along PA=97.4°. Profiles were taken both with respect to the center of the circular K-band cavity (left column), and from the stellar location (right column). In all cases, solid lines represent profiles through the Eastern half of the disk and dashed lines represent profiles through the Western half of the disk. Black profiles were taken along the disk major axis and blue and red profiles were taken 20° North and South of the major axis, respectively. (a) and (b) Radial profiles through the Ksband halo-subtracted and deprojected PI image. Profiles have been binned to 6 pixels (0.06") radially, equivalent to half of the stellar	
	FWHM.; (c) and (d) Radial profiles through a smoothed Ks-band halo-subtracted and deprojected PI image. The original image was smoothed with a 5.65 pixel (0.05") Gaussian in order to match the stellar FWHM of the H-band image and allow for a direct comparison of the Ks and H-band profiles. Profiles have been binned to 12 pixels/0.11" to match the H-band data.; (e) and (f) Radial profiles through the H-band halo-subtracted and deprojected PI image. Profiles have been binned to 12 pixels/0.11" radially, equivalent to half of the stellar FWHM	291
D.1 D.2	Legend applying to all subsequent figures in this appendix Response distributions for astronomy and numeracy educators normalized to the number of respondents in each group for the six skills deemed most important by both groups	
D.3	Response distributions for astronomy and numeracy educators normalized to the number of respondents in each group for skills ranked	294
D.4	Response distributions for astronomy and numeracy educators normalized to the number of respondents in each group for skills ranked	
D.5	13-18 in importance	

LIST OF TABLES

2.1 2.2	Position Angle Estimates
3.1 3.2	Previously published photometry and spectra of Oph IRS 48 127 Relevant Whitney Model Input Parameters for Oph IRS 48 136
4.1	f_d represents the fraction of the disk's dust mass in large grains, Z_{100} is the scale height at 100 AU, α is the exponent for the radial midplane density, and β is the exponent for the radial scale height. The stellar and extinction parameters are from Hillenbrand (1997); the disk mass is from Mann & Williams (2010). All other parameters were derived from our observations as described in the text with the exception of Z_{100} , for which the model was found to be insensitive to a range of reasonable parameters (values 5-10AU), and $r_{disk,in}$, which was determined iteratively from the fit to existing J-L' photometry. The large grain dust prescription is Model 1 of Wood et al. 2002, a mixture of amorphous carbon and astronomical silicates with a maximum grain size of $1000\mu m$. The small grain dust prescription is the ISM grain model of Kim et al. 1994, a mixture of silicate and graphite with a maximum grain size of $0.28\mu m$
6.1	Common mathematical misconceptions encountered frequently in our classrooms
7.1 7.2 7.3 7.4 7.5 7.6 7.7	Reason for Choosing Major by Major216Number of Questions by Skill and Difficulty236Instrument Evolution246Item Statistics248Completion Statistics256Filter Statistics262Effort Statistics266
A.1	SIRPOL Aperture Polarimetry
E.1 E.2	QuaRCS Demographic Questions

ABSTRACT

What processes are responsible for the dispersal of protoplanetary disks? In this dissertation, beginning with a brief Introduction to planet detection, disk dispersal and high-contrast imaging in Chapter 1, I will describe how ground-based adaptive optics (AO) imaging can help to inform these processes.

Chapter 2 presents Polarized Differential Imaging (PDI) of the transitional disk SR21 at H-band taken as part of the Strategic Exploration of Exoplanets and Disks with Subaru (SEEDS). These observations were the first to show that transition disk cavities can appear markedly different at different wavelengths. The observation that the sub-mm cavity is absent in NIR scattered light is consistent with grain filtration at a planet-induced gap edge.

Chapter 3 presents SEEDS data of the transition disk Oph IRS 48. This highly asymmetrical disk is also most consistent with a planet-induced clearing mechanism. In particular, the images reveal both the disk cavity and a spiral arm/divot that had not been imaged previously. This study demonstrates the power of multiwavelength PDI imaging to verify disk structure and to probe azimuthal variation in grain properties.

Chapter 4 presents Magellan visible light adaptive optics imaging of the silhouette disk Orion 218-354. In addition to its technical merits, these observations reveal the surprising fact that this very young disk is optically thin at H-alpha. The simplest explanation for this observation is that significant grain growth has occurred in this disk, which may be responsible for the pre-transitional nature of its SED.

Chapter 5 presents brief descriptions of several other works-in-progress that build on my previous work. These include the MagAO Giant Accreting Protoplanet Survey (GAPlanetS), which will probe the inner regions of transition disks at unprecedented resolution in search of young planets in the process of formation.

Chapters 6-8 represent my educational research in quantitative literacy, beginning with an introduction to the literature and study motivation in Chapter 6. Chapter 7 describes the development and validation of the Quantitative Reasoning for College Science (QuaRCS) Assessment instrument. Chapter 8 briefly describes the next steps for Phase II of the QuaRCS study.

CHAPTER 1

Introduction

Exoplanets were largely synonymous with science fiction until the discovery of the planet 51 Peg b in 1995 (Mayor & Queloz 1995). In the nearly two decades since the groundbreaking radial velocity exoplanet discoveries of the late 1990s and early 2000s, astronomers have discovered not just a huge number of exoplanets (1516 confirmed as of October 14, 2014), but have also uncovered entirely new flavors of planets without solar system analogs (e.g. Hot Jupiters, Super-Earths, Super Jupiters). In addition to the now well-established radial velocity detection method, astronomers have pioneered a number of alternative methods for finding exoplanets, and each method probes a different region of planetary parameter space. In this introduction, I will provide a brief context for the work presented in this dissertation by outlining a few key aspects of the fields of exoplanet formation, circumstellar disk imaging, and high-contrast adaptive optics.

1.1 Exoplanet Discovery Methods

As of 2014, the four successful exoplanet detection methods are as follows:

1. Radial Velocity. Planets are detected through high spectral resolution monitoring of nearby stars for periodic Doppler shifts that correspond to the gravitational tug of an orbiting planet. This method is most successful at detecting the larger gravitational influence of massive companions, and outer planets with long orbital periods are only detectable through decades-long monitoring. This method has been very successful and has contributed robust statistics regarding the prevalence of massive, nearby planets, including the surprising occurrence of Hot Jupiters. RV techniques are limited,

however, in their ability to detect signals from low mass planets by our current ability to remove noise caused by stellar activity. It remains to be seen whether this technique will be able to achieve the $\sim 10 \, \text{cm/sec}$ precision necessary to detect the effect of earth-like planets in the habitable zones of their host stars, although promising techniques are being developed toward this end.

- 2. Transits. Planets are detected through the periodic dimming of their host star as the planet passes in front of it. Pioneering ground-based transit detections in the early 2000s (e.g. Charbonneau et al. 2000; Henry et al. 2000; Konacki et al. 2003) paved the way for the Kepler spacecraft mission, which has drastically increased the number of planets detected in this way. This technique is also best-suited for detecting large radii inner planets, whose transits will be deeper and more frequent, however the exquisite photometric sensitivity, low noise level, and five year mission of Kepler allows for the detection of transits as shallow as 100 or fewer ppm and semimajor axes of 1AU and beyond. Through pairing with radial velocity detections and/or by detecting transit timing variations, a number of transiting exoplanet densites are now known, allowing us for the first time to probe planetary composition.
- 3. Microlensing. Photometric monitoring of stars in our Galaxy occasionally reveals brief brightening events that correspond to the gravitational lensing of a background star. Occasionally, these brightening events exhibit peaks that correspond to an additional lensing by a planet orbiting the primary stellar lens. This technique is much less sensitive to the mass and separation of planetary companions than other methods, and therefore might of-

fer the most complete demographics as the sample size increases, however confirmation and characterization of these planetary companions is made impossible by the low probability of a successive lensing event for the same star.

4. Direct Imaging. The development of adaptive optics technologies, which allow for greatly improved ground-based image stability, as well as advancements in coronagraphy and post-processing techniques, have allowed astronomers in recent years to isolate emission from exoplanets. Unlike the other planet detection methods, which rely on observations of the host star to infer planetary properties, direct imaging is just that - direct. Direct thermal emission from the planets themselves is detected and the spectra of exoplanets can be explored. This technique is highly complementary to other planet detection techniques in that it probes preferentially the outer reaches of planetary systems.

With the combination of decades-long radial velocity campaigns, the Kepler mission's tremendous success at monitoring exoplanet transits, global followup of microlensing events, and large ground-based direct imaging surveys, all four of these exoplanet detection methods have reached the realm of sample statistics, rather than mere stamp collecting. Particularly exciting in the coming decades will be the point when radial velocity and transit surveys moving from the inside out and direct imaging surveys moving from the outside in are able to push to the same planetary separation regimes. Eventually, this will allow astronomers to get a handle on the demographics of entire planetary systems, and characterize planets in those systems using multiple methodologies, rather than just sampling the very inner planets or the very outer planets with separate methods.

1.2 Circumstellar Disk Imaging

As the field of exoplanetary astronomy grows and more and more planets are detected, some of the most pressing unanswered questions are where, how, and from what material extrasolar planets form. Although slightly less "sexy" in igniting the imagination of astronomers and the public alike, the field of circumstellar disks is where these questions will be answered.

1.2.1 Disk Formation

Circumstellar disks are an artifact of the star formation process, and a natural consequence of angular momentum conservation within the dense star-forming cores of molecular clouds. As the central core of the cloud collapses to form a star, it spins faster to conserve angular momentum and heats up to conserve energy during collapse. Both changes serve to increase collisions between atoms and molecules within the core. Their random motions perpendicular to the spin axis will tend to cancel one another out, while their coherent motion in the spin direction will tend to add, and the disk will broaden and flatten, eventually forming planets that are roughly coplanar.

Estimates of primordial disk masses are complicated by the fact that these disks are optically thick at most wavelengths. One theoretical method for estimating disk mass involves using the current state of our own solar system to place constraints on the mass of primordial material available to form it. This approach, called the minimum mass solar nebula, uses the observed composition of solar system planets scaled by cosmic abundances to estimate the amount of material in the protosolar disk, and it is strictly a lower limit on the mass, as much of the material in the protosolar nebula was likely dispersed as the disk evolved (as described in section 1.2.4).

Observed Spectral Energy Distributions (SEDs) of young disks can also be used to estimate the total disk mass. In particular, the integrated flux at mm wavelengths is a measure of the amount of dusty material in the optically thin outer disk. By assuming a gas:dust mass ratio (often 100), this value can be translated to a total disk mass. For the youngest disks, these values are around $0.1 M_{\odot}$ (Andrews & Williams 2005).

Evidence for the presence of circumstellar disks is abundant, both through the presence of infrared excess in the SEDs of young stars (a result of the reprocessing of stellar emission by dusty circumstellar material), and through direct imaging at a variety of wavelengths. The disk location, dust/gas species, temperature, and physical emission mechanism vary widely according to wavelength. For example, sub-mm continuum imaging probes thermal emission from large dust grains near the disk midplane where the majority of the disk mass resides. Line emission from gaseous species in the disk such as CO are also present in this wavelength regime. In the optical and NIR, disk emission is dominated by scattered light from small dust grains near the disk surface layers. A small number of disks in nearby star forming regions are also seen in silhouette against the bright nebular background, as described in detail in Chapter 4.

The fraction of young stars in nearby star forming regions showing evidence of circumstellar material approaches 100% in the youngest regions, but drops rapidly as these clusters age. At 5 Myr, fewer than 50% of young stars in these clusters show evidence for disks, and the proportion approaches zero at >10Myr, though the scatter in large, perhaps due to the unreliability of stellar age estimates for young stars. This trend holds true whether the indicator of disk fraction is an inner disk (NIR excess) or outer disk (mm excess) tracer, thus it would appear that clearing of the inner and outer disk happens simultaneously.

The results discussed in Section 1.1 reveal the ubiquity of exoplanetary systems. Modern techniques have been particularly successful at detecting gas giant planets, which require that a large amount of disk gas be present during formation. Thus, circumstellar disks must succeed in forming planets with regularity. The total mass remaining in planets after dispersal of the disk is very small (at most tens of Jupiter masses) when compared to the mass of primordial disks. The process of planet formation from circumstellar disk material is thus both inefficient, incorporating only a small fraction of the primordial disk mass into planets, and rapid, as it must occur within the observed $\sim \! 10 \, \mathrm{Myr}$ disk lifetime.

1.2.2 Disk Classification

Disks are most easily identified observationally through infrared excess in the SEDs of young stars. Although all disks share the observational trait of infrared excess, the amount and wavelength of the excess varies significantly from disk to disk. In particular, very young so-called "protoplanetary" disks exhibit strong infrared excess at all wavelengths, and their SEDs are roughly flat from the stellar blackbody peak out to the mm regime. Other young star systems, particularly in older star forming regions, have a double peaked SED profile in which disk emission picks up only in the mid-infrared. This excess only at long wavelengths and a simultaneous lack of gas tracers in these disks has led to the conclusion that they host less massive disks of tenuous dusty debris, called "debris disks.

It is believed that protoplanetary disks become debris disks through an evolutionary process in which gas and dust is dispersed and/or incorporated into forming planets, clearing their thick disks of material. This clearing process reduces the disk mass from a few Jupiter masses for protoplanetary disks (Taurus median $\sim 5 \times 10^{-3}$ M $_{\odot}$ Andrews & Williams 2005) to fractions of an earth mass for debris disks (~ 0.001 -0.1M $_{Earth}$ Panić et al. 2013). However, it is still unclear

precisely how one type of system becomes the other. By studying systems that host disks, and in particular those most likely to also host young planets, we can begin to answer these questions.

A subclass of disks termed "transitional disks is a particularly relevant population to study in the context of disk evolution because transitional disks appear to be, as their name suggests, in transition between a young, gas-rich protoplanetary disk phase and an older gas-poor debris disk phase, though uncertainty in ages for young stars has made it difficult to determine unequivocally whether transition disks are a true part of this sequence or a separate evolutionary pathway. This class of disks was first identified through their SEDs as sources with disk excess emission at most long wavelengths, but without NIR excess. This has been interpreted as an indication of the dispersal of the innermost warm NIR-emitting regions of the disk, as represented schematically in Figure 1.1.

In recent years, technological advancements in radio interferometry and high contrast adaptive optics imaging have allowed us to resolve cleared inner gaps in many nearby transitional disks at both sub-mm (e.g. Andrews et al. 2011, 2010) and near infrared (NIR, 1-5 μ m, e.g. Mayama et al. 2012; Thalmann et al. 2010) wavelengths, and their morphologies in the two wavelength regimes are often quite discrepant. For example, as will be explored in Chapter 2, the transitional disk SR21 has been shown to host a large (r \sim 36AU) heavily-depleted inner clearing in the sub-mm (Andrews et al. 2011). However, 1.6 μ m NIR images show no evidence of a cavity (Follette et al. 2013b). Many other disks reveal similar morphological discrepancies at different wavelengths (e.g. Dong et al. 2012).

If transition disks represent a common evolutionary pathway, then their low incidence in start forming regions (<10%) suggests that the clearing phase must be rapid - on the order of 1Myr.

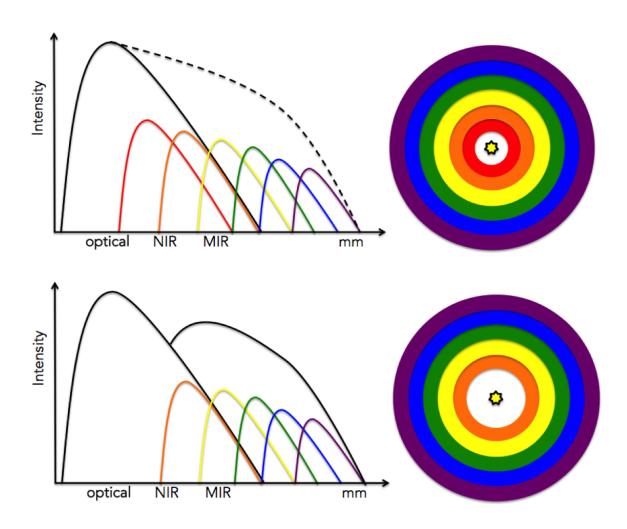


Figure 1.1: Schematic representation of a full disk with annular regions that emit as black-bodies at successively longer wavelengths toward the cooler outer regions. If the innermost disk component is removed, it manifests itself as a lack of excess in the NIR.

1.2.3 Disk Structure

The radial extent of circumstellar disks is well constrained by observations. Silhouettes, thermal sub-mm emission and scattered light disks have all been observed to extend to hundreds of AU in radius. This presents somewhat of a conundrum in the context of planet formation as all of the planets in our own solar system orbit at radii of less than 50AU. It is unclear whether the protosolar disk was somehow truncated, whether the outer regions of disks are universally incapable of sustaining planet formation or whether the current state of our solar system is a result of the dynamical evolution of planets that formed at larger disk radii.

The vertical extent of a protoplanetary disk at a given radius is set by the balance between the vertical component of the stellar gravitational force, which accelerates particles toward the disk midplane, and the opposing pressure force caused by the vertical pressure gradient from the dense disk midplane to the more tenuous upper layers.

In a vertically isothermal disk, this relationship amounts to a simple equation for the scale height h of the gas disk, namely:

$$h = \frac{c_s}{\Omega} \tag{1.1}$$

where Ω is the Keplerian angular velocity $\sqrt{GM_*/r^3}$ and c_s is the isothermal sound speed determined by:

$$c_s^2 = \frac{k_B T}{\mu m_p} \tag{1.2}$$

where μ is the mean molecular weight in units of the proton mass m_p . The net result of this relationship is that as radius in the disk increases, the Keplerian angular velocity Ω falls and the scale height increases. The precise relationship between scale height and radius is a function of the change in sound speed with

radius. In particular, if sound speed follows a power law such that:

$$c_s \propto r^{-\beta}$$
 (1.3)

then the ratio of the disk scale height to the radius (h/r) varies as:

$$\frac{h}{r} \propto r^{-\beta + 1/2} \tag{1.4}$$

Assuming that the mean molecular weight of the disk doesn't change with radius (which may not be an entirely reasonable assumption), then the scale height is primarily a function of the disk midplane temperature. If the temperature falls off as 1/r or shallower, then the ratio h/r will increase with radius and the disk will be flared. This shape has implications for both the SED and the rate of disk dispersal as it affects the amount of disk surface at a given radius on which stellar light is directly incident.

1.2.4 Disk Evolution

The decline in disk fraction with time, as well as observed accretion signatures in young stars, make it clear that disks must evolve with time. In order to flow onto the star as observed, gas in the disk must lose angular momentum through viscous processes. For terrestrial fluids, the dominant mechanism for diffusion is molecular collisions, however the mean free path under protoplanetary disk conditions is too long for this collisional mechanism to dominate disk evolution on the scale of observed disk lifetimes. For the purposes of numerical simulations, the turbulent viscosity that drives disk evolution is generally defined as:

$$\nu = \alpha c_s h \tag{1.5}$$

where α is a dimensionless quantity called the Shakura-Sunyaev α parameter. It's precise value depends on the underlying turbulent mechanism(s), and this is still

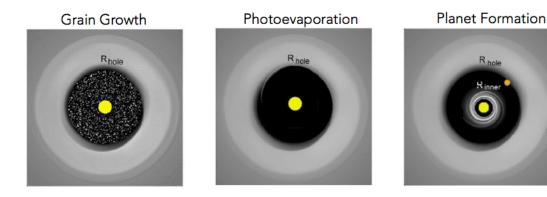


Figure 1.2: Schematic representation for the three main candidates for the clearing of transitional disk cavities. Each makes its own predictions about the degree of clearing inside of the disk cavity

a very active area of research. Candidates include various forms of turbulence such as Ohmic dissipation, the Hall effect and Ambipolar diffusion, all of which rely on charged particles within the disk. Because ionization is expected to occur mostly in the surface layers of disks (though contributions from more penetrating cosmic rays as well as radioactive decay may also be relevant), dissipation of angular momentum may be mainly a surface effect and disks may accrete primarily through active surface layers, with a quiescent "dead zone" near the disk midplane.

1.2.5 Clearing Mechanisms

Even among the small number of resolved disks, there is a wide range in their sizes, morphologies and compositions. The extent to which these variations are evolutionary, and what physical processes in the disk drive this evolution, is unknown. I outline each of the three most prominent theories invoked to explain the presence of transitional disk gaps below.

1. Photoevaporation

Young stars are strong emitters of high energy radiation due to elevated levels of chromospheric activity and high-energy accretion processes. This radiation will tend to photoionize gas in the upper layers of the disk atmosphere. Once ionized, gas will flow along the open magnetic field lines that thread the disk, transporting both mass and angular momentum away from the disk.

The precise level in the disk atmosphere where photoevaporation occurs depends on the geometry of the disk (flared vs. flat) and the nature of the high energy radiation. In particular, photoevaporation driven by x-rays, far ultraviolet (FUV) and extreme ultraviolet (EUV) radiation each drive winds at different column densities and therefore at different levels in the disk atmosphere and at different rates.

Generally speaking, however, the process occurs in the same way regardless of the precise mechanism. At the smallest disk radii, the atmosphere of the disk will tend to be bound gravitationally despite a high degree of photoionization, however beyond a certain characteristic radius, equal to:

$$r_g = \frac{GM_*}{c_*^2} \tag{1.6}$$

the orbital velocity of ionized gas falls below the sound speed (c_s) and the gas becomes unbound. For a photoionized gas with a temperature of $10^4 \mathrm{K}$ and a sound speed $c_s \sim 10 \mathrm{km/sec}$ around a solar mass star, this gravitational radius corresponds to 8.4AU. However, a more sophisticated treatment of photoevaporation suggests that photoevaporative winds are likely launched from much closer - just $0.2r_g$ - corresponding to a few AU from the central star. The now isolated inner disk will fall onto the star on a viscous

timescale, and once this happens, the high-energy flux from the star will be directly incident on the outer disk, increasing the rate of photoevaporation and rapidly clearing the disk from the inside out.

Photoevaporation only becomes a dominant mechanism in disk evolution once the rate of replenishment of disk material falls below the wind loss rate, thus young actively-accreting disks are unlikely to be affected by this process. The question is not truly whether photevaporation occurs in disks - it inevitably does - but when it begins, whether it is driven primarily by EUV, FUV or x-ray radiation, and, most relevant for the purposes of this thesis, whether it is the dominant mechanism creating the central clearings in transitional disks.

The predictions of photoevaporative clearing suggest that the size and symmetry of the cavity may be a good indication of whether or not this process dominates. In particular, it is difficult to create wide cavities, asymmetric cavities and cavities in disks whose central stars are still actively accreting, through photoevaporation.

2. Grain growth

The orbital velocities of solid materials in protoplanetary disks are dominated by gravity, and they will naturally follow Keplerian orbits around the central star. Gaseous particles, however, feel an additional outward force due to the pressure gradient from the inner hot dense regions of the disk to the outer cool and less dense regions. Gas in a protoplanetary disk will therefore orbit the star at sub-Keplerian speed.

Solid particles are affected by their differential speed with respect to gas at the same orbital radius, however the effect varies according to particle size. Small particles are well coupled to the gas due to their short stopping times, and they are dragged along at sub-Keplerian speed, however as they do not experience the same outward pressure support, their velocities are not sufficient to balance inward centrifugal force and they drift slowly inward. Larger particles are not well coupled to the gas and they experience the relative velocity as a headwind as they plow through the slower orbiting gas particles at the same radius. This results in loss of energy and angular momentum to the gas, and these larger particles spiral inward. This effect dominates the motion of large particles until they grow to sizes beyond the mean free path of the gas (\sim 10cm or more for most disk regions), and inward drift speed peaks at \sim 10cm-1m. This makes it difficult for particles of this size to remain in the disk, and is referred to as the "meter size barrier".

A similar process occurs vertically in the disk, where solid particles feel a gravitational pull toward the disk midplane but do not feel the pressure force described in section 1.2.3. The speed with which a solid particle settles toward the disk midplane depends on the turbulence prescription, but the general tendency is for larger particles to have larger vertical velocities toward the midplane. They will thus tend to sweep up smaller, slower moving particles in their path as they settle toward the midplane. For particles in the μ m size regime, coagulation to larger sizes through this mechanism is rapid and particles on the order of \sim 1mm in size quickly populate the disk midplane.

The process of grain coagulation should be most efficient where particles are abundant and the vertical component of the gravitational force is largest, so in the inner disk. If particles in the inner disk are able to grow to mm sizes rapidly, then this could be part of the explanation for missing NIR ex-

cess in the SEDs of transitional disks. This mechanism therefore predicts that transitional disk cavities will be populated primarily by large grains.

3. Planet formation

Under the core accretion model of planet formation, a forming protoplanet will begin to sweep up disk material within its orbital annulus. The width of this annulus will be on the order of the so-called "Hill radius", which is the radius inside of which the gravitational influence of the planet dominates over the shear in the Keplerian disk and material can remain bound to the planet. This radius is represented as:

$$r_H = (\frac{M_p}{3M_*})^{1/3}a\tag{1.7}$$

Even for a 10 M_J planet, this corresponds to less than 0.5 AU at an orbital radius of 5 AU around a 1M_{\odot} star, so even if a single massive planet can accrete particles within several Hill radii of its orbit (which numerical simulations suggest it can, e.g. Dodson-Robinson & Salyk 2011), it is difficult to explain why we observe cleared inner holes in transition disks rather than annuli at intermediate radii. It may be that multiple planets are needed to maintain large (tens of AU) observed gaps.

If the opening of disk gaps was able to isolate the inner disk from replenishment by material in the outer disk, then it would disappear on a viscous timescale and create a cavity, however numerical simulations of gap opening by planets suggest that significant amounts of material make it past the planet and into the inner disk (Zhu et al. 2012). These so-called "dust-filtration models" predict that the degree of clearing inside of a planet-induced disk gap will vary according to the size of the particle, being successively less well cleared for smaller particles, as shown in Figure 1.3.

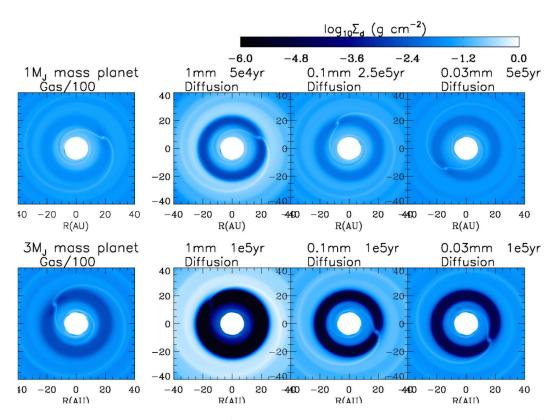


Figure 1.3: Numerical simulations of gap opening by a single planet showing the differences in gap size and degree of clearing according to both the size of the planet and the size of the grains in question. Larger planets open wider, more cleared cavities. Larger grains are most efficiently cleared, while small grains are less affected. These simulations extend only to grain sizes of 0.03mm. For grains an order of magnitude smaller - the μ m sized grains probed by NIR scattered light imaging - the cavity all but disappears. Reproduced from Zhu et al. (2012).

Any model that purports to explain the mechanism through which these clearings are generated must be able to explain differences in the observed distribution of grains at varied wavelengths. Multiwavelength high-resolution imaging of transitional disks is thus essential in order to understand the exact mechanism that is responsible for clearing them, and I will give examples of such imaging campaigns in Chapters 2 and 3 of this dissertation.

Furthermore, disentangling the competing theories for gap formation involves resolving the innermost part of the disk, where planet formation is likely to occur, and this requires high resolutions and effective removal of residual starlight. In recent years, much indirect evidence has accumulated suggesting that planet formation is, if not the dominant mechanism for clearing these cavities, at least a contributor. In particular, a wide range of asymmetries most consistent with a planet-induced origin have been observed in transition disks, including warps, divots, spiral arms, dust traps and pericenter offsets. A few of these non-axisymmetric features are highlighted in Figure 1.4.

Although there has yet to be a successful, independently confirmed direct detection of a planet inside of a transitional disk clearing (though there are several intriguing candidates, notably those of Kraus & Ireland (2012) and Quanz et al. (2011)), imaging campaigns such as that contained in Chapter 2 and Chapter 3 highlight the power of high contrast imaging to detect and characterize potential planet-induced features of transitional disks.

1.3 High-Contrast Direct Imaging of Young Exoplanets and Circumstellar Disks with Modern Adaptive Optics Systems

This dissertation will focus on the direct imaging method and its utility not just for discovering and characterizing massive outer planets, but also for character-

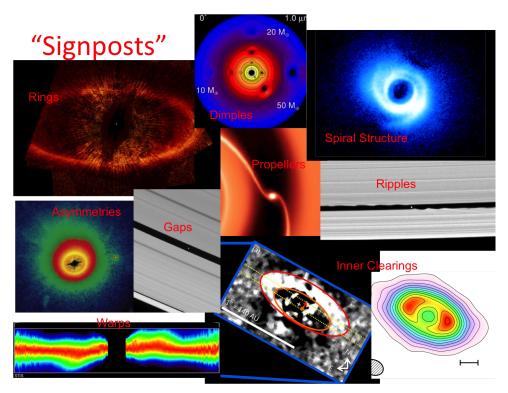


Figure 1.4: A selection of observations, model images and solar system analogs of features that may be indicative of the influence of planetary perturbers in circumstellar disks, many of which are non-axisymmetric.

izing the circumstellar disks in which those planets are formed. I provide in this section an extremely abbreviated description of the history of the field, focusing on developments relevant to the work presented herein.

1.3.1 Adaptive Optics

Many recent successes in the field of direct imaging were made possible through the development of adaptive optics technologies, and The University of Arizona has played a critical role in the development of AO techniques and technologies. To summarize the entire field in a single paragraph: Adaptive Optics (AO) imaging utilizes deformable mirrors (DMs) to correct for the effect of atmospheric turbulence/image blurring by diverting some portion of the light entering the instrument into a wavefront sensor (WFS). The WFS measures the deformation of the wavefront from either a bright star in the field, a Natural Guide Star (NGS), or from an artificial Laser Guide Star (LGS), and signals the DM to deform in such a way as to cancel out the corrugation of these wavefronts. Upon reflection off the DM, wavefronts from the science object are flattened and the image is stabilized, allowing us to image it at much higher resolution, separation and contrast than is possible with natural seeing. AO systems are complex, and several choices must be made in their design, including the nature of the DM and the WFS. In terms of the AO systems used to complete the current and future work described in this thesis, I will briefly describe several particularly relevant AO technologies here.

1.3.2 Deformable Mirrors and Wavefront Sensing

Deformable mirrors fall roughly into two categories - large deformable secondary mirrors and smaller deformable mirrors that lie much farther along the optical path, although there are several varieties. The University of Arizona has invested heavily in the development of Adaptive Secondary Mirrors (ASMs) and ASMs

are currently installed on the MMT, LBT and Magellan telescopes. Advantages of ASMs include:

- 1. By repurposing the secondary mirror for adaptive optics correction, no additional reflecting surfaces are required. Since aluminum reflective optics are at best 98% efficient, and in practice more like 95%, this maximizes optical throughput. The light entering the MagAO WFS, for example, has undergone only five reflections (and only 3 for the IR science camera Clio2). Assuming 95% reflectivity for each optic, this amounts to an optical throughput of 77% The optics of the Keck AO system, on the other hand, require at least 9 reflections before the light enters the WFS, lowering the optical throughput to 63%.
- 2. The thermal emissivity of physical optics also contributes to thermal background emission, particularly at NIR wavelengths. This contribution scales linearly with the number of optics, so fewer optics are also an advantage here (ASMs have ~3 times lower emissivity, which allows for an equal increase in exposure time at thermal wavelengths before sky saturation).
- 3. Due to their large physical size and the flexibility of ASM thin shells, the stroke of each actuator is large, allowing for correction of large amplitude aberrations in the wavefront.
- 4. The large size of an ASM also allows for preservation of high etendue and a correspondingly wide field of view (FOV), as for a given f-ratio beam, FOV is roughly proportional to the size of the optic.

Non-secondary AO mirrors, such as those installed in GPI and on Subaru, are optically flat, smaller, and cheaper, and there are several varieties. Particularly

promising in this arena are MicroElectroMechanical Systems (MEMS) Mirrors, which use silicon micromachining technologies pioneered in the semiconductor industry to create very small and potentially low cost tertiary mirrors that are actuated through signals applied to electrostatic parallel plates rather than through physical actuation. The Gemini Planet Imager uses a MEMS mirror as a "tweeter" to complete high spatial frequency correction. However, the small stroke capability ($\sim 3\mu$ m peak to peak) of the MEMS requires a second "woofer" mirror to correct larger amplitude wavefront variations.

There are also a number of available options for Wavefront Sensing architecture, but the AO systems used for this work fall into just two categories. The "classic" wavefront sensing architecture is called a Shack-Hartmann Wavefront Sensor (SHWFS). It relies on a grid of lenslets placed in the pupil plane to create a corresponding grid of subapertures. The signal in and spacing of the "spots" created by each subaperture are used to reconstruct the local slope of the wavefront. This is a well-tested and very successful architecture for wavefront sensing.

One disadvantage of the Shack-Hartmann architecture is that because the lenslets are physical optics, they cannot be binned, and so there is a fundamental limitation to how bright the guide star must be in order to have a detectable amount of signal in each subaperture. Choosing the lenslet pitch for a Shack-Hartmann wavefront sensor is therefore a balance between the desired level of correction, which tends to require more lenslets, and the desired limiting magnitude of the guide star, with fainter limits requiring fewer lenslets.

Arcetri Observatory and The University of Arizona have pioneered the implementation of an alternative architecture called the Pyramid Wavefront Sensor (PWFS, Ragazzoni 1996). Pyramid WFS focus the telescope beam on the tip of a glass pyramid, creating four images of the telescope pupil on a fast readout

CCD. The focused beam is modulated rapidly (~100Hz-1kHz) around the tip of the Pyramid so that large amplitude wavefront aberrations are dispersed across all four pupils and do not saturate. This increases the dynamic range of the sensor and makes intensity a linear function of wavefront slope over a wider range of amplitudes. Different optical wavefront aberrations manifest themselves differently across these subapertures and the aberrated wavefront can be deconstructed into a series of modes and measured across these subapertures. Advantages of PWFS include:

- 1. With a diffraction limited spot on the tip of the pyramid, the PWFS is sensitive to tip/tilt errors on the order of λ/D , where D is the diameter of the primary mirror. For a SHWFS, tip/tilt can only be corrected to λ/d , where d is the size of each subaperture (corresponding to the pitch between lenslets mapped to the primary mirror). For visible wavelength imaging in particular, this additional tip/tilt stability is essential since D is much greater than d.
- 2. The contrast achieved by many AO cameras is limited by the sampling of the wavefront, and only spatial frequencies greater than or equal to twice the pitch of the actuators (44cm for MagAO) can be measured and corrected via Nyquist sampling. A particular problem is the likelihood that undersampled high spatial frequency wavefront aberrations might manifest themselves as low spatial frequency modes a problem called "aliasing". Because the tip of the pyramid is at a focal plane and each face of the pyramid is a continuous sheet, the wavefront is not sampled until the four pupils are imaged on the camera, so the detector pixel size determines the sampling in a PWFS. This generally allows for fewer aliasing problems, as evidenced by the tremendous success of the LBT and Magellan AO systems

in achieving a "dark halo". This dark central region surrounding the core of the PSF, shown in Figure 1.5, represents a well sampled (non-aliased) region of parameter space corresponding to large spatial frequencies. The edge of the dark halo corresponds to the "control radius" of the system - the limiting size of spatial frequencies (~44cm in the case of MagAO) that can be controlled with the system.

3. The PWFS architecture means that binning of subapertures takes place on the CCD, after the beam has passed through the pyramid, and it can thus be varied according to the faintness of the guide star. This means that compromises between the amount of signal in each subaperture and the degree of correction can be made and varied in real time, as opposed to a SHWFS where these choices are fixed, and rebinning requires inserting a different optic into the beam.

1.3.3 MagAO and Subaru AO Systems

Two systems - MagAO and Subaru - were used to conduct the bulk of the work described herein, so I will briefly outline the relevant capabilities of each here.

For the MagAO system (Close et al. 2013, 2014; Males et al. 2012; Follette et al. 2010), I have worked primarily with its visible light camera VisAO. It is a system designed to make use of the exceptionally high number of controllable modes of the Magellan ASM, which map to a 22cm pitch on the 6.5m primary mirror, in order to push Adaptive Optics into the visible wavelength regime. The schematic shown in Figure 1.6 highlights the optics of the visible wavelength system.

Two of the chapters in this thesis, Chapters 2 and 3, are related to my work as a member of the Strategic Exploration of Exoplanets and Disks with Subaru (SEEDS, Tamura 2009) team, for which I have used Subaru's High-Contrast Im-

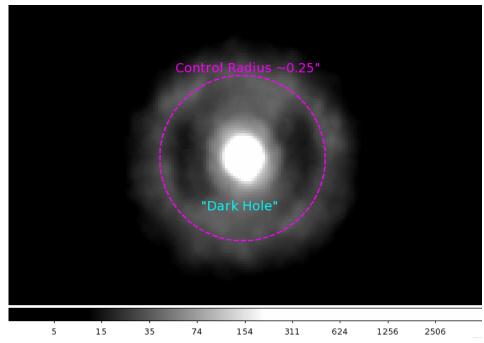


Figure 1.5: A sample MagAO H α PSF, showing a clear dark hole inside of the system's control radius at 0.25". This is the region of Fourier space corresponding to well-sampled and corrected wavefront aberrations

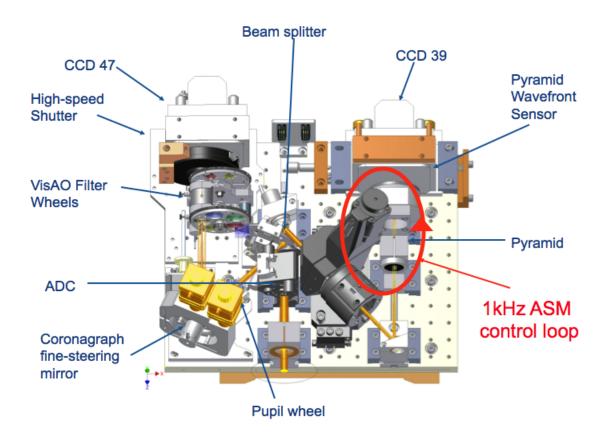


Figure 1.6: The mechanical design of the Magellan VisAO and PWFS system, with the key features for visible AO science indicated by blue arrows. The red loops indicate the two control loops for wavefront correction. Light from the ASM enters from the bottom.

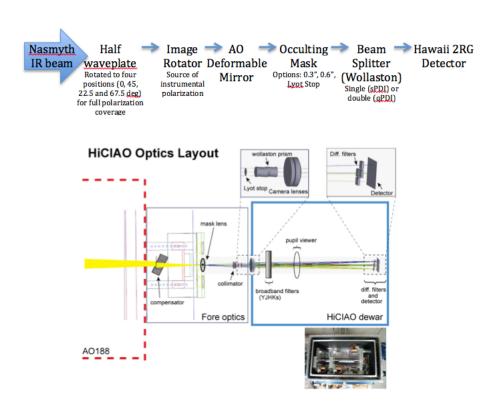


Figure 1.7: Top: Schematic of the Subaru + AO188 + HiCIAO optical system. Bottom: More detailed schematic of the HiCIAO optics.

ager for Adaptive Optics (HiCIAO) camera and 188-actuator tertiary AO system (AO188) in Polarized Differential Imaging (PDI) mode. The optical architecture of this mode is described in Figure 1.7. Principally, it relies on a Wollaston prism to split the beam into two polarization states, and these are modulated by rotating a half-waveplate to four different angular positions. The principal choices made by observers are whether to use an occulting mask and whether to use a single or double Wollaston to split the beam into two or four channels, respectively.

1.4 Differential Imaging Techniques

Nearly all circumstellar science relies on one or more forms of differential imaging in order to isolate disk or planet emission from the direct stellar emission that dominates in most cases. The work in this thesis utilized three forms of differential imaging, each of which are described briefly below.

1.4.1 Simultaneous Differential Imaging

Simultaneous Differential Imaging (SDI) mode is a differential imaging technique that relies on simultaneous acquisition of a science image at a wavelength of interest and in another nearby wavelength where the phenomenon of interest will not be visible. For example, SDI has been attempted on extrasolar planets in and out of the methane band at $1.6\mu m$, where their atmospheres were expected to have significant absorption, though this is proving to be largely not the case. In principle, continuum objects will have approximately equal fluxes in the two filters, while giant planets and brown dwarfs with methane in their atmospheres should presumably "drop out" at the methane band. In MagAO, SDI mode operates through a Wollaston Prism that is placed in the VisAO beam and functions as a 50/50 beamsplitter for unpolarized sources. Each half of the beam is passed through a different narrowband filter, one on a spectral line of interest and one on the neighboring continuum (see Figure 1.8). The continuum channel functions as a simultaneous probe of the point-spread function (PSF) of the object of interest at a nearly identical wavelength, making subtraction and removal of continuum emission remarkably simple. The robustness of SDI-mode PSF subtraction allows us to probe the disk closer to the central star ($r\sim50$ mas) than can be achieved by any other AO imager, and the visible wavelength regime allows for higher resolutions than are available elsewhere. This opens up a window into disk regions that have previously been accessible only through interferometry, and never at visible wavelengths. There are currently three SDI filter sets (H α , [OI] and [SII]), each of which provides a unique view on a particular process believed to occur in young stars.

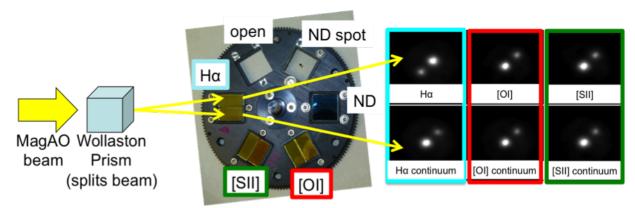


Figure 1.8: A schematic of MagAOs SDI imaging mode. The visible light beam is passed through a Wollaston prism, which splits it, allowing half of the light to pass through a narrowband filter centered on a spectral line of interest (H α , [OI] or [SII]), and half through a filter centered on the neighboring continuum. Each spectral line probes a different type of phenomenon, described in detail in the text. In the case shown at right, H α emission is being used as a probe of accreting material. In five of the six filters shown, the primary (lower) star is brighter, however the additional accretion luminosity of the secondary (upper) star makes it the brighter one in the H α image (upper left panel).

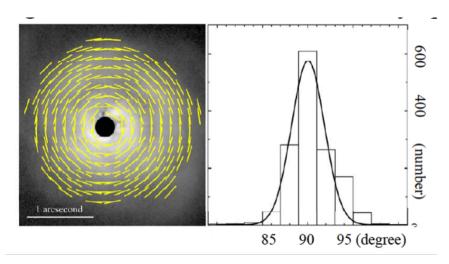


Figure 1.9: Sample Polarized Intensity image with centrosymmetric polarization vectors overlain (left) and histogram of orientations relative to the radial direction (right). Image credit: Jun Hashimoto

1.4.2 Polarized Differential Imaging

Polarized Differential Imaging takes advantage of the fact that direct starlight is (generally) unpolarized, while starlight is scattered most efficiently off of disk grains when it is polarized in a preferential direction (mutually orthogonal to the disk/star plane and the line of sight to the Earth). The polarized signal from a Young Stellar Object (YSO), therefore, consists of scattered light from dust grains and has very little direct starlight component (except in the case of instrumental polarization, which can be measured and removed), and mapping the polarization direction can further verify that this scattered light originates in a circumstellar disk, for which the pattern will be centrosymmetric, such as shown in Figure 1.9.

In the most basic sense, Polarized Differential Imaging (PDI) relies on splitting the telescope bean into two separate polarization states (l and r here), and modulating these by rotating a half-waveplate two four different angular positions generally 0° , 22.5° , 45° and 67.5° . This allows for extraction of the Stokes Q and

U parameters as follows.

$$Q' = I'(0^{\circ}, l) - I'(0^{\circ}, r)$$
(1.8)

$$-Q' = I'(45^{\circ}, l) - I'(45^{\circ}, r)$$
(1.9)

$$U' = I'(22.5^{\circ}, l) - I'(22.5^{\circ}, r)$$
(1.10)

$$-U' = I'(67.5^{\circ}, l) - I'(67.5^{\circ}, r)$$
(1.11)

$$Q = (Q' - (-Q'))/2 (1.12)$$

$$U = (U' - (-U'))/2 (1.13)$$

As the above equations reveal, the purpose of four rather than two half waveplate positions is that it provides two independent estimates of Q and U, based on images from separate parts of the detector. This technique, called the doubledifferential technique, reduces location-specific artifacts. These Q and U images can them be combined to create a PI image as follows.

$$PI = \sqrt{Q^2 + U^2} \tag{1.14}$$

Q and U can also be combined to create a polarization angle at each point in the image via:

$$\theta_P = \frac{1}{2} tan^{-1} \left(\frac{U}{Q} \right) \tag{1.15}$$

Since total intensity information is collected at the same time (by adding rather than subtracting the polarization channels), in principle a diskless PSF star can be observed and its PSF subtracted to obtain total intensity imagery of the disk. In practice, however, this proves quite difficult, as is discussed in Chapter 5.

1.4.3 Angular Differential Imaging

Angular Differential Imaging (ADI)is another frequently used differential imaging technique, and it can be combined with both of the other differential techniques described in this section to allow even more robust extraction of planetary signals. Quasi-static stellar speckles caused by imperfections in the telescope optics, unlike speckles due to atmospheric turbulence, evolve very slowly and can mimic a planet signal. Because these speckles are caused by the telescope optics, they are fixed in the detector frame. This need not be the case for planetary signals. If the telescope rotator is disabled during AO observations and the science object is allowed to rotate on the sky, then the planet signal will not be fixed in the detector frame. This provides a way to separate planet signals from quasi-static speckles. Since initial development of the ADI technique (Marois et al. 2006), several more complex algorithms have been developed to optimize this process. These include Locally Optimized Combination of Images (LOCI, Lafrenière et al. 2007) and Principal Component Analysis (PCA, Soummer et al. 2012). These techniques are optimized for point source extraction, and their utility for extracting light from disks is minimal in their current form, as discussed in Chapter 5.

CHAPTER 2

H-BAND SCATTERED LIGHT EMISSION FROM THE MYSTERIOUS SR21 TRANSITIONAL DISK

We present the first near infrared (NIR) spatially resolved images of the circumstellar transitional disk around SR21. These images were obtained with the Subaru HiCIAO camera, adaptive optics and the polarized differential imaging (PDI) technique. We resolve the disk in scattered light at H-band for stellocentric $0'.1 \le r \le 0'.6$ $(12 \le r \le 75 \text{AU})$. We compare our results with previously published spatially-resolved sub-mm Submillimeter Array (SMA) images that show an inner r≤36AU cavity in SR21. Radiative transfer models reveal that the large disk depletion factor invoked to explain SR21's sub-mm cavity cannot be "universal" for all grain sizes. Even significantly more moderate depletions (δ =0.1, 0.01) than those that reproduce the sub-mm cavity ($\delta \sim 10^{-6}$) are inconsistent with our H-band images when they are assumed to carry over to small grains, suggesting that surface grains scattering in the NIR either survive or are generated by whatever mechanism is clearing the disk midplane. In fact, the radial polarized intensity profile of our H-band observations is smooth and steeply inwardly-increasing (r^{-3}) , with no evidence of a break at the 36AU sub-mm cavity wall. We hypothesize that this profile is dominated by an optically thin disk envelope or atmosphere component. We also discuss the compatibility of our data with the previously postulated existence of a sub-stellar companion to SR21 at $r\sim10-20$ AU, and find that we can neither exclude nor verify this scenario. This study demonstrates the power of multiwavelength imaging of transitional disks to inform modeling efforts, including the debate over precisely what physical mechanism is responsible for clearing these disks of their large midplane grains.

2.1 Introduction

Transitional disks are of particular interest to the astronomical community because they have features intermediate between optically thick gas-rich protoplanetary disks and optically thin dusty debris disks, suggesting that they may be an intermediate evolutionary step. Transitional disks were first classified based on their Spectral Energy Distributions (SEDs) as objects with significant mid infrared excess emission indicative of starlight reprocessed by a circumstellar disk, but a lack of strong near infrared (NIR, \sim 1-5 μ m) excess. This NIR deficit implies a dearth of the very warmest grains in the inner disk (Strom et al. 1989; Skrutskie et al. 1990). Spatially-resolved sub-mm imaging of a number of transitional disks has revealed that many of them are indeed cleared of large thermally-emitting dust grains out to 10AU or more, well beyond the dust sublimation radius, thereby suggesting that an additional clearing mechanism may be at work (Andrews et al. 2009; Hughes et al. 2009; Brown et al. 2007; Andrews et al. 2011, referred to as A11 throughout the remainder of this paper).

The precise mechanism by which these clearings are made is widely debated in the literature, with several viable alternatives put forward, including clearing by massive forming protoplanets (e.g., Dodson-Robinson & Salyk 2011) and photoevaporation (e.g, Owen et al. 2011). Each theoretical explanation for the observed presence of sub-mm disk cavities makes its own predictions for clearing timescales, efficiencies and radii. However, the small amount of available spatially-resolved multiwavelength data on transitional disks has limited the ability of modelers to test these theories against observation. As such data accumulate, they reveal an increasingly complex picture of the cleared regions. In fact, it is clear that the nature of these cavities, and therefore an understanding of the mechanism or mechanisms that create them, can only be grasped through accu-

mulation of a statistically significant sample of multiwavelength data on transitional disks.

Thus far, the most success in spatially resolving transitional disks has been achieved in the sub-mm regime (e.g., A11), with a small number of additional studies using mid-infrared (MIR, $10\text{-}40\mu\text{m}$) aperture masking (e.g. Eisner et al. 2009, hereafter E09), NIR spectroastrometry (e.g. Pontoppidan et al. 2008, hereafter P08) and MIR interferometry (e.g. Ratzka et al. 2007) to infer inner disk structure. Efforts are currently underway to image a sample of transitional disks at shorter wavelengths, including the Strategic Exploration of Exoplanets and Disks with Subaru (SEEDS; Tamura 2009) campaign.

Despite initial success at resolving the rim of the LkCa 15 cavity in scattered light (Thalmann et al. 2010) with a Locally Optimized Combination of Images (LOCI) method, additional transitional disks imaged by SEEDS are revealing that these cavities, including that of LkCa 15, are more elusive when viewed in polarized light (Hashimoto et al. 2012; Muto et al. 2012; Dong et al. 2012; Mayama et al. 2012, Wisniewski et al. 2013, in prep, Grady et al. 2013, in prep). On the surface, this suggests that NIR scattering material is present, and even abundant, inside the cavities. NIR clearings thus may not be as ubiquitous nor as fully cleared in transitional disks as their sub-mm cousins appear to be.

Polarized NIR scattered light and sub-mm imaging, though both sensitive to circumstellar dust, probe very different regimes in the disk, both in terms of density and height above the midplane. The sub-mm emission from transitional disks is thermal, tracing mainly mm-sized grains that are concentrated ("settled") toward the disk midplane (e.g., Dullemond & Dominik 2004; Brauer et al. 2008). These large (a>10 μ m) grains are very important for understanding the large scale structure of the disk, as they house most of the disk's dust mass and their coag-

ulation drives the process of planet formation under the core accretion scenario (e.g., Pollack et al. 1996). The density and temperature of the grains at the disk midplane also determine the size and shape of the remainder of the disk.

At NIR wavelengths, however, the source of emission is less clear, an uncertainty that is compounded by the small number of spatially resolved disks in the literature. It is unknown whether disks with large grain clearings in the sub-mm are also cleared of small grains, nor is it understood what proportion of the disk mass lies in the small grain population. The fact that we see axisymmetric polarization structure in the NIR from transitional objects reveals that they house circumstellar disks with some small grain material. However, very little dust mass is needed to scatter at H-band, so it is unclear how much mass resides in this portion of the disk and whether it is optically thin or optically thick in this wavelength regime. If transitional disks are optically thick in the NIR, as is the case for protoplanetary disks, then this scattered emission will probe only the τ =1 surface. Small grains are believed to be vertically extended by a factor of two or more relative to larger grains (Dullemond & Dominik 2004; Brauer et al. 2008), so NIR scattering in an optically thick disk should occur well above the disk midplane. It is unclear whether or not clearing mechanisms acting on the large grains at the disk midplane would have an observable effect on the NIR emission under this scenario.

Conversely, if transitional disks are optically thin in this wavelength regime, as is the case for debris disks, then scattered light emission probes the entire vertical extent of the disk. In this case, scattered light emission originates from small grains throughout the vertical extent of the disk, down to and including the disk midplane. Observations of the same disk at both sub-mm and NIR wavelengths are compelling in that they allow for comparison of the settled and upper surface

layers of the disk, the large (mm-sized) and small (μ m-sized) grains, separately. A good disk model should thus be able to self-consistently reproduce observations at both wavelengths simultaneously, as well as any additional multiwavelength data. Our efforts to do just that for the SR21 disk are described in detail in Section 4. Section 2 describes our data collection methodology, and our result are summarized in Section 3.

2.1.1 Observational History of SR21

SR21A is a $2.5 M_{\odot}$, ~ 1 Myr old Class II Young Stellar Object (YSO) in the Ophiucus Star Forming Region. Recent astrometric observations suggest that the distance to the Ophiuchus star forming region is 119 ± 6 pc (Lombardi et al. 2008). However, we adopt a distance of 135pc (Mamajek 2008) in order to remain consistent with previous studies. SR21 has been classified as a 6'7 binary (Barsony et al. 2003); however, Prato et al. (2003) concluded that the two components were not coeval, and that the B component does not exhibit infrared excess emission and is therefore unlikely to host a circumstellar disk of its own. The A and B components also have discrepant proper motions (Roeser et al. 2010), so the "binary" does not appear to be a bound, coeval system. We will therefore refer to SR21A as simply SR21 throughout the remainder of this paper.

SR21 was first identified as a transitional disk by Brown et al. (2007), who inferred a 0.45-18AU gap in the disk based on a fit to the \sim 1-100 μ m SED. In order to explain the NIR excess seen at <5 μ m, they invoked a small uncleared region in the inner disk from 0.25-0.45AU. Since this initial identification, a variety of puzzling and seemingly contradictory data have emerged that continue to favor a disk that is not simply cleared of all material in this 0.45-18AU region.

The first evidence of a lack of complete clearing in the inner disk came from CO rovibrational spectroastrometry done with CRIRES on the VLT by P08, who

observed molecular gas emission originating from a region well inside of the dust gap inferred by Brown et al. They concluded that this emission arose from a narrow 7-7.5AU ring of molecular gas. It is not surprising to have found gas emission inside of the dust gap, as SR21 may still be accreting, albeit weakly $(<10^{-8.84} M_{\odot}/yr$, a relatively high upper limit, Natta et al. 2006). However, the physical mechanism causing an apparently truncated ring of molecular gas at this radius is still an open question. P08 suggest that the inner 7AU cutoff corresponds to a physical truncation of the disk, while a rapid drop in gas temperature outward of 7AU, perhaps due to a thick disk component in the region, causes the CO flux to quickly drop below the detection threshold. This explains the observed truncation of CO emission at 7.5AU without requiring that the disk material be physically truncated at this radius.

E09 also postulated that there is emission from within the \sim 18AU 'cavity' based on MIR (8.8 and 11.6 μ m) aperture masking data taken with TReCS on Gemini South. They put forward an alternate interpretation to Brown et al. (2007), capable of fitting both the SR21 SED and their visibility data. They suggest that the disk that is completely cleared within 10AU, but with a warm companion at 10-18AU whose circum(sub)stellar material creates the NIR excess seen in the SED.

The first spatially resolved imaging study of SR21 was done at 340GHz (880 μ m) and was published by Brown et al. (2009) among a sample of four disks with large central clearings. Followup data, also at 340GHz, were presented in Andrews et al. (2009). We base our sub-mm modeling efforts on A11, a subsequent detailed imaging and modeling study of 12 transitional disks, including SR21. Using two-dimensional Monte Carlo radiative transfer modeling to fit the sub-mm data and the infrared SED, A11 concluded that the 36AU sub-mm central

cavity in SR21 was depleted in material by a factor of $\sim 10^{-6}$ relative to the outer disk.

As multiwavelength data accumulate on SR21, it is clear that its structure, particularly in the region interior to the 36AU sub-mm clearing, cannot be understood through SED modeling alone because there are several degenerate solutions capable of explaining it. Spatially resolved images are needed at multiple wavelengths, probing multiple regions of the disk, before good theoretical constraints can be put on the mechanism clearing the interior of SR21. In the next section, we describe new spatially resolved NIR scattered light imaging, which contributes further to the puzzle of the SR21 "cavity".

2.2 Observations and Data Reduction

Polarized Differential Imaging (PDI) of SR21 was done in H-band ($1.6\mu m$) on 2011 May 22 with the high-contrast imager HiCIAO (Tamura et al. 2006; Hodapp et al. 2006) and adaptive optics system AO188 (Minowa et al. 2010) on the 8.2m Subaru telescope on Mauna Kea. These observations were conducted as part of the larger SEEDS survey, which was begun in October 2009 and is currently entering its third year.

The data were taken in "qPDI" mode in which a double Wollaston prism is used to split the beam into four 512x512 pixel channels (0′.0095/pixel), two each of o and e-polarizations. The splitting of each polarization state into two separate channels reduces saturation effects on the Hawaii 2RG detector. In order to obtain full polarization coverage and minimize artifacts, a half waveplate was rotated to four different angular positions (0°, 45°, 22°.5 and 67°.5). This cycle was repeated 18 times on SR21, with a 15 second exposure per waveplate position, for a total of 18 minutes of integration time.

Each image was bias subtracted, flat fielded and bad pixels removed in the standard manner for infrared data analysis using custom IDL scripts. Pinhole data were then used to create a distortion solution for each channel, which was applied before the images were rotated to a common on-sky geometry. All of the 15 second science exposures of SR21 were saturated to a radius of 6-8 pixels. For this reason, all subsequent combinations have had an inner r=8 pixel mask applied.

Each individual channel was then aligned using Fourier cross-correlation. These aligned images were combined to create Stokes Q and U images using standard differential polarimetry methods, namely adding together each set of two identically polarized channels and then subtracting these combinations from one another. The four channels were also added directly together to form a total intensity image.

To estimate the performance of the Subaru AO system during our observations, we used two unsaturated images of the Point Spread Function (PSF) reference star HD 148212, and 12 and 10 unsaturated images of SR21 taken before and after the main SR21 data set respectively. The average Strehl Ratio of the HD 148212 images is 0.27, and the average Strehl Ratios of the SR21 before and after images are 0.22 and 0.20 respectively. The standard deviations of the Strehl Ratios in the two SR21 data sets are 0.018 and 0.016 respectively, indicating that AO correction was maintained steadily.

2.2.1 Polarized Intensity Images and Polarization Vectors

The polarized intensity image shown in Figure 1 is a median combination of individual Polarized Intensity (PI) images for each of the 18 waveplate cycles, computed via $PI = \sqrt{Q^2 + U^2}$. Polarization angles were calculated for each image according to the formula $\theta_p = 0.5tan^{-1}(U/Q)$ and median combined. These are

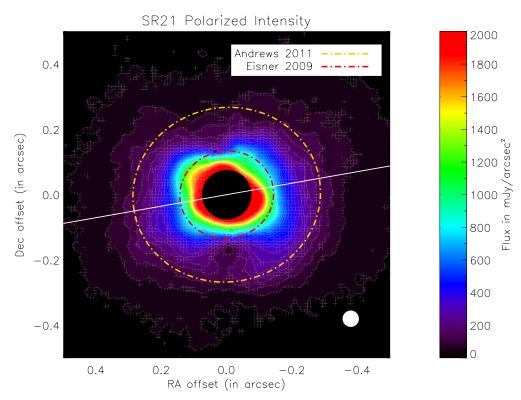


Figure 2.1: Median polarized intensity (PI) image of SR21 at H-band. The gap radius inferred by A11 and the companion orbit inferred by E09 are indicated with white and red dashed ellipses respectively.

shown overplotted on the PI image in five pixel bins in the lefthand panel of Figure 2. The length of each vector has been scaled according to the magnitude of the polarized intensity in that bin. A first order instrumental polarization correction has been applied, as described in detail in Hashimoto et al. (2011).

As has been demonstrated by other groups (Close et al. 1997; Oppenheimer et al. 2008; Apai et al. 2004; Quanz et al. 2011), PI images are a powerful data product for studying circumstellar disks. Unlike most other varieties of disk observations, PI data are generally spared the difficult and imperfect task of PSF subtraction under the assumption that direct starlight is randomly polarized, while light scattered off of a circumstellar disk shows a preferred polarization direction according to the scattering location in the disk. Specifically, the scattering process will preferentially scatter light polarized perpendicular to the direction to the illuminating source, while light coming directly from the star will be unpolarized and randomly oriented. Extended polarized emission is therefore unlikely to occur in systems without scattered light disks, and light scattered from a disk should show a centrosymmetric pattern of polarization vectors.

Although direct starlight is randomly polarized when it reaches the telescope, multiple reflective surfaces in the HiCIAO optical path are capable of polarizing some fraction of this incoming light. This phenomenon is referred to as instrumental polarization (IP). In order to estimate the quality of IP removal via standard techniques in our data, we reduced our data on the PSF star HD 148212 using the same methodology as SR21, including a first-order instrumental polarization correction. HD 148212 is a G2V star at d=105±15pc and ~10′ from SR21. It was observed just before SR21 on the same night and in the same observing mode, with 17 cycles of waveplate rotations and 5 seconds per exposure, for a total of 5.7 minutes of exposure time. As shown in the righthand panel of Figure

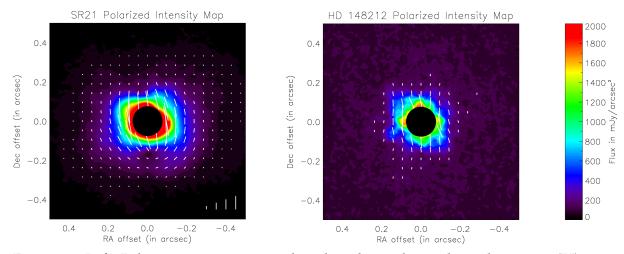


Figure 2.2: Left: Polarization vectors overplotted on the median polarized intensity (PI) map of SR21. A general centrosymmetric geometry is evident in the SR21 vectors, suggesting that this emission does indeed trace scattered light from a circumstellar disk. Right: Polarization vectors for the PSF star HD 148212 overplotted on its median PI image, which has been scaled by a factor of 3 to compensate for the difference in exposure time. In this case, the vectors do not show a coherent pattern, and PI emission is weak and compact. This is interpreted as photon noise residuals and is used to estimate the error in the polarized intensity profiles shown in figure 3. In both cases, vectors are scaled by the strength of the polarized intensity signal in the region and are not plotted when they fall at or below noise level. Vectors whose lengths correspond to 500, 1000, 1500 and 2000mJy/arcsec² are shown in the bottom right of the SR21 panel as a scale reference.

2, no polarized emission was observed beyond $r \gtrsim 0''\!2$ in the initial first order IP corrected HD 148212 polarized intensity image. Although the polarization signal inside of this radius is small compared to the Polarized Intensity of SR21 in the same region, it does show the characteristic four-lobed astigmatic pattern of the HiCIAO PSF. Additionally, while polarization vectors for HD 148212 do not show centrosymmetry, they do show some alignment along the vertical axis of the instrument. These are both hints that there is a residual uncorrected instrumental polarization signature, and further correction is warranted.

To first order, we can approximate the residual instrumental polarization by measuring the average polarization strength and direction of the signal in the HD 148212 polarized intensity image. We assume that HD 148212 is an ordinary diskless unpolarized star, and that all of its extended emission is due to residual instrumental polarization or photon noise. We construct an artificial "halo" by multiplying the average strength (1.1%) in the HD 148212 halo into the raw intensity image, then creating artificial "halo Q" and "halo U" images from it using the measured average polarization angle (-11°). We combine these synthetic Q and U images to create a "halo PI" image and subtract it from the original HD 148212 PI image. We also subtract the "halo Q" and "halo U" images from the raw Q and U images, and use the resulting images to calculate corrected polarization vectors, which we overplot on the halo-subracted PI image for HD 148212 in the righthand panel of Figure 2. The extended polarized intensity emission from the PSF star is reduced to r < 0'.1 through this process, and the resulting vectors are randomized, no longer showing a preferred polarization direction. We are therefore confident that this provides a reasonable approximation of the residual instrumental polarization.

The need to apply this same "halo removal" procedure to the SR21 PI data

can be seen in the Figure 2 SR21 polarization map. Although the SR21 vector pattern is clearly centrosymmetric along and around the disk major axis even in the uncorrected image, it deviates somewhat from centrosymmetry along the same instrumental axis as the aligned vectors in the HD 148212 image (approximately the minor axis of the disk). Since not all SR21 polarized emission is spurious, approximating the strength and direction of the halo is nontrivial in this case. We used a region with low intrinsic disk emission to construct a "halo" for subtraction, namely 100 pixel (0'95) wide regions centered on the minor axis at 30 pixels < r < 50 pixels (0'285 < r < 0'475) on either side of the disk. Following the same procedure as with HD 148212, we created polarized intensity "halo" Q, U and PI images from the polarization direction (-8.8°) and strength (1.02%) measured in this region and subtracted.

Subtraction of this first-order approximation reveals a more centrosymmetric pattern, as evidenced in the lefthand panel of Figure 3. However, it also introduces significant PSF artifacts to the polarized intensity map, as can be seen in the background polarized intensity contours of Figure 3. For this reason, we have chosen to model the "raw" unsubtracted radial profile in Section 4, noting, however, that some of the emission interior to 0.15 is likely to be contaminated by a small degree of instrumentally polarized stellar flux. The polarized halo strength of 1.02% is an order of magnitude smaller than our lower limit for the intrinsic SR21 polarization (10%), so this contamination is minimal.

Despite its crudeness, the "halo subtraction" method highlights an interesting feature of the SR21 polarization vector structure, which is that several regions in the disk with high signal (notably the northeast and southwest quadrants) show deviations from centrosymmetry even after halo subtraction. This may indicate that there is unresolved structure in the disk at these locations. Interestingly,

the deviation in the northeast is consistent with the location of the companion inferred by E09, which could be explained if the companion had an accretion disk of its own, with its own polarization structure superimposed on that of the SR21 disk.

2.2.2 Intensity Images

Intensity images were also constructed from the PDI observations of SR21 by cross correlating, centroiding and aligning the images in each of the four channels at each of the four waveplate (0°, 45°, 22°.5 and 67°.5) positions for each of the 18 observation cycles. The four channels at each waveplate position were summed for each cycle, and the resulting 18 images were then median combined to create an SR21 intensity image.

Unlike polarized intensity images, total intensity images are dominated by stellar flux, and PSF subtraction is necessary. For this purpose, we created intensity images of HD 148212 using the same methodology described above. The final reduced (but unsubtracted) median combinations of the SR21 and HD 148212 total intensity images are shown in Figure 4. Both reveal significant uncorrected astigmatism, as evidenced by the four-lobed pattern centered on the star. The SR21 disk is not visible in the unsubtracted total intensity image. The PSF subtraction process is described in detail in Section 3.3.

2.3 Results and Analysis

Our principle observational results are the final polarized intensity (Figure 1) and total intensity (Figure 4) images and polarization vectors (Figures 2 and 3). These images were analyzed in order to estimate the extent, inclination, major axis orientation and polarization percentage of the SR21 disk. The derivation of each of these properties is described in detail in this section.

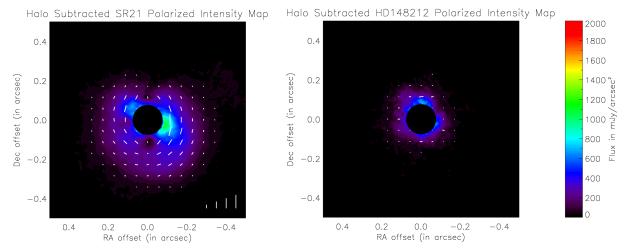


Figure 2.3: SR21 polarized intensity image (background) and polarization vectors (overplotted) after first order correction for a seeing limited halo component. This method reveals the polarization vectors to be more centrosymmetric, particularly along the disk minor axis) than originally inferred. Local deviations from centrosymmetry persist in the northeast and southwest quadrants of the disk, where unresolved disk structures and/or companions could induce local polarization structure that deviates from the larger centrosymmetric pattern. The process of subtracting this halo component has introduced additional PSF artifacts into the background PI map, most evident in the negative astigmatic pattern centered on the star. We have thus chosen to model the uncorrected polarized intensity map, which we believe more closely resembles the true structure of the SR21 disk. Vectors whose lengths correspond to 500, 1000, 1500 and 2000mJy/arcsec² are shown in the bottom right of the SR21 panel as a scale reference.

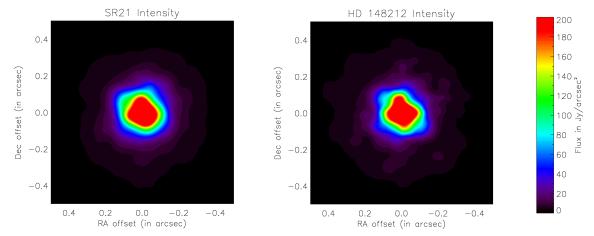


Figure 2.4: Intensity images of SR21(left) and HD 148212 (right). HD 148212 has been scaled according to differences in exposure time and intrinsic H-band magnitude (based on published 2MASS values) to match SR21 for subtraction. Some uncorrected astigmatism is evident in the four lobes of the PSF in both stars.

2.3.1 Polarized Intensity Isophotes

We began by fitting elliptical isophotes to our PI data in order to provide our own measure of the inclination and major axis of the SR21 disk, independent of the values in the literature, which were determined at other wavelengths. These elliptical fits are shown overplotted on Figure 5 in white, beginning at a radius of 10 pixels from the center (well outside the saturation radius) and extending outward to 50 pixels, where the PI flux approaches the noise level. Errors in isophotal fits were determined by summing in quadrature the intrinsic error estimation in the IRAF¹ ellipse procedure, which was used to fit the isophotes, and the standard deviation of the parameter across the image.

The average ellipticity of these isophotes suggests a disk inclination of $14\pm2^{\circ}$ (where 0° would be a face-on geometry). This is smaller than the value reported

¹IRAF is distributed by the National Optical Astronomy Observatory, which is operated by the Association of Universities for Research in Astronomy (AURA) under cooperative agreement with the National Science Foundation.

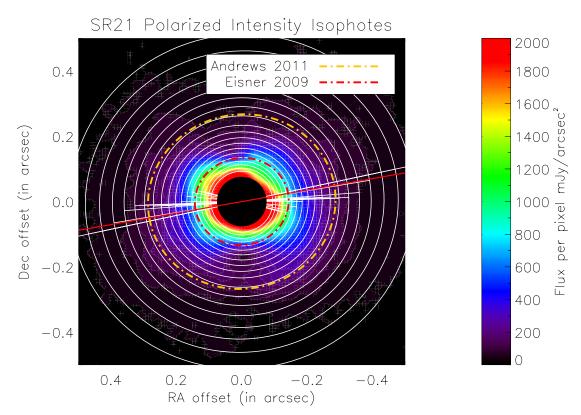


Figure 2.5: Elliptical isophotal fits to the SR21 Polarized Intensity image. The fits are shown overplotted on the PI image in white starting at a semimajor axis of 10 pixels (well outside the saturation radius) to a semimajor axis of 50 pixels (where the PI flux drops to noise level). The major axis of each isophote is overplotted as a straight white line. The average value of the major axis position angle is clocked by $\sim 15^{\circ}$ from the inner to outer disk, suggesting unresolved disk structure. The ellipticity of the isophotes was also used to estimate the inclination value of 14° (where 0° would be face-on).

by A11; however, their estimate of 22° was made with the caveat that the submm ring seen in SR21 is very narrow, and the geometry is not well constrained by their observations. The H-band emission is significantly more spatially extended; therefore, we adopt 14° as the preferred value in our modeling efforts, which are described in Section 4. There is no statistically significant deviation in our calculated inclination value between the inner and outer disk of SR21.

There is a clear counterclockwise progression from the inner to the outer disk in the major axis of the elliptical isophotal fits (overplotted in white in Figure 5). The average orientation of the major axis over the entire disk is $86\pm11^{\circ}$. The large uncertainty is due to the fact that the PA migrates from inner to outer disk. This happens relatively smoothly, but with a pronounced jump between 17 and 20 pixels. Inside of this radius, which is of particular interest because it corresponds to the location of E09's hypothesized companion, the average major axis position angle is 76±8° (measured East of North), while outside of this radius it is $96\pm7^{\circ}$. This shift in the average major axis PA of $\sim 20^{\circ}$ between these two regions, and of 42° in total between inner (68°) and outermost (110°) isophotes, suggests that the SR21 disk may contain unresolved structures such as a warp or a spiral arm. Some of this PA migration might be explained by instrumentally polarized emission with the shape of the HiCIAO PSF, particularly because the most pronounced turnover occurs at approximately the same radius as the extent of the HD 148212 halo. However, the observed SR21 isophote major axis migration continues well beyond that radius, and we therefore conclude that it is a real feature of the data.

We combine our results with previous estimates of the major axis PA from the literature in Table 1. The A11 estimate for the major axis PA of 100° (overplotted on Figure 5 in red) is consistent with our measurement of the PA of the outer disk.

We find that previous estimates of the PA of the disk are strongly dependent on the radius being probed, and our observation of a counterclockwise trend with radius fits well into the broader picture. This suggests that there is unresolved non-axisymmetric structure in the SR21 disk, and provides further confirmation that the trend we observe is not caused purely by residual PSF-shaped instrumental polarization signal.

Table 2.1: Position Angle Estimates

PA	Radius probed	Reference
0	AU	
15±4	$7.0{\pm}0.4$	P08 (v=1-0)
16±3	7.6 ± 0.4	P08 (v=2-1)
44 ± 7	11±1	E09
61±8	15±1	E09
76±8	12-25	this work
96±7	25-65	this work
100	30-120	A11
105	30-120	Brown 2009

2.3.2 Radial Brightness Profiles and Asymmetries

We used the final median combination of unsaturated total intensity images and the published 2MASS photometry values of SR21 and HD 148212 to establish a flux conversion for our data. In doing so, we make the assumption that the H-band magnitudes of both objects remained steady in the decade between 2MASS and HiCIAO observations. The ADU-to-Jansky conversion factor calculated using the HD 148212 images differs by only 1.6% from that calculated using the

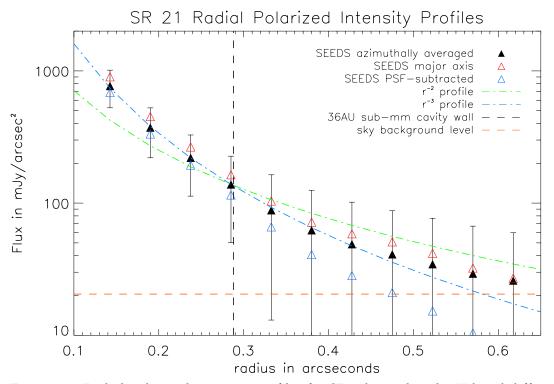


Figure 2.6: Radial polarized intensity profiles for SR21 binned to the H-band diffraction limit (0.5") and shown as triangles. The black filled triangles represent the azimuthallyaveraged radial profile and the red open triangles represent the profile as measured in an 11 pixel wide region centered on the major axis of the disk. The blue open triangles represent the profile of the PSF-subtracted image. All three profiles follow a very similar slope, with the major axis profile systematically higher in flux per pixel by 20% relative to the azimuthally-averaged profile, and the PSF subtracted image systematically lower, as is to be expected. Because the images were saturated interior to 0.07", radial profiles are only shown for bins that fall completely outside of the saturated region (from 0.1" outward). Also shown is a canonical r^{-2} profile in green and a r^{-3} profile in blue. The r^{-3} profile is a better fit to the data, although there is some evidence that it becomes shallower in the outer disk beyond 0.45". The veracity of this shallower profile in the outer disk is brought into question by the steepness of the PSF subtracted profile, suggesting that the flattening in the outer profile may be due to halo effects. Errors were estimated from the photon noise in the four individual channels used to create the PI image summed in quadrature (black error bars). The location of the sub-mm cavity wall at 36AU and the noise floor are indicated with vertical and horizontal dashed lines, respectively.

SR21 images; therefore, we believe that this is a robust assumption and have adopted the average of the two values as our flux conversion factor. We have applied a software mask to the saturated region interior to r=8 pixels; however, astigmatic PSF artifacts extend out to r~13 pixels (0″.13, ~18AU). Our data are therefore most robust at $18AU \le r \le 80AU$.

Computed flux-converted radial polarized intensity profiles for SR21 are shown in Figure 6. Brightness profiles are shown for both the azimuthally-averaged case (filled black points) and for the case where we limit analysis to a region 11 pixels wide centered on the disk major axis (open red points). We also over plot (open blue points) the radial profile of a PSF-subtracted Polarized Intensity image. Although PSF subtraction is nominally unnecessary for polarized intensity images, our analysis in section 2.1 revealed that residual instrumental polarization contributes at a low level. We therefore utilized the polarized intensity image of the PSF star HD 148212, and scaled it such that the noise in the background region was equivalent to that of SR21 and subtracted. This provided a subtraction with fewer evident PSF artifacts than our earlier "halo subtracted" PI image and serves to verify that the slope and steepness of the radial polarized intensity profile is insensitive to contamination by residual instrumental polarization.

In all three cases, we plotted the profile in five pixel radial bins excluding the innermost two bins (so interior to 0'.1), which have pixels inside the saturation radius. This binning was chosen to match the diffraction limit of 0'.05 at H-band. The brightness profile along the major axis is systematically higher than the azimuthally-averaged profile, as is to be expected based on the increased brightness of the disk in this region; however, the shapes are virtually identical. The PSF subtracted profile is systematically lower than that of the unsubtracted profile, as is to be expected. Unlike the other two profiles, the PSF-subtracted profile does

not flatten in the outer disk, suggesting that this trend may be caused by residual instrumental polarization in the unsubtracted profiles. We have elected to fit the azimuthally-averaged profile of the unsubtracted image in our modeling, which is described in Section 4.

These radial PI profiles reveal that H-band scattered light emission exists in the SR21 disk out to a distance of at least \sim 0%, which is comparable to the disk extent in the A11 SMA images. This corresponds to a physical scale of \sim 80AU, or more than twice the radius of the observed sub-mm cavity.

Errors in the radial polarized intensity measurements were estimated using both the background photon noise in the four individual qPDI channels of the SR21 raw images and using the observations of the PSF star HD 148212. The residual polarized intensity flux surrounding HD 148212 was summed in the same manner as the SR21 radial profile and is comparable to the photon noise. The error bars in Figure 6 and all of the radial profile plots shown in the modeling section correspond to the photon noise.

A canonical r^{-2} power law for the surface brightness of a flared disk is shown overplotted in green on all radial profiles and is a poor fit to the data. The best fit radial profile to the data is of the form r^{-3} . Although there is some indication that it may flatten beyond 0'45, this trend is not exhibited in the PSF subtracted PI profile, so we believe it to be spurious. The smooth, steep incline in polarized intensity from the outer disk in to the saturation radius is the crux of the new NIR data, and explaining it is the main driver of the modeling efforts described in Section 4.

2.3.3 Intensity Images and Compatibility with the Existence of a Substellar Companion

Direct addition of the 4 qPDI channels yields a total intensity image, which may also be diagnostic of the disk, but is dominated by stellar emission. PSF subtraction is required in order to isolate the disk component in these images. PSF subtraction is an imperfect process in systems such as SR21 for a number of reasons. First, Strehl Ratio may vary between the two sets of observations, and within each set itself, and Strehl Ratio mismatch can introduce artificial structure into the PSF-subtraction. The SR21 and HD 148212 data sets were calculated to have comparable Strehl Ratios (\sim 20%); therefore, we do not believe that Strehl mismatch is a dominant effect. Secondly, speckles that survive the median combination process are not necessarily cospatial in the two data sets; therefore, local excesses and deficits must be viewed with some suspicion. Thirdly, the SR21 data set was centrally saturated, while the HD 148212 data set was not, and so the counts in the SR21 image are an underestimate of the total emission. Finally, and perhaps most importantly, the total intensity of SR21 includes both a direct stellar and scattered light disk contribution, while tHD 148212's emission corresponds to direct starlight alone. These last two effects make estimation of the appropriate scaling of the PSF image for subtraction somewhat ad hoc, but by employing a series of reasonable approximations we can place bounds on the appropriate scale factor for subtraction.

We chose to scale the PSF star HD 148212 first by the relative ratio between its 2MASS H-band magnitude and that of SR21, a quantity which we will call *a*. Assuming that the total H-band magnitude of these systems has not varied significantly with time, subtracting the PSF star intensity image scaled by this value is an oversubtraction, both because some portion of SR21's H-band magnitude

corresponds to its disk contribution and because the SR21 image is saturated at the center. Therefore we use this subtraction (SR21 - a*HD 148212) as an "upper limit" on the disk's total intensity in scattered light.

We then created a grid of PSF subtractions with successively lower scale factors applied to the PSF star. We used the ratio of scattered polarized light to total scattered light in the disk as a metric in establishing the lower limit for the PSF scale factor. This ratio of polarized to total scattered light (P=PI/ I_{disk}) should be on the order of 50% in the case of 100% efficient scattering of unpolarized starlight off of spherical grains; however, variables such as porosity, oblateness, and alignment as well as particle size and composition effects, multiple scattering effects, wavelength dependence and inclination effects have been invoked to explain observed deviations from this value (Whitney & Wolff 2002; Simpson et al. 2009, e.g). Observed scattered polarization fractions vary from \sim 10% to \sim 70% in transitional disks, with typical values of approximately 40% (Dean Hines and Glenn Schneider, private communication based on HST program GO 10178, October 2012).

We deem values of <10% to be unphysically low based on these results; therefore, we apply the scale factor resulting in P=10% (0.5a) as our lower limit for PSF subtraction. Shown in Figure 7 is the PSF subtraction resulting from a scale factor 0.75a, which is halfway between the upper (a) and lower (0.5a) limits established above. The image has been smoothed by a gaussian with FWHM=6 pixels to mitigate the effects of speckles on the inferred structure of the disk.

In estimating the ratio of polarized to total disk intensity P (PI/I_{disk}), we use this intermediate PSF scaling and an annulus of 15 < r < 60 pixels. This region begins just outside the radius of the PSF astigmatism and ends at approximately the outer radius of the polarized intensity disk. Within this region, we have rejected

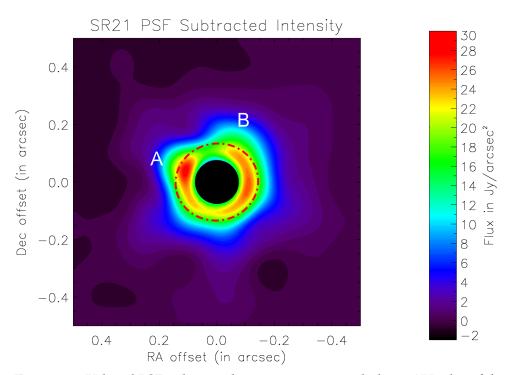


Figure 2.7: H-band PSF-subtracted intensity image with the 18AU orbit of the companion proposed by E09 overplotted. There is a distinct excess of emission in the Northeast quadrant (A) at a location that is consistent with their hypothesized companion, although its proximity to PSF artifacts makes robust identification impossible with this data. There is also a northward extension of the disk that is nominally present in the PI image (B). The image has been smoothed with a FWHM=6 pixel Gaussian in order to suppress artificial structures induced by speckle noise.

all pixels where (i) the polarized intensity flux is below noise level, (ii) P>1.0, indicating a PI speckle, or (iii) P<0, indicating a total intensity speckle. Our upper and lower limits for the subtraction tell us that SR21's P value lies somewhere between 10 and 27%, and a=0.75 gives P=18%.

The final PSF-subtracted and masked intensity images of SR21 for all scale factors investigated reveal two structures, both of which are also present in the PI data. The first is an excess in the NE quadrant near the location of the E09 companion, and the second is an arc extending upwards towards the north. These are marked as A and B in Figure 7. Although it is entirely possible that these features are spurious, they are intriguing in the context of the E09 observations and our own observation of the trend in PA with radius in the SR21 disk, which is suggestive of structure.

E09 inferred a warm (\sim 700 K), extended (\sim 40 R_{\odot}) companion 18 AU from SR21 based on the system's mid-IR flux and visibilities. Such a source would contribute \sim 5% of SR21's bolometric flux, or \sim 0.4 L_{\odot} , and would need to be a low-mass star surrounded by a dusty shroud or disk. The enshrouded companion would likely be considerably brighter in the H-band than a \sim 700 K blackbody, itself a magnitude 16 source at SR21's distance. We note that a bolometric luminosity of 0.4 L_{\odot} is characteristic of a \sim 0.2 M_{\odot} T-Tauri star at 1 Myr, which would have an effective temperature of \sim 3000 K (Kenyon & Hartmann 1995).

We obtained our images of SR21 with the image rotator off, allowing us to use Angular Differential Imaging (ADI) to search for a companion. A traditional ADI analysis (Marois et al. 2006) detected no statistically significant sources. Because of SR21's declination, we were only able to obtain $\sim 10^{\circ}$ of field rotation, making it difficult to apply more sensitive techniques like Locally Optimized Combination of Images (LOCI, Lafrenière et al. 2007). We therefore attempted a variation of

LOCI using the ACORNS-ADI pipeline (Brandt et al. 2012), in which we supplmented the comparison frames with 30 principal component images culled from a large library of observations of single stars.

We detect no statistically significant companion candidates down to a limiting contrast of 100 at 0'.14 (18 AU at the distance to SR21). This is not a meaningful limit for a substellar companion. However, it does likely rule out a stellar companion sufficiently luminous to account for the mid-IR excess reported by E09. Follow-up observations at longer wavelengths or with extreme adaptive optics will be necessary to reach fainter detection limits.

2.4 Modeling

To assist in interpreting our results, we ran a suite of disk models using the three dimensional radiative transfer Whitney code (see Whitney et al. 2003a,b, for details). The key capabilities, modifications from previously reported versions and input assumptions of the Whitney code are described in Section 4.1

In Section 4.2, we describe our efforts to reproduce the sub-mm observations of A11 using an independent code, and our verification of the A11 best fit parameters for the large grain dust disk.

In Section 4.3, we report on several degenerate explanations for the three features that make the SED of SR21 unique: (1) the steep $12-20\mu m$ rise, (2) the marked NIR excess and (3) the very small size of the $10\mu m$ silicate feature.

In Section 4.4 we report on our attempts to reproduce the observed steep r^{-3} polarized intensity radial profile, and its ability to constrain the geometry and optical depth of the disk.

Section 4.5 details our efforts to merge these three pieces into a single selfconsistent model, and in Section 5 (Discussion) we speculate on the implications of our best fit model.

2.4.1 The Whitney Code and Input Assumptions

The Whitney code mimics settling and grain growth by specifying the composition and structure of a "large grain" disk separately from that of a "small grain" disk. The large grain disk can be artificially settled relative to the small grain disk by setting its scale height parameters to lower values. Although both disks contain a distribution of grain sizes, we will refer to them simply as the "large grain disk" and "small grain disk" throughout.

The Whitney code treats both thermal emission and scattered light in the disk. Grain temperatures are computed using the method of Lucy 1999. and the code treats anisotropic scattering using a Henyey-Greenstein function (Henyey & Greenstein 1941). The albedo, average scattering angle and maximum polarization are pre-computed at each wavelength based on a Mie calculation for the ensemble of grains in each population. The polarization phase function is approximately dipolar, following White 1979.

The Whitney code utilized herein has been modified from previously reported versions and is described in detail in Dong et al. (2012) and Whitney et al. 2013 (in prep). The modifications of particular relevance to this paper were threefold. First, the disk density distribution was modified in order to allow for an exponential cutoff in mass as a function of radius, making it identical to the disk mass distribution of models run by A11. Secondly, the new models allow for a "puffing up" of the inner disk edge and the inner edge of the disk gap, a physical situation widely invoked to explain NIR spectral excess (e.g., Dullemond & Monnier 2010) and increasingly invoked to explain the observed features of gapped disks (e.g., Jang-Condell & Turner 2012). The final modification made to the code allows for the differential depletion of large and small grains, in addition to independent

geometries between the two disks (a feature included in earlier versions of the code). This means that the disk radial density exponent $(\alpha, \rho \propto r^{-\alpha})$, radial scale height exponent $(\beta, Z \propto r^{-\beta})$, absolute scale height at 100AU and depletion (δ) are all specified independently inside and outside of a disk gap or cavity. In effect, this allows for four independent disk components - large and small grain disks inside and outside of a specified gap radius. The size distribution and composition (and therefore opacity as a function of wavelength) of large and small grains are also specified independently.

The basic model input parameters (e.g., stellar mass, radius, temperature) were determined based on values in the literature, and those adopted by A11 in particular. The choices of model parameters are given in Table 2. In total, over 250 independent models of the SR21 disk were calculated in this modeling effort.

Initial input assumptions included a disk mass of $0.006 M_{\odot}$, a gas:dust ratio of 100:1, and a large grain dust mass fraction of 85%. Neither of these values is well constrained by our H-band observations, so we have assumed the same values as A11. We have fixed the composition of the settled dust disk to the grain prescription of (Wood et al. 2002), a power-law grain size distribution with an exponential cutoff (Kim et al. 1994) composed of amorphous carbon and astronomical silicates with solar abundance constraints for carbon and silicon (C/H \sim 3.5 x10⁴ and Si/H \sim 3.6 x10⁵, respectively; Anders & Grevesse 1989; Grevesse & Noels 1993). The functional form of this size distribution is

$$n(a)da = C_i a^{-p} exp\left[-\left(\frac{a}{a_c}\right)^q\right] da$$

, where a is the grain size, n is the number of grains, and the parameters p, a_c , and q, which control the distribution shape, are adjusted to fit the wavelength dependence of the observations. In particular, the long wavelength slope of the SED and the peak of the MIR excess are sensitive to these parameters. Following



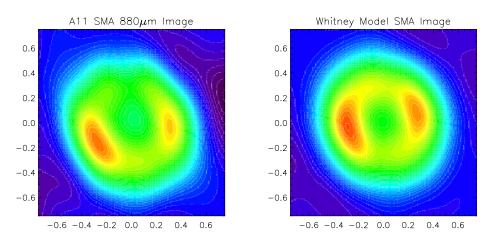


Figure 2.8: Left: The 880μ m SMA image of A11. Right: A large grain disk model with A11's best-fit input assumptions produced by the Whitney code.

the Wood et al. 2002 best fit model for the HH30 disk, we use p=3.0, q=0.6 and a_c =50 μ m. This is a good match to the grain properties inferred by A11.

Several small grain dust prescriptions were also investigated, and the observed SED and H-band images were found to be relatively insensitive to their particulars. We use the interstellar medium dust prescription of Kim et al. 1994 which has a smaller average grain size than our other models. Like the large grain disk, the small grains follow a power law size distribution with an exponential cutoff, however in this case they are composed of silicates and graphite, with maximum grain sizes of $0.16\mu m$ and $0.28\mu m$ respectively.

2.4.2 Sub-mm Image Modeling

Our H-band PDI observations probe only scattered light from the upper surface of the disk and therefore do not constrain the properties of the large grain disk where the majority of the disk mass lies. However, we have endeavored to match

Table 2.2: Model Input Parameters

Parameter	Model Input(s)	
	Input Assumptions	
M_{star}	2M _☉	
R_{star}	$3.15 \widetilde{\mathrm{R}_{\odot}^{*}}$	
T_{star}	5750K ⁺	
A_V	6.3	
M_{disk}	$0.006\mathrm{M}_{\odot}^*$	
${ m r}_{gap}$	none, 18, 36* AU	
${ m r}_{disk,in}$	$ m r_{\it sub}$, 7AU	
$\mathbf{r}_{disk,out}$	200AU	
i	14, 22*	
r_C	15AU*	
Large Grain Disk Parameters		
composition	Wood et al. 2002 ⁺	
f_d	0.85^{*}	
Z_{100}	1.52AU*	
α_{outer}	2.15^{+}	
β_{outer}	1.15+	
α_{gap}	2.15 ⁺	
$eta_{m{gap}}$	1.15^{+}	
δ	10^{-8} , 10^{-7} , 10^{-6^+} , 10^{-5} , 10^{-4} , 10^{-3} , 10^{-2} , 10^{-1} , 1	
Small Grain Disk Parameters		
composition	Kim et al. 1994; Cotera et al. 2001	
f_d	0.15	
Z_{100}	2-15AU	
α_{outer}	0.5-4	
eta_{outer}	1-3.15	
α_{gap}	-5-10	
eta_{gap}^{gap}	1-3.15	
δ^{gap}	$10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}, 10^{-1}, 1, 2, 5, 10$	
of Whitney mode	el input parameters investigated for SR21. Where several	

The range of Whitney model input parameters investigated for SR21. Where several discreet values were attempted, the table entry is given in the form of a list, and for parameters where a range of values was attempted, the table entry gives the full range of attempted values. Parameters identical to those of A11 are marked with an asterisk, while equivalent parameters (rounded for convenience, compatibility, or translated to Whitney code syntax) are marked with a + sign. R_C is the characteristic radius for the exponential density cutoff as defined in A11, f_d is the fraction of the disk's dust mass, Z_{100} is the disk scale height at 100 AU, α is the exponent for the radial midplane density dependence, β is the exponent for the radial scale height dependence, and δ is the depletion factor of the gapped region relative to the outer disk

the sub-mm observations of A11 in addition to our newly-acquired data in order to develop the most physically motivated disk model possible. As a first test of the compatibility of the A11 models with the Whitney code, we began with input parameters for the disk that were identical (or nearly) to the best-fit model of A11.

Figure 8 shows the SMA image taken by A11 in panel 1, and the equivalent $880\mu\text{m}$ image produced by our model in panel 2. We are able to produce a similarly well-fitting model with the Whitney code and identical input parameters. Model outputs were sampled with the same spatial frequencies observed with the SMA, and images were created with the same inversion and deconvolution parameters as the observations.

Having shown that the A11 and Whitney models produce equivalent results with identical input assumptions, we then investigated the effect of gap depletion on the sub-mm images to test the robustness of the A11 assumption that the gap is depleted by a factor of δ =10^{-5.9} relative to the outer disk. Keeping all other disk parameters fixed, we varied only the depletion factor δ in the interior of the gap by an order of magnitude for each model from δ =10⁻⁸ (two orders of magnitude higher depletion than the A11 assumption of \sim 10⁻⁶) to zero depletion (which, at least on the surface, is the most consistent with the evidently undepleted cavity seen in our observations).

We break these models into three categories: "high" depletion models (δ =10⁻⁸-10⁻⁶), "moderate" depletion models (δ =10⁻⁵-10⁻³) and "low" depletion models (δ =10⁻²-10⁰) and show the real 880 μ m image of A11, our simulated image, residuals and visibility fits for each of these models in Figures 9-11. In each case, the model images were scaled such that the total flux was equivalent to that of the real SMA observation, and the required scaling is given in the model panel of the figures.

The "low" depletion models provide poor fits to both the 880 μ m image and the visibility profile, so we conclude, as did A11, that the SR21 disk is depleted in large grains interior to 36AU. All of the moderate and high depletion models provide reasonable fits to both the 880 μ m image and the visibility profile; however, the δ =10⁻⁵ model shows the lowest residuals and the δ =10⁻⁶ model provides the best fit to the visibility profile. The δ =10⁻⁶ model requires just a 2% increase in flux to sum to the observed flux density (easily explained by small differences in input opacity assumptions), while the δ =10⁻⁵ model requires an 8% scaling, so we adopt the δ =10⁻⁶ value for the large grain r<36AU disk in all further modeling. This value is closest to the A11 best fit value of δ =10^{-5.9}; therefore, we consider our large grain models to be functionally equivalent.

As a test of the robustness of the large grain disk assumptions, we verified that the small grain disk parameters specified in our models did not create observable variation in the simulated $880\mu m$ image. As expected, we found that wide variation in input small grain parameters has no observable effect on the $880\mu m$ images produced by our models. The $880\mu m$ emission of the large grain disk is therefore strongly decoupled from the small grain disk parameters.

Having verified that we can reproduce the A11 result, that the A11 inferred gap radius and depletion factor for large grains is a good match to the models produced by both codes, and that this result is unaffected by variation in small grain parameters, we treat the A11 large grain disk parameters (with the exception of disk inclination, for which we adopt our measurement) given in Table 2 as fixed throughout the remainder of our modeling efforts.

2.4.3 SED Modeling

There are sufficient data in the literature to construct a well-sampled observed SED for SR21. In addition to 2MASS J, H and K photometry, SR21 photometry

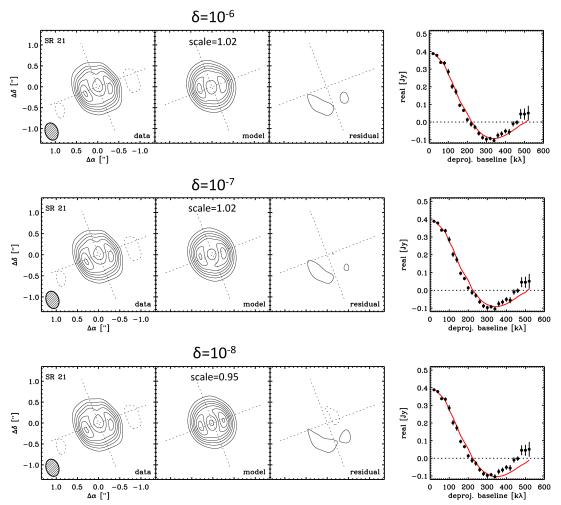


Figure 2.9: Highly depleted (δ =10⁻⁸-10⁻⁶) large grain disk models produced by the Whitney Code. From left to right, the panels are (a) the observed SR21 SMA image of A11 (b) the final Whitney model image, where the raw model output was sampled with the same spatial frequencies observed with the SMA (c) the residuals and (d) the model fit (in red) to the SMA visibilities. Input parameters were identical to those of A11, except that the depletion factor (δ) for large grains inside of 36AU was varied. The depletion factor is given above each model row. A scale factor was also applied to each model image to equalize the model and SMA fluxes before subtraction from the observed image. This factor is given in panel (b) of each model. The minor discrepancies between modeled and observed fluxes can be easily explained by slight differences in input opacities.

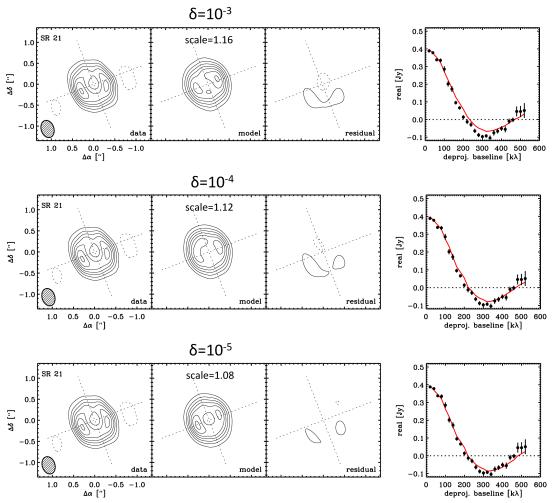


Figure 2.10: Moderately depleted (δ =10⁻⁵-10⁻³) large grain disk models produced by the Whitney Code. See Fig. 9 caption for details.

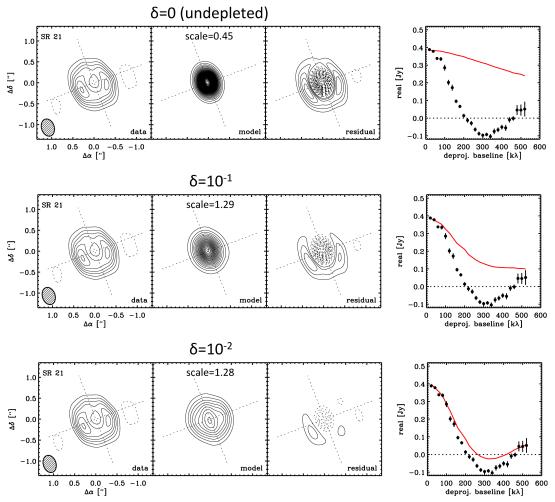


Figure 2.11: Minimally depleted (δ =10⁻²-10⁰) large grain disk models produced by the Whitney Code. See Fig. 9 caption for details.

has been obtained with Spitzer's Infrared Array Camera (IRAC) and Multiband Imaging Photometer (MIPS) at 3.6, 4.5, 5.8, 8 and 70μ m, the AKARI spacecraft at 18μ m, the Infrared Astronomical Satellite (IRAS) at 12, 25 and 60μ m, the Submillimeter Array (SMA) at 880 μ m, and the James Clerk Maxwell Telescope's Submillimetre Common-User Bolometer Array (SCUBA) at 350, 450, 850 and $1300\mu m$. $10-35\mu m$ spectra from the Spitzer Infrared Spectrograph (IRS) and $52-97\mu m$ from the MIPS SED mode are also available. These are shown overplotted on all SEDs generated by our models. Although we show it for completeness, the MIPS 70μ m photometry point has been excluded from our fits due to its inconsistency with both the MIPS SED spectrum and IRAS $60\mu m$ photometry. All reported photometric points and spectra have been extincted according to a standard r_V =3.1 interstellar extinction law (Cardelli et al. 1989). The extinction for SR21 has been estimated to be as high as A_V =9.0 (Prato et al. 2003), but A11 have argued that the published photometry is most consistent with a lower value of $A_V \sim 6.3$. We also find that lower extinctions are most consistent with the observed optical/NIR SED, so we adopt the same A_V =6.3 value for SR21.

The SR21 SED shows circumstellar dust excess emission longward of the stellar blackbody peak, as evidenced by the long wavelength excess (see Figure 12). Unlike protoplanetary systems, which tend to have flatter transitions between the stellar and dust peaks in the SED indicating significant dust populations at all temperatures, the SR21 SED has a distinct double peaked profile, reaching a local minimum near 12μ m. This suggests that there is a dearth of hot grains in the disk, a conclusion supported by the observed sub-mm cavity. Debris disk systems, on the other hand, are much more depleted of small grains, and their emission traces the stellar photosphere out to wavelengths of 20μ m or more. In contrast, the SR21 SED rises rapidly beyond 12μ m, indicating a healthy popula-

tion of warm grains that emit at $\lambda > 12 \mu m$.

Shortward of the 12μ m local minimum in the SR21 SED, the emission is also greater than photospheric, which indicates a small population of very warm grains (a hot inner disk component) or, alternatively, a warm companion. The SR21 SED shows little evidence of a 10μ m silicate feature, while many similar transitional disk systems exhibit strong, broad silicate emission at this wavelength. A good disk model for SR21 therefore needs to reproduce all three of these distinct SED features: (a) the steep $12\text{-}20\mu$ m rise, (b) the sub- 12μ m excess and (c) the lack of a strong, broad 10μ m silicate feature.

The Whitney code produces stable SEDs with relatively few (\sim 10⁶) photons, making it much less computationally expensive than generating images/radial profiles. We were therefore able to explore much of the available parameter space for SR21 to find several classes of reasonable models capable of reproducing the observed features of the SED.

2.4.3.1 NIR Excess Fits

The NIR excess shortward of $10\mu m$ is a distinct feature in the observed SED. The (small) excess in this region was originally attributed to a 0.07-0.25AU warm inner disk component. Indeed, adding this component to our models produces a well-fitting NIR excess. However, the MIR observations of Pontoppidan et al. (2008) are somewhat inconsistent with the presence of this interior hot disk component. Their data show a sharp truncation in gaseous CO at 7AU, suggesting a high flux of UV photons. This feature would be difficult to produce with an inner hot disk component at r<7AU capable of intercepting them. Furthermore, extending the large grains in to 0.07AU in our models, even while depleting them by a factor of 10^{-6} , results in point-like sub-mm emission at the disk center, which is not present in the A11 observations.

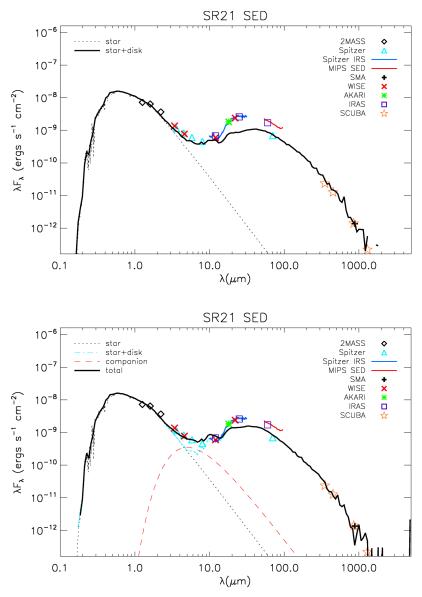


Figure 2.12: Two model SEDs produced by otherwise identical models with (top) a small grain disk that extends inward to 0.07AU and (bottom) a more distant 7AU truncation in the small grain disk and a warm companion with E09 parameters. Both models provide good fits to the λ <12 μ m NIR excess

The presence of the warm companion predicted by E09 is another potential mechanism for creating excess NIR flux in SR21, and in this case allows for a larger inner disk truncation radius. Our observations can neither confirm nor exclude the presence of a companion, yet they do provide indirect support of its existence, as detailed in Section 3.3.

Otherwise identical models with an (a) 7AU truncation and a warm (T=730K, R=30r_☉, E09) blackbody companion (b) a disk that extends inward to the sublimation radius at 0.07AU are shown in Figure 12 to demonstrate that both assumptions reproduce the observed NIR excess. Without any means of probing disk structures interior to the saturation radius of 0′.1 in our observations, we cannot break the degeneracy between inner disk extent and the existence of a companion in the NIR SED fits using scattered light imagery.

We hold a slight preference for the warm companion hypothesis based on the P08 CO results and our own observations of excess intensity and deviation from polarization axisymmetry (indirectly supporting the companion hypothesis and described in detail in Section 3.3), so we assume a 7AU truncation in all of our scattered light models. We add the warm companion component to all of our modeled SEDs in order to reproduce the observed NIR excess, but acknowledge that a significant inner disk extent is also a plausible mechanism for creating this feature. We do not add polarized intensity emission from the companion into our simulated images because the polarized intensity contribution of such a companion is (a) not well constrained by models and (b) should not have a significant effect on the azimuthally-averaged radial profile.

2.4.3.2 10μ m Silicate Fits

The final feature of note in the SR21 SED is the lack of a prominent 10μ m silicate feature, which tends to be rather strong and broad in disks with a significant in-

ner small grain disk component. The IRS spectrum spans this region of the SED and also that of the weaker 18μ m silicate feature without any evidence of a strong line flux at either wavelength. We find that the majority of the 10μ m silicate emission comes from the inner disk wall, and we can suppress it by either restricting the scale height of the small grain inner disk wall or by restricting small grains to a very small proportion of the total mass (<1%). Generally speaking, many otherwise well-fitting models overproduce 10μ m silicate emission. Changing the small grain prescription may be warranted in future modeling attempts. Specifically, the silicate features could be suppressed by either increasing the grain sizes or limiting the proportion of silicate grains in the disk. We leave this for future modeling studies.

Our scattered light imagery cannot give any information about the size of the small grain disk wall where the 10μ m feature originates, whether it lies at 7AU or 0.07AU (the sublimation radius). We limit the size of the inner disk wall to small scale heights in all of our scattered light models in order to suppress the 10μ m feature, but in some cases not severely enough, resulting in overproduction of 10μ m emission in the SEDs. We could presumably correct this by running additional models with smaller inner wall scale heights or reduced small grain populations, but have elected not to based on the computational expense of full scattered light models as well as our inability to resolve the region of the inner wall with our observations.

2.4.3.3 Fits to $10-20\mu m$ rise

The steepness of the rise from $12\text{-}25\mu\text{m}$ is perhaps the most unique and significant feature of the SR21 SED and is the primary feature that led to its "transitional disk" classification. The presence of this feature requires that there be a sharp increase in disk emitting area (e.g. a "wall") at intermediate (20-40AU) radii.

However, the precise geometry of the "wall", as well as its composition, is not well constrained by the observed SED.

The simplest way to create a "wall" is to allow for depletion inside of the r=36AU gap, so that the disk makes a transition from optically thin to optically thick at this radius. We know that the disk is optically thin at r<36AU for long wavelengths (and therefore in large, thermally-emitting grains), and optically thick beyond; therefore, the large grains certainly do present a "wall" at 36AU. All of the models that we consider are therefore "gapped" in large grains.

However, the sub-mm data places no such constraint on small grain depletion. In fact, there is reason to believe that small grain dust filtration occurs at gap radii (Zhu et al. 2012), and that material must survive into observed sub-mm dust gaps, as a number of gapped disks are still observed to be actively accreting (Salyk et al. 2009). The observed H-band radial profile of the SR21 disk would seem to support this hypothesis, as it shows no evidence of a change in small grain disk properties at r=36AU. Until our Subaru H-band data were obtained, there were no spatially resolved data at shorter wavelengths to constrain the properties of the small grains and the amount to which they are depleted inside of the sub-mm cavity.

The simplest hypothesis, put forward by previous SED modelers, is that the small grain disk follows the same structure as the large grain disk, including the r<36AU gap. This hypothesis reproduces the observed SED, but it is inconsistent with our new Subaru H-band polarized intensity profile. Simulated radial profiles and SEDs for various small grain depletion factors are shown in Figure 13. They show that although small grain depletion in the r<36AU region reproduces the $12-20\mu m$ rise in the SED, it is strongly inconsistent with the observed H-band radial profile. All models in which the inner disk is strongly depleted, and

therefore optically thin, suggest that we should have seen an excess in H-band emission at the cavity wall and a turnover in the radial profile at this radius.

Undepleted and low depletion models, on the other hand, are inconsistent with the observed SED. Additionally, their radial profiles are too shallow to be consistent with the H-band data. Appendix A details our efforts to find a scenario in which the small grain disk, whether optically thin or optically thick at r<36AU, contributes to the $12-20\mu m$ rise in the SED.

In the end we find that the steep $12-20\mu m$ rise can be created with a disk depleted only in large grains in its interior. Our data suggest that the small grain disk is not "gapped".

2.4.4 Scattered Light Imagery

At first glance, the H-band radial polarized intensity profile is difficult to reconcile with the large, heavily depleted sub-mm cavity, as it spans the region of the sub-mm wall without any evident break or turnover. In fact, the principal features of the observed radial polarized intensity profile, namely its smoothness and its steepness ($\sim r^{-3}$), were not reproducible under any of the scenarios we investigated in which the small grain disk contained a depleted cavity. Furthermore, we were unable to reproduce the radial profile under any scenario in which the H-band scattered light originates from the surface of an optically thick small grain component, whether gapped or not. This effort is described in detail in Appendix A. This is a surprising result, as we had no reason to expect that transitional disks are optically thin at short wavelengths.

We find that the steepness of the radial profile requires that we invoke an optically thin small grain component that spans the entire disk. In order to create this, we must either (a) relax the A11 assumption that the small grains compose 15% of the disk's dust mass and restrict their contribution to <1%, making the

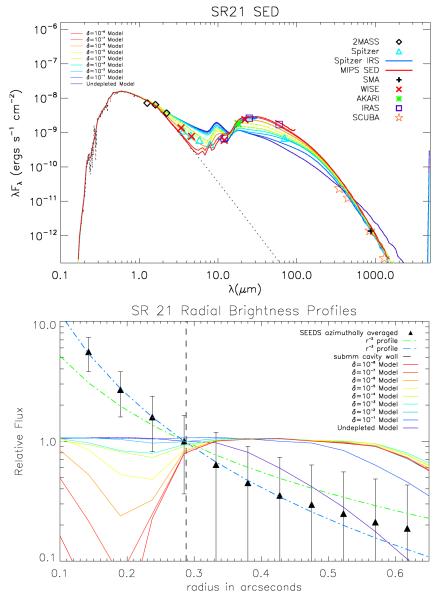


Figure 2.13: Top: model SEDs of varying r<36AU depletion factors overplotted on the observed SR21 SED. The nine "Gapped" model runs of different depletion factors are shown with solid colored lines, ranging from 10^{-8} depletion in red to zero depletion in purple. Bottom: H-band radial polarized intensity profiles for the same models overplotted on the observed profile. The flux of the undepleted model was fixed to the data at r=36AU, and all other models scaled by the same amount, in order to highlight changes in the radial profile at the cavity wall. None of the models provide a good fit to the data, revealing that the small grain disk is not only undepleted, its geometry is different from that of the large grain disk.

small grain dust disk optically thin throughout, or (b) invoke a third optically thin disk atmosphere or envelope component that has a larger vertical extent than either the large or small grain disks.

Under scenario (a) the optically thick "wall" producing the $12\text{-}20\mu\text{m}$ SED rise must be composed entirely of large grains in order to allow the small grain disk to be optically thin. As we noted in section 4.3, we do not require a small grain wall in order to reproduce the $12\text{-}20\mu\text{m}$ rise in the SED. In fact, we can reproduce the SED structure at $\lambda > 12\mu\text{m}$ with a large grain gap alone. We do need to increase the scale height of the gap wall in order to provide enough flux at these wavelengths (relative to models where an optically thick small grain component contributes to the rise), but not by such a degree as to lead us to declare these models unphysical $(Z_{wall} \gtrsim 1\text{AU})$.

Since an optically thin small grain disk means that we inevitably "see" the large grains at the midplane, these models require that the large grain disk not scatter efficiently at H-band. Otherwise, they suffer from the same local maximum problem that led us to exclude "Gapped" small grain disks. As can be seen in Figure 14, which shows a scenario under which the small grain disk is optically thin throughout, the Wood et al. (2002) dust prescription chosen for the large grain disk (and found to be a good match to that of A11) does scatter weakly at H-band, resulting again in a turnover in the radial profile at 36AU. These models would require an alternate grain size distribution and/or composition in order to prevent a local excess at 36AU. We have not endeavored to explore this possiblity here.

We disfavor explanation (a) because it requires that small grains be optically thin throughout the vertical extent of the disk. An entirely optically think small grain disk disk is very unlikely given the small amount of disk mass allowed

Large grain density distribution Spectral Energy Distribution 50.0 10 16.7 10 10-10 -16.7 10⁻¹² 10.0 λ(μm) 0.1 1.0 100.0 1000.0 33.3 66.7 100.0 0.0 Small grain density distribution Radial Polarized Intensity Profile 50.0 10.0 SEEDS H—band Model r² profile r³ profile J sub—mm cavity wall sky background level (24.7 log(density) 16.7 Scaled Flux -16.7

Small Grain Optically Thin Disk Model

Figure 2.14: Sample model outputs for a representative optically thin disk model. The lefthand panels represent the density of large grains (upper) and small grains (lower). The righthand panels provide model fits to the observed SED (upper) and H-band radial polarized intensity profile (lower)

0.1

0.1

0.2

66.7

100.0

and the observed optically thick r>36AU large grain component. In order to be optically thin throughout, we require not only that <0.1% of the disk mass lie in the small grain disk, representing a likely unphysical degree of grain growth or small grain depletion, but also that the large grain disk's grain size distribution contain no grains that scatter at H-band, an equally unlikely scenario.

Under scenario (b), which we believe to be far more likely, an envelope component with a significant vertical extent lies above the optically thick small grain dust disk. Representative model outputs for this scenario are shown in Figure 15, and a schematic of the disk structure is shown in Figure 16.

Although we call this third disk component an "envelope", it is not clear that it fits into the canonical definition. There has been no evidence in the literature to suggest that SR21 is still actively accreting, and the only constraints on the accretion rate are upper limits (Natta et al. 2006). This is in contrast to a number of other gapped disks that do show evidence of accretion (Salyk et al. 2009).

It seems that this optically thin component may be more readily explained as a vertically extended disk "atmosphere" of sorts, rather than an accretion envelope. Whether this is part of the natural structure of the disk, a remnant of the envelope, or a disk wind launched from the optically thick disk surface cannot be determined from our data. In fact, our observations are relatively insensitive to its precise geometry. The radial profile of the optically thin component follows an r^{-3} power law in all cases where the disk α - β =-1, which allows for a large range in flaring and density structures.

Under this three component disk model, the radial profile is still sensitive to the structure of the underlying optically thick small grain disk surface. In particular, the addition of the optically thin component that gives rise to the $\rm r^{-3}$ power law cannot mask a discontinuous small grain distribution at the midplane, even

Large grain density distribution Spectral Energy Distribution 100 (ergs s⁻¹ cm⁻²) 10-33 10 10-10 -33 10-11 10-12 10.0 λ(μm) -100 0.1 1.0 100.0 1000.0 133 0 67 200 Small grain density distribution Radial Polarized Intensity Profile 33 Scaled Flux -33

133

200

Three Component Disk Model

Figure 2.15: Sample model outputs for a representative three component disk model. In this case, the underlying optically thick small grain disk is of the 'Steeply Curved' variety, and the scattered light contribution from the optically thin component is thirty times that of the optically thick component The lefthand panels represent the density of large grains (upper) and small grains (lower). The righthand panels provide model fits to the observed SED (upper) and H-band radial polarized intensity profile (lower)

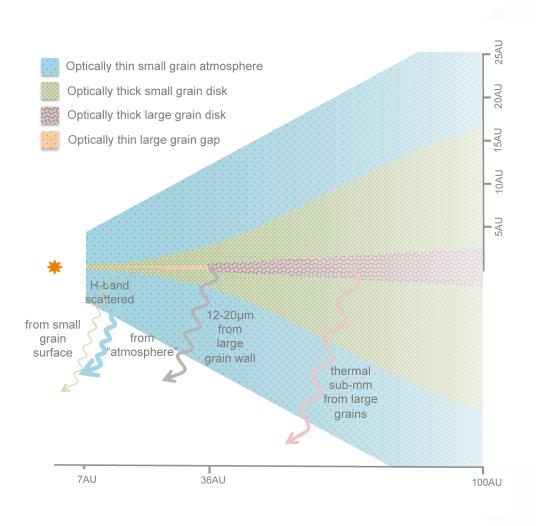


Figure 2.16: A schematic view of the three component disk model that provides the best fit to our data. The vertical scale has been inflated by a factor of two relative to the horizontal. The origins of sub-mm light, H-band scattered light and the steep $12-20\mu m$ SED rise are indicated.

when the optically thin component is allowed to contribute an order of magnitude more scattered light than the optically thick component.

2.5 Modeling Discussion

Our requirements to fit the radial profile are threefold. First, we require that there be enough small grain material inside the 36AU large grain cavity to scatter direct NIR starlight before it reaches the cavity wall, effectively shadowing it. Although this gap must be optically thin to sub-mm light in order to reproduce the SMA observations of A11, we find no viable scenarios in which it can be optically thin at shorter wavelengths and remain consistent with the observed H-band radial polarized intensity profile.

Secondly, the smoothness of the observed H-band polarized intensity profile suggests that the surface of this underlying optically thick small grain disk must follow a smooth surface geometry. We cannot fit our data under scenarios where there is a marked effect, either in differential depletion or a scale height discontinuity, of the large grain dust gap on the small grain distribution. This is perhaps the most interesting and puzzling result of our observations and analysis – that the small H-band scattering grains in the SR21 disk must follow a smooth distribution that is relatively unaffected by the goings-on at the disk midplane. This supports scenarios under which dust filtration occurs at the large grain gap wall, allowing the small grains to be carried into the gap along with the disk's gas.

Finally, we find that the steepness of the H-band radial profile of SR21 is indicative of an optically thin small grain disk component. In order to remain consistent with the first two criteria, this requires that we invoke a third vertically extended optically thin disk component overlying the optically thick portion of the disk.

As noted in Section 2, the AO system performance was sub-optimal during the SR21 observations, and uncorrected effects such as those of a polarized seeing halo may mimic disk emission, particularly close to the star. Thus it is possible that the true SR21 radial profile flattens out in the interior and we are mistaking uncorrected PSF effects for disk emission in the inner region. However, all pixels in the saturated region were excluded from the radial profiles, and the PSF artifacts do not extend beyond the innermost bin in our radial profile plots, so this is unlikely.

The computational expense of scattered light models means that we did not investigate the scattered light parameter space as systematically as the SED parameter space, and a more thorough and systematic exploration of the available parameter space is warranted to ensure that our solution is unique.

It could be argued, however, that ADI observations with more field rotation, higher AO performance (Strehl Ratio) H-band observations and/or further multiwavelength data that resolve the mysterious SR21 "gap" for other varieties of disk grain are equally warranted. Observing methods that are sensitive to sharp disk features (e.g., ADI) could help verify whether there is any local excess in scattered light at the 36AU dust wall in addition to resolving the potential companion.

It is also worth noting that all of our disk models have assumed axisymmetry. Unresolved spiral structure, emission from a companion, disk warps, etc. could be present in the SR21 disk and may have an effect on the radial profile that has not been examined in this analysis.

Uniformity of grain composition as a function of radius is another aspect of the models that is perhaps unphysical. Compositional differences, such as higher volatile abundances in the gapped region, could allow for a more marked effect of density on surface brightness in the inner disk, allowing optically thick large- β large- α models to increase more steeply towards the star.

2.6 Conclusion

We have spatially resolved the SR21 transitional disk for the first time at NIR wavelengths. Our data show the signature of an optically thin small grain disk component, and place strong constraints on the underlying distribution of any optically thick surface grains.

Our Subaru data resolve the disk in scattered polarized light for stellocentric $r\geqslant0'.1$ ($\gtrsim12$ AU). Extended polarized emission is present above noise level out to $r\sim0'.6$ (~80 AU) and exhibits a centrosymmetric pattern of polarization vectors. The azimuthally-averaged radial profile of the polarized intensity emission reveals that the disk flux drops off steeply (r^{-3}) and smoothly with radius. The disk is also evident in the PSF subtracted total intensity image, and the range of viable scale factors for PSF subtraction give us an estimated polarization percentage for the disk of $10\% \lesssim P \lesssim 27\%$.

We confirm with an independent model that scenarios in which large midplane grains are depleted by a factor of $\sim 10^{-6}$ inside of a ~ 36 AU gap reproduce the SMA observations of A11; however, small grains cannot be depleted to the same degree inside the cavity.

Our analysis eliminates a variety of scenarios under which there is a discontinuity at the 36AU large grain gap radius, both in terms of physical likelihood and ability to fit the observational data. In order to be consistent with the lack of an equivalent gap in the Subaru HiCIAO H-band polarized intensity images and radial intensity profiles, there must be a decoupling of large and small grain disk parameters in the SR21 disk. Our data rule out scenarios in which the SR21 disk

contains an optically thin small grain dust gap with a wall at 36AU.

Although we can neither confirm nor deny its existence definitively, we present several pieces of anecdotal evidence in favor of a companion in the SR21 disk. These include an intensity excess near the appropriate location of E09's hypothesized companion, a deviation from centrosymmetry in the polarization vectors in this region, and an apparent migration of the disk major axis from inner to outer disk, suggesting a disk warp or potentially unresolved spiral structure.

We conclude that the best fitting models for the H-band radial polarized intensity profile, the observed sub-mm emission and the multiwavelength SED for the SR21 disk invoke three separate components. These axisymmetric, uniform radial grain composition scenarios suggest an optically thin envelope or atmosphere component lying above an optically thick, geometrically smooth small grain disk component. Further constraints on the geometry of this very interesting transitional disk may be obtained by follow-up at higher spatial resolutions and at other wavelengths.

CHAPTER 3

SEEDS ADAPTIVE OPTICS IMAGING OF THE ASYMMETRIC TRANSITION DISK OPH IRS 48 IN SCATTERED LIGHT

We present the first resolved near infrared imagery of the transition disk Oph IRS 48 (WLY 2-48), which was recently observed with ALMA to have a strongly asymmetric sub-millimeter flux distribution. H-band polarized intensity images show a \sim 60AU radius scattered light cavity with two pronounced arcs of emission, one from Northeast to Southeast and one smaller, fainter and more distant arc in the Northwest. K-band scattered light imagery reveals a similar morphology, but with a clear third arc along the Southwestern rim of the disk cavity. This arc meets the Northwestern arc at nearly a right angle, revealing the presence of a spiral arm or local surface brightness deficit in the disk, and explaining the East-West brightness asymmetry in the H-band data. We also present 0.8-5.4 μ m IRTF SpeX spectra of this object, which allow us to constrain the spectral class to A0±1 and measure a low mass accretion rate of $10^{-8.5} M_{\odot}/yr$, both consistent with previous estimates. We investigate a variety of reddening laws in order to fit the mutliwavelength SED of Oph IRS 48 and find a best fit consistent with a younger, higher luminosity star than previous estimates.

3.1 Introduction

Circumstellar disks have long been hypothesized to be a breeding ground for planets, however only recently has the spatial resolution and inner working angle necessary to directly test this hypothesis been achievable. Giant planets are expected to perturb the disk as they form, changing the spatial distribution of gas and dust by opening gaps and supporting wholly or partially cleared cen-

tral cavities spanning tens of AU (Andrews et al. 2011; Williams & Cieza 2011). These cavities are also detected in unresolved photometry of young stellar objects (YSOs) as "missing" hot thermal emission in the spectral energy distribution (SED) (Strom et al. 1989; Calvet et al. 2005).

Although planet formation is a leading hypothesis for how these observed cavities are formed (Zhu et al. 2011; Dodson-Robinson & Salyk 2011), the waters are muddied by several competing, and evidently co-occurring, mechanisms for creating these gaps (Alexander et al. 2013), which also include grain growth (Dullemond & Dominik 2005; Birnstiel et al. 2012) and photoevaporation (Pascucci & Sterzik 2009; Clarke et al. 2001). Each mechanism makes different predictions for the size, structure, symmetry, and depletion factor of various disk constituents (gas, large dust grains, small dust grains), thus high resolution multiwavelength data that probe both gas and dust properties and morphologies are needed to disentangle them.

Recent observations have revealed dust disk distortions that, unlike cleared cavities, are not conspicuous in the SEDs. These include spiral arms (Muto et al. 2012; Grady et al. 2013; Casassus et al. 2012; Rameau et al. 2012), divots (Hashimoto et al. 2011; Mayama et al. 2012), and asymmetries (van der Marel et al. 2013; Casassus et al. 2013; Fukagawa et al. 2013; Isella et al. 2012; Pérez et al. 2014). Per current understanding of disk clearing mechanisms, these deviations from axisymmetry are most easily explicable via interactions with giant planets embedded in these disks, although internal disk dynamical processes cannot be excluded. These include gravitational instability (Jang-Condell & Boss 2007), photoelectric instability (Lyra & Kuchner 2013), Rossby wave instability (Lyra & Kuchner 2013) and MRI-induced asymmetries (Heinemann & Papaloizou 2009).

Oph IRS 48 (WLY 2-48 Wilking et al. 1989) is an $A0_{-1}^{+4}$ (Brown et al. 2012a) star

in ρ Oph, with high foreground extinction (A_V =12.9, McClure et al. 2010) at d=121 pc (Loinard et al. 2008). The first resolved images of this disk are from Geers et al. (2007), who imaged it at five MIR bands from 8.6 to 18.7 μ m with VISIR on the VLT. While their shorter wavelength PAH images (7.5-13 μ m) all peak at the star, though with varying degrees of spatial extension, the 18.7 μ m thermal emission reveals a ring-like structure peaking at 55AU from the star. They attribute the lack of a strong 10μ m silicate feature in the spectrum of this object to clearing of small μ m sized dust inside the cavity seen at 18.7 μ m, though they note that centrally peaked and spatially extended PAH emission at other MIR bands indicates that the cavity cannot be wholly cleared of material. Indeed, subsequent modeling of PAHs in Oph IRS 48 by Maaskant et al. (2014) find that the spectrum is dominated by highly ionized PAHs in optically thin regions inside of the disk cavity, and that the neutral PAH contribution from the outer disk is very low.

Brown et al. (2012a) obtained high resolution (R \sim 100,000) CRIRES spectra of Oph IRS 48 at the 4.7 μ m CO fundamental rovibrational band, and found a \sim 30 AU ring, which they interpreted as arising from a dust wall at this radius. They modeled the excitation of the M-band CO emission and found a good fit to the optical spectrum with a 2.0 \pm 0.5 solar mass star with L=14.3 L $_{\odot}$.

Subsequent ALMA 0.44mm observations of Oph IRS 48 (van der Marel et al. 2013) revealed that the sub-millimeter dust distribution is highly asymmetric, and is concentrated in an arc subtending $\sim 100^\circ$ on the south side of the star, with a factor of >130 contrast between that structure and the north side of the disk, where no emission is detected. 12 CO 6-5 data, on the other hand, are consistent with a continuous, rotating Keplerian disk. Though a slight North/South asymmetry was noted in this data and in Brown et al. (2012b), potential effects of foreground absorption called its reality into question. van der Marel et al. (2013) in-

terpreted the Southern 0.44mm structure as a dust trap associated with an asymmetric pressure bump, potentially caused by a 10 Jupiter mass body located at 17 < r < 20 AU. They note that smaller dust grains are less efficiently trapped under such a model (Birnstiel et al. 2013), and are expected to be distributed more axisymmetrically about the star. Furthermore, warm H_2CO emission was detected in the southern part of the disk (van der Marel et al. 2014), hinting at either a density or temperature contrast between the north and the south.

A subsequent study by Bruderer et al. (2014) conducted detailed gas modeling on these data. These models point to an axisymmetric disk with two distinct regions inside of the observed $18.7\mu m$ cavity - one with r<20AU in which the gas is heavily depleted and one with 20AU < r < 60AU in which the gas surface density is only marginally depleted relative to the outer disk. This complex depletion of the gas surface density inside of the disk cavity is most consistent with clearing of the disk by one or more planetary companions.

In this paper we report the first scattered-light imagery of the disk of Oph IRS 48 and compare it with previous observations, described above, in an effort to develop a consistent multi-wavelength picture of the system.

3.2 Observations and Data Reduction

3.2.1 Subaru/HiCIAO

Polarized Differential Imaging (PDI) of Oph IRS 48 was done in H-band (1.6 μ m) on 2013 May 19 and in Ks band (2.2 μ m) on 2014 June 9 with the high-contrast imager HiCIAO (Tamura et al. 2006; Hodapp et al. 2006) and adaptive optics system AO188 (Minowa et al. 2010) on the 8.2m Subaru telescope on Mauna Kea. These observations were conducted as part of the Strategic Exploration of Exoplanets and Disks with Subaru (SEEDS) survey, which began in October 2009. Oph IRS

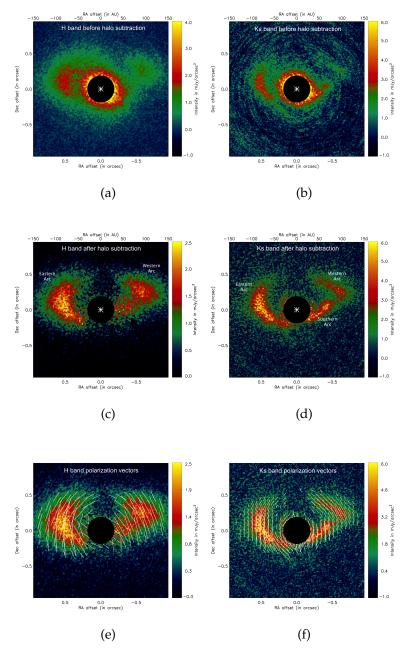


Figure 3.1: Oph IRS 48 HiCIAO H-band (left) and Ks-band (right) PI images. In all cases, an r=20 (0.19") pixel mask has been applied to cover the stellar residual. Polarization vectors are shown overplotted in bins equivalent to half of the FWHM of the star at each wavelength, and the length of each vector corresponds to the strength of the signal in that bin.

48 is too faint (R=16.66, Erickson et al. 2011) for use as a natural guide star (NGS), therefore a sodium laser guide star (LGS) was used to obtain high order correction, with Oph IRS 48 itself serving as the natural tip-tilt guide star.

The data were taken in "qPDI" mode, in which a double Wollaston prism is used to split the beam into four $\sim 5 \times 5$ " channels, two each of o and e-polarizations. The splitting of each polarization state into two separate channels reduces saturation effects on the Hawaii 2RG detector. In order to obtain full polarization coverage and minimize artifacts, a half waveplate was rotated to four different angular positions (0°, 45°, 22°.5 and 67°.5) to create each PI image. This cycle was repeated 14 times at H band and 17 times at Ks band, however the Ks data included 2 waveplate cycles during a seeing spike, which have been excluded in this analysis. With 30 seconds per exposure, this translates to a total of 28 minutes of integration time at H band and 30 minutes at Ks band.

Each image was bias subtracted, flat fielded and cleared of bad pixels in the standard manner for infrared data analysis using custom IDL scripts. Observations of the globular cluster M5 were then used to create a distortion solution for each channel, which was applied before the images were rotated to a common on-sky geometry.

Each individual channel was aligned using Fourier cross-correlation. These aligned images were combined to create Stokes Q and U images using standard differential polarimetry methods, namely adding together each set of two identically polarized channels and then subtracting these combinations from one another. The four channels were also added directly together to form a total intensity image. The FWHM of the star in the final, median-combined total intensity images was 0″23 (24 pixels) at H and 0″12 (12 pixels) at Ks. The factor of two improvement in FWHM at Ks is largely due to the better performance of the AO

system at longer wavelength, as the seeing was approximately 0'5 on both nights.

The polarized intensity images shown in Figures 3.1a (H-band) and 3.1b (Ksband) are median combinations of Polarized Intensity (PI) images computed via $PI = \sqrt{Q^2 + U^2}$ for each waveplate cycle. These images suffer from a residual polarized halo as described in detail in Follette et al. (2013b). In the case of Oph IRS 48, which is embedded fairly deeply in the cloud (A_V =12.9), polarization due to intervening dust may also contribute significantly to this effect. Figures 3.1c and 3.1d show the same images after removal of the polarized halo. The halo was computed by calculating the polarization P and polarization angle θ_p in the region 0 < r < 40 pixels (0'38) and $-180^{\circ} < PA < -150^{\circ}$ at H-band and 0 < r < 15 pixels (0'.14) and 160° < PA < 200° for Ks band. These regions were selected iteratively from among the regions near the star with little or no disk emission. The quality of halo subtraction was judged in each case by the strength of the stellar residual remaining at the center of the image. The halo values that minimized these stellar residuals were P=0.6% and PA=41° at H band and P=0.5% and PA=36° at Ks band. The halo was removed most effectively in the H-band data, while the K-band data, though significantly improved through this process, still house a fairly strong stellar residual at the center of the cavity that may be reflective of real inner disk emission. These residuals have been masked in all images.

Polarization angles were calculated for each halo-subtracted image according to the formula $\theta_p = 0.5 \arctan(U/Q)$ and median combined. These are shown overplotted on the PI image in 12/6 pixel (0'.11/0'.06) bins for H/Ks (half of the stellar FWHM in each case) after halo subtraction in Figures 3.1e and 3.1f. The length of each vector has been scaled according to the magnitude of the polarized intensity in that bin, and vectors in bins at or below the background level have been clipped. Halo subtraction results in strong centrosymmetry, as expected for

polarized disk emission.

A PSF star, HD148835, was also observed at H band on 2013 May 19 just after Oph IRS 48 observations were completed. Five waveplate cycles of 15 second images were completed for this brighter star for a total of 5 minutes of integration time. The same reduction procedure was followed for this data set in order to verify that nothing mimicking disk structure was induced by the halo subtraction procedure. In the case of the PSF star, halo subtraction both reduces the strength of the stellar residual and randomizes the direction of the polarization vectors, as expected for a star without any circumstellar material.

We also attempted classical PSF subtraction using the total intensity images of Oph IRS 48(V=17.7, R=16.7) and HD148835 (V=10.07, R=9.26), but found the PSFs to be poorly matched and were not able to isolate disk emission in this way. The large difference in brightness between the two stars should result in significantly different AO performance and therefore PSF structures, so this poor match is to be expected.

3.2.2 IRTF/SpeX

We obtained NIR spectra of Oph IRS 48 on 2013 May 16 using the SpeX (Rayner et al. 2003)) spectrograph on NASA's Infrared Telescope Facility (IRTF). The spectra were recorded using the echelle grating in both short-wavelength mode (SXD, 0.8-2.4 μ m, Figure 3.2a) and long-wavelength mode (LXD, 2.3-5.4 μ m, Figure 3.2b) using a 0'8 slit. The spectra were corrected for telluric extinction and flux calibrated against the A0V calibration star HD 145127 using the Spextool data reduction package (Vacca et al. 2003; Cushing et al. 2004). In addition to the 0'8-slit spectra, we also recorded data with the SpeX prism disperser and a wide 3.0" slit, which allows us to retrieve the absolute flux levels when the sky transparency is good and the seeing is 1" or better (it was 0'6 arcsec at the time the data were

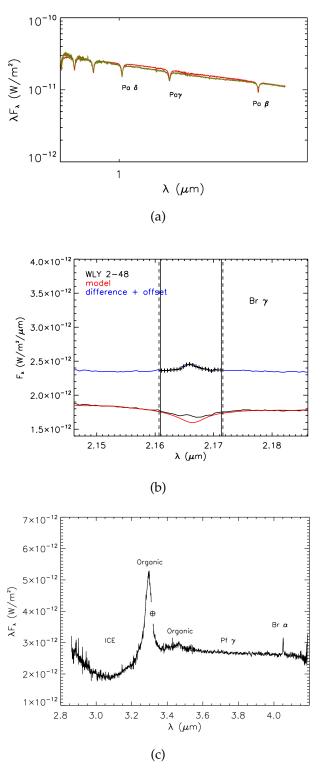


Figure 3.2: (a) 0.8-2.4 μ m SpeX spectrum, with the spectrum of another A0 star, SAO206463 overplotted.; (b) Pa β line fit used together with a similar fit to the Br γ line to derive a weak accretion rate of \dot{M} =10^{-8.5} M $_{\odot}$ /yr for Oph IRS 48.; (c) 2.3-4.3 μ m SpeX spectrum, with pertinent lines labeled.

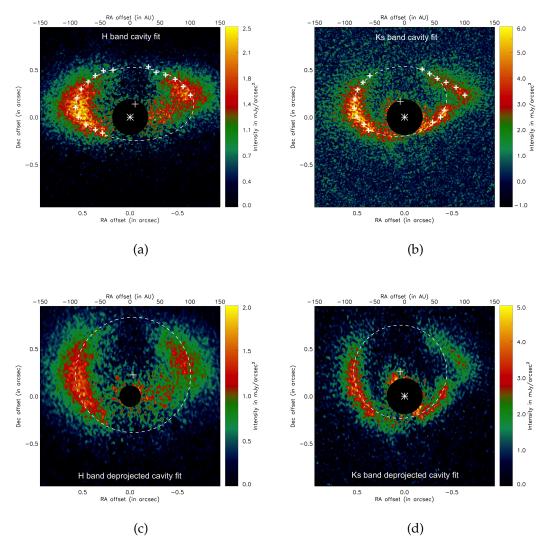


Figure 3.3: Top panels: Elliptical fits to the Oph IRS 48 scattered light cavity at H (left) and Ks (right) bands. The locations of the peaks for each 10° azimuthally binned radial profile are shown with white crosses. The best fit ellipse to these points is shown overplotted as a dashed white line in each case. The center of that ellipse is shown with a large white cross. As discussed in the text, the Ks band cavity fit is significantly more robust, as it is (i) higher resolution data, (ii) brighter and (iii) the morphology is clear enough at this wavelength that profiles dominated by the asymmetry can be excluded. Bottom panels: The same images, but deprojected along PA=97.4 $^{\circ}$ to account for an inclination of 50° and with the deprojected circular cavity fit shown. The large northern offset between the stellar centroid and the cavity center is discussed in detail in Section 3.4.5

acquired).

3.3 Results

In this section, we present the results of our analysis of the new HiCIAO and SpeX data. In Section 3.1, we discuss the geometry of the disk in scattered light. In Section 3.2, we present fits to the Oph IRS 48 cavity. In section 3.3, we discuss the azimuthal symmetry, or lack thereof, in the data based on radial profiles through the deprojected disk image. In section 3.4, we present a revised age and spectral type derived from our analysis of the multiwavelength SED, and in section 3.5, we present accretion estimates derived from our SpeX spectra.

3.3.1 Observed Disk Geometry

The disk as seen in PI extends approximately $1''2\pm0''2$ (\sim 145±25AU) from the star at H-band and $1''.0\pm0''.1$ (\sim 120±12AU) at Ks-band. This is similar to the extent of 12CO gas emission traced by ALMA (Bruderer et al. 2014, \sim 160AU), as expected for small grains, which should be tightly coupled to the gas in the disk. Emission from the cavity rim begins \sim 0''3 (\sim 20AU) from the center of the cavity and peaks at \sim 0''5±0''.1 (55±15AU) at most azimuths for both wavelengths. Disk parameters derived from fits to this cavity are described in detail in Section 3.3.2.

The scattered light emission from this disk has marked asymmetries both in the North/South and East/West directions that cannot be explained purely by inclination effects, as discussed in Section 3.4.4. In particular, the H-band emission is concentrated in two arcs, which we will call the "Eastern" and "Western" arcs from here forward. The Eastern arc extends for $\sim 100^\circ$ from the Northeast to the Southeast of the star (20<PA<120). The Western arc is much shorter, and lies entirely northward of the stellar position, extending just 40° along PAs -70<PA<-40.

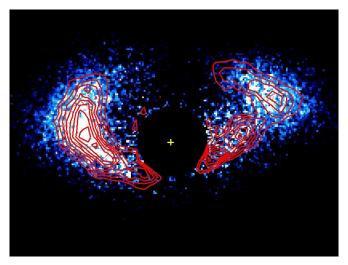


Figure 3.4: Ks band polarized intensity contours overplotted on the H band polarized intensity image of OPH IRS 48. The improvement in image quality in the K band image clarifies that the excess emission seen to the West of the star at H band is probably the Southern disk rim, clearly visible in the K band image.

With one notable exception, the morphology of the K-band PDI data are similar to the H-band data. The location and extent of emission along the Eastern and Western arcs is similar between the two image sets, however the Ks-band data reveal a third arc in the South that is not readily visible at H-band along -150° <PA<-70°. The morphology of this arc relative to the others, and its implications for disk structure and composition are discussed in detail in section 3.4.6.

3.3.2 Cavity Morphology

To asses the morphology of the Oph IRS 48 cavity in a quantitative sense, we created radial profiles binned to half of the FWHM of the stellar PSF for each image set (12 pixels/0'.11 at H, 6 pixels/0'.06 at Ks) for each 10° azimuthal slice in the disk. We fit the peak of each radial profile with a Gaussian and inspected the quality of fit by eye. Only those profiles well fit by a Gaussian (namely those

with a clear cavity rim feature present) were included. The peak location for each profile was weighted by its intensity and fit with an ellipse to produce a least-squares elliptical cavity fit. These fits are shown overplotted on the H and Ks band data in Figures 3.3a and 3.3b respectively.

In the Ks-band PI data, the Southern and Western arcs appear to intersect one another at nearly a right angle, and the radial distance from the star to the rim also jumps rapidly at this intersection. We interpret this as non-axisymmetric structure in the disk - either a spiral arm that extends from the Western rim out to the Southwest or a local surface brightness deficit at the cavity wall. Both scenarios are investigated in detail in Section 3.4.6. As revealed in the overlay in Figure 3.4, the majority of the Western H-band arc lies along this spiral arm, which tends to degrade the quality of the cavity fit at this wavelength, however we report its parameters here for completeness.

The best-fit ellipse to the H-band data, shown in Figure 3.3a, has an eccentricity of 0.78 (a=0'.62, b=0'.39), a PA of 92°, and is offset relative to the stellar centroid by +6 pixels (0'.06) in RA and +15 pixels (0'.14) in Declination.

The higher resolution of the Ks band data allows us to exclude PAs of -50°, -60° and -70° from the elliptical cavity fit, since they lie along the asymmetry. This leads to an elliptical cavity that is well fit to all of the other radial profile peaks. We thus consider the fit to the Ks-band cavity to be much more robust than the H-band fit.

The best fit ellipse at Ks-band, shown in Figure 3.3b has an eccentricity of 0.66 (a=0'49, b=0'37, consistent with an inclination of $42\pm10^{\circ}$), a PA of $97\pm2^{\circ}$, and a center that is offset from the stellar residual by -5 pixels (0'05) in RA and +18 pixels (0'17) in Declination. These values are in keeping with the published values for the inclination (50°) and PA (100.3°) of the disk. The inclination, in particular,

is not well constrained by our data, and so we adopt the literature value of 50° in deprojecting our images. The offset of the cavity center from the stellar residual is discussed in detail in Section 3.4.5.

The K band fit corresponds to a deprojected circular cavity size of $59\pm10\mathrm{AU}$ at 121pc, where error bars were derived from the average width of the Gaussian fit to each radial profile weighted by the peak. This is also consistent with the radius of the cavity as seen at other wavelengths. The circular cavity fits are shown overplotted on the deprojected disk images in Figures 3.3c and 3.3d

The degree of depletion of small grains inside the cavity is not well constrained by our data and is severely limited by stellar residuals. The smallest ratio between the trough (inside the cavity) and peak (at cavity rim) of a radial profile in our data is 0.40 at H-band and 0.33 at Ks-band, thus the NIR-scattering grains inside of the cavity are depleted by, at minimum, a factor of 2.5-3 relative to the outer disk.

3.3.3 Radial Profiles and Azimuthal Structure

The disk images were deprojected to remove the effect of the i=50° inclination, and allow for a direct comparison of radial profiles through the disk in any given direction. These are shown for H and Ks band in Figure 3.5a, where in each case the disk has been deprojected along PA=97.4° for an inclination of 50°.

The profiles were taken with respect to the cavity center and not the stellar residual, and the Ks-band data were convolved with a 5.65 pixel (0'05) Gaussian in order to simulate similar AO system performance and allow for a direct comparison of the two wavelengths. These choices are described in detail in Appendix C.

The asymmetry in the disk is immediately apparent in Figure 3.3.3, as the profile along the Western major axis (dashed black line) peaks \sim 0'.15 more distant

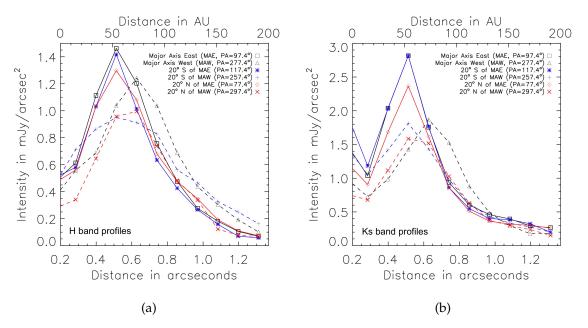


Figure 3.5: Radial profiles taken through the H (left) and Ks (right) images. Azimuthal asymmetry is especially apparent along the Western major axis (dashed black lines), where profiles at both wavelengths peak ~0.15" farther from the cavity center, though the asymmetry is also hinted at in the Southwestern profiles (red dashed lines). The blue scattering properties of the asymmetry are readily apparent - the H-band Western major axis profile peaks at ~85% of the Eastern major axis value, while at Ks-band this value is ~25% lower (~60%). Profiles were taken with respect to the apparent cavity center rather than the location of the stellar residual, and this choice is described in detail in Appendix C. The disk images were deprojected to account for an inclination of 50° along PA=97.4°, and the Ks band image was convolved with a 5.65 pixel (0.05") Gaussian in order to simulate similar Strehl-ratio imagery. In all cases, solid lines represent profiles through the Eastern half of the disk and dashed lines represent profiles through the Western half of the disk. Black profiles were taken along the disk major axis and blue and red profiles were taken 20° North and South of the major axis, respectively.

Wavelength(s)	Instrument	Reference
(μ)		
0.43, 0.64	NOMAD	Zacharias et al. (2005)
0.65, 0.8	Hydra	Erickson et al. (2011)
1.2, 1.6, 2.2	2MASS	Cutri et al. (2003)
3.4, 4.6, 12, 22	WISE	Wright et al. (2010)
3.6, 4.5, 5.8	Spitzer IRAC	van Kempen et al. (2009)
5.9-36.9	Spitzer IRS	McClure et al. (2010)
18.7	AKARI IRC	Yamamura et al. (2010)
60-181	Herschel PACS	Fedele et al. (2013)
70	Spitzer MIPS	van Kempen et al. (2009)
450, 850, 1300	SCUBA	Andrews & Williams (2007)

Table 3.1: Previously published photometry and spectra of Oph IRS 48.

from the cavity center than the other profiles at both wavelengths. The profiles 20° South of the Western major axis are also somewhat deviant, peaking between those of the rest of the disk (at $\sim 0'/5$) and the Western major axis.

The Ks-band profiles are a factor of 2 greater in absolute intensity than the H-band profiles. This is to be expected given the region of high extinction in which Oph IRS 48 resides. Indeed the 2MASS K-band magnitude (7.582) of the star is 3 times brighter than the H-band magnitude (8.815). The profiles are generally similar, however, in relative brightness between the H and Ks band data, except in the region of the asymmetry. Deviations in this region are discussed in detail in Section 3.4.6.

3.3.4 Oph IRS 48 Age and Spectral Type

In order to derive the intrinsic unreddened properties of IRS 48 we first investigated the spectral classification using the SpeX spectrum of the star. We compared the spectrum directly to similar data on other stars. The spectrum of IRS 48 exhibited noticeable line emission within the photospheric features at wavelengths longer than 1.0 μ m (Pa γ , Pa β , etc.) so only photospheric lines at shorter

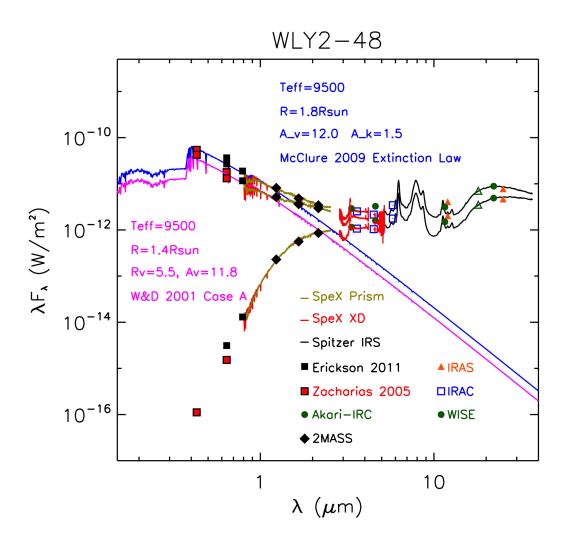


Figure 3.6: The spectral energy distribution (SED) of Oph IRS 48. The lower set of data are observed values, while the upper sets represent the data de-reddened using two different reddening laws. Both reddening laws assume $\log(g)$ =4.0 and a distance of 121pc. The magenta curve is an A0 stellar template dereddened according to Weingartner & Draine (2001) Case B with R_V =5.5 and A_V =11.5, which yields L=14.3L $_{\odot}$. This was the assumption of Brown et al. (2012b). The blue line is an A0 template dereddened according to McClure (2009), which may be more appropriate for this high extinction region. This reddening assumption yields a factor of \sim 1.7 higher dereddened luminosity of 23.6L $_{\odot}$. The upper curve tracks the stellar photosphere more closely out through the NIR, calling the presence of a hot dense inner disk component into question

wavelength were employed (Pa δ , Pa ϵ , Pa 09, etc.). The lines were found to be a good match to SAO 206463 (A0V, shown overplotted in Figure 3.2a) so we adopted a spectral type of A0 \pm 1. This is in agreement with the spectral type determined by Brown et al. (2012a).

To place Oph IRS 48 on a pre-main sequence evolutionary track, the absolute luminosity needs to be determined by dereddening the data. We investigated dereddening the spectrum using two reddening laws, that of Weingartner & Draine (2001) and that of McClure (2009). For the former, we used their Case B reddening law, with R_v =5.5. In order to match the spectral slope of a A0 star, an extincion of A_V =11.5 mag was required. This was the methodology used by Brown et al. (2012b) in obtaining their luminosity estimate of 14.3 L_{\odot}.

However, McClure (2009) found that for $A_K > 1$ mag in dark clouds, a much "greyer" extintion curve applies, with R_V closer to 8.5. Using this extinction law, we derive $A_V = 12.0$ mag (determined independently from, but consistent with, the McClure (2009) estimate of $A_V = 12.9$ mag for this object) and $A_K = 1.7$ mag.

The greater total reddening correction using this extinction law requires a larger stellar radius and luminosity, the latter increasing by a factor of \sim 2 compared to that retrieved by Brown et al. (2012b) using the Weingartner & Draine (2001) Case B extinction curve, to 23.6L $_{\odot}$. Hence, there is a large intrinsic uncertainty in the stellar luminosity derived, through the choice of extinction law adopted.

A comparison between the two dereddened spectra is shown in Figure 3.6. While both provide reasonable fits to the existing photometry and spectra, the McClure (2009) fit follows the stellar photosphere much more closely out through the NIR. This calls into question the existence of a hot inner disk component in this object, which has been inferred from the presence of excess emission at these

wavelengths (e.g. Bruderer et al. 2014; Maaskant et al. 2013).

We note that neither reddening law provides a particularly good fit to the available optical photometry (Zacharias et al. 2005; Erickson et al. 2011). Given the inconsistency between the two published R-band magnitudes, we believe that additional photometry or spectroscopy at visible wavelengths is advisable in order to further constrain the best choice of reddening law for this object.

These two different reddening laws also put the star at different places along pre main sequence evolutionary tracks. The Weingartner & Draine (2001) reddening law adopted by Brown et al. (2012b) puts the star at an age of \sim 15Myr along the 2.0 M $_{\odot}$ PMS evolutionary track of Siess et al. (2000). Our factor of 1.7 higher luminosity using the McClure (2009) law puts the star at an age of \sim 8Myr along a 2.2M $_{\odot}$ evolutionary track.

We can derive well-fitting SEDs up to significantly higher stellar luminosities, corresponding to masses up to $2.5 M_{\odot}$ and ages as young as 5 Myr, this the error in our (and other) luminosity determinations is quite large. We note that although Pre-Main Sequence evolutionary models do not agree well in general, especially among higher mass stars (e.g. Palla & Stahler 1993; D'Antona & Mazzitelli 1994; Siess et al. 2000), this difference in inferred age as a result of reddening assumptions is largely model independent.

Although we have used Siess et al. (2000) here to match Brown et al. (2012b), the age and mass inferred for Oph IRS 48 should be systematically younger and higher respectively for all evolutionary models under this McClure (2009) reddening assumption due to the higher inferred luminosity of the star. If we use our derived luminosity and measured temperature and compare them to other available pre-main sequence isochrones for A stars (e.g. Palla & Stahler 1999; Dotter et al. 2008), our age estimate can fall to just a few Myr.

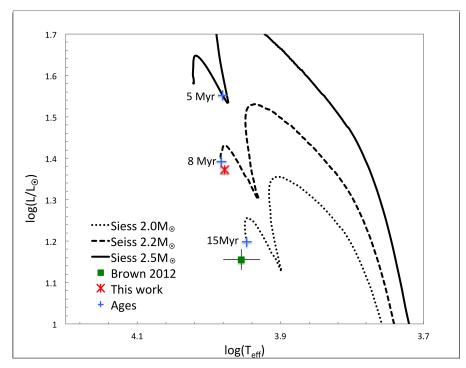


Figure 3.7: An H-R Diagram showing the location of Oph IRS along Siess et al. (2000) PMS evolutionary tracks. The green square shows the location of Oph IRS 48 according to the Brown et al. (2012b) reddening assumption, while the red asterisk shows the location of Oph IRS 48 derived from our best-fit dereddening. The factor of \sim 1.7 difference in the luminosity derived from the two reddening assumptions results in different assumptions about the mass and age of Oph IRS 48. Brown et al. (2012b)'s Weingartner & Draine (2001) law puts Oph IRS 48 at \sim 15Myr along the 2.0M $_{\odot}$ evolutionary track. We find reasonably well-fitting SEDs with Luminosities as high as 45L $_{\odot}$, corresponding to a 2.5M $_{\odot}$ star at 5Myr, and our best fit McClure (2009) law puts it at \sim 8Myr along the 2.2M $_{\odot}$ evolutionary track. Whether or not these isochrones are the most appropriate for young A-stars is a matter of contention, however our higher derived luminosity puts Oph IRS48 along consistently higher mass isochrones and at a younger age than previous work regardless of chosen evolutionary model, as described in the text.

Age estimates for other rho-Ophiuchus members within 10 arcminutes of Oph IRS48 range from <0.1Myr to 14Myr (Wilking et al. 2005), so cannot place any meaningful constraints on evolutionary models, although the younger age that we derive using a Siess et al. (2000) evolutionary track is more closely consistent with the median age of 2-5 Myr for stars in the region (Wilking et al. 2008).

3.3.5 Accretion Rate Estimate

We were also able to estimate the stellar accretion rate from the SpeX data by fitting the Pa β and Br γ lines, and a sample fit is shown in 3.2b. To do this, we subtracted our best fit stellar template (A0) from the dereddened SpeX spectrum to get the accretion line luminosities. Following Muzerolle et al. (1998) for both lines and Calvet et al. (2004) for an independent estimate of Br γ , we derive log(L_{acc})=-1.1 for Pa β and log(L_{acc})=-1.0 for Br γ , which translate to \dot{M} =10^{-8.5} and 10^{-8.4} M $_{\odot}$ /yr respectively using the relationship:

$$\dot{M} = \frac{L_{acc}R_*}{0.86GM_*} \tag{3.1}$$

Our best estimates for the accretion luminosity and mass accretion rate are thus $\log(L_{acc})$ =-1.1 and \dot{M} =10^{-8.5} M_{\odot} /yr for Oph IRS 48. These are slightly lower than the $\log(L_{acc})$ =-0.9 and \dot{M} =10^{-8.4} derived in Salyk et al. (2013) using the Pfund β line at 4.7 μ m, but well within the \sim 0.5dex spread in their empirical relationship.

3.4 Discussion

3.4.1 Monochromatic Radiative Transfer Modeling wth Sprout

We began our modeling efforts by performing monochromatic radiative transfer calculations using a Monte Carlo code developed for use with SEEDS observations (the Sprout code¹, Takami et al. 2013) in order to constrain disk parameters,

¹Available at http://www.asiaa.sinica.edu.tw/~jkarr/Sprout/sprout.html

particularly flaring and scale height, for full MCRT modeling. With Sprout, the modeled system consists of an illumination source with an axisymmetric circumstellar disk following a standard flared accretion disk prescription (e.g., Shakura & Sunyaev 1973; Lynden-Bell & Pringle 1974) and described in cylindrical coordinates (r,z) by:

$$\rho(r,z) = \rho_0 \left[1 - \sqrt{\frac{R_*}{r}} \right] \left(\frac{R_*}{r} \right)^{\alpha} \exp \left\{ -\frac{1}{2} \left[\frac{z}{h} \right]^2 \right\}, \tag{3.2}$$

where ρ_0 is a constant to scale the density, R_* is the stellar radius, α is the radial density exponent, and h is the disk scale height. The scale height h increases with radius as $h = h_0(r/r_0)^{\beta}$, where β is the flaring index ($\beta > 0$) and h_0 is the scale height value at a specific radius r_0 , which we fix at 60AU to match our data. This schema is commonly used by other researchers (e.g., Cotera et al. 2001; Whitney et al. 2003b; Robitaille et al. 2006; Follette et al. 2013b; Takami et al. 2013).

We also use the common assumption $\alpha=\beta+1$ (e.g., Robitaille et al. 2006; Robitaille et al. 2007; Follette et al. 2013b; Takami et al. 2013). This provides a surface density of $\Sigma \propto r^{-1}$, the same prescription used by Bruderer et al. (2014). For these simulations, we use the dust model developed by Cotera et al. (2001) to explain the scattered light in the HH 30 disk.

The free parameters for the density distribution of the disk are therefore ρ_0 , h_0 , β . We assume that the mass of small dust, which is responsible for scattered light from the disk surface is $2.4 \times 10^{-6} \mathrm{M}_{\odot}$, 15% of the total (large + small grain) dust mass derived in Bruderer et al. (2014). For a given β , we adjust the disk scale height h_0 to match the Northward offset of the observed PI distribution from the star, and ρ_0 to achieve the above dust mass. We find that the flaring parameter β primarily affects the brightness at outer radii compared with that at inner radii and that we generally need values of $\beta > 1$ to reproduce our data.

We used 10^6 photons for each simulation, collect the scattered light at viewing

angles of $50^{\circ} \pm 10^{\circ}$, and convolve the simulated images with a Gaussian whose FWHM matches the observations. Photons that do not interact with the disk are not collected. Figure 3.8 shows one of the best fit models (β =1.3, h_0 =15AU at r=60 AU).

This β is similar to the model by Bruderer et al. (2014) used to explain images at millimeter and mid-infrared wavelengths (1.2), and is within the typical range for a flared disk (to first approximation, temperature varying as $r^{-1/2}$ in a disk typically leads to a scale height proportional to $r^{5/4}$), so we adopt it for all further modeling efforts.

The scale height required in the Sprout model, however, is \sim 2x the Bruderer et al. (2014) small grain model (8.4 AU), corresponding to a midplane temperature difference of 4x, which is not explicable simply by the greater stellar luminosity derived in this work (our factor of 1.7 increase in luminosity corresponds to just a factor of 1.2 in temperature at a given location). As this large scale height is required to reproduce the observed Northward offset, we utilize large gap edge scale heights in our future modeling efforts, though we were somewhat limited in our ability to do so given the constraints of the models, as described in detail in Section 3.4.5.

3.4.2 SED modeling

Informed by the results of our Sprout simulations, we ran a suite of disk models using the updated 2013 version of the three dimensional Whitney monte carlo radiative transfer (MCRT) code (see Whitney et al. 2013, 2003a,b, for details). We began our modeling investigation by attempting to fit the multiwavelength observed SED of Oph IRS 48, which includes a wide range of broadband photometry as well as optical, NIR and MIR spectra. Sources for this data are listed in Table 3.1, and our best fit model is overplotted on data from the literature in Fig-

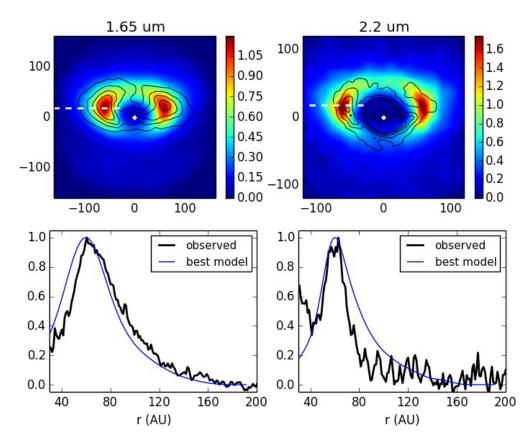


Figure 3.8: Top: Simulated images of Oph IRS48 at H-band (left) and Ks-band (right) generated by the monochromatic radiative transfer Sprout code. The data are shown overplotted in each case as black contours. The location of the illumination source is marked with a white diamond, and the locations where profiles were extracted are marked with white dashed lines, corresponding roughly to the "major axis" relative to the apparent cavity center. The combination of an elevated flaring parameter (β =1.3) and a larger scale height at the cavity rim (15AU) are able to reproduce the apparent Northward offset of the scattered light cavity from the star, verifying that it is likely a result of viewing geometry and is not a true offset, as described in detail in the text. Bottom: Simulated radial polarized intensity (blue) profiles for these models shown relative to the data at H (left) and Ks (right).

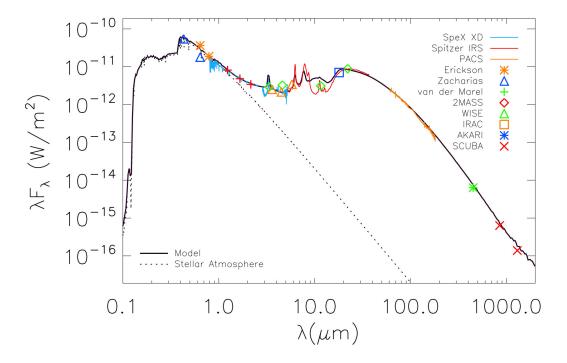


Figure 3.9: The best fitting SED model generated with the Whitney code, shown in black and following the parameters outlined in Table 3.2, overplotted on observational data from the literature. The sources for the literature data are described in detail in Table 3.1. The poor quality of the fit in the $10\mu m$ region is discussed in Appendix B.

Table 3.2: Relevant Whitney Model Input Parameters for Oph IRS 48

Parameter	Value	Reference
$\overline{\mathrm{M}_{star}}$	2.2M _☉	this work
T_{star}	9500K	this work
R_{star}	$1.8R_{\odot}$	this work
$\mathrm{M}_{disk,dust}$	1.6×10^{-5}	Bruderer et al. (2014)
R_{max}	160AU	Bruderer et al. (2014), this work
R_{min}	1AU	this work
R_{gap}	60AU	this work
i	50°	van der Marel et al. (2013)
β	1.3	this work
δ_{gap}	0.004	this work
H_{60AU}	10AU	this work

ure 3.9. This best fit SED is an excellent match to the literature photometry and spectra except at the 10μ m silicate feature. Attempts to fit this portion of the SED, and their incompatibility with scattered light images generated by the models, is described in detail in Appendix B.

Given the wide-ranging degeneracies of SED modeling, input parameters were fixed to observed values wherever possible, and only parameters having a marked effect on the quality of the SED fit were investigated in detail. Relevant parameters are listed in Table 3.2.

Whereas previous modelers have assumed a lower intrinsic stellar luminosity, and this has required an additional inner dust disk component (e.g. Bruderer et al. 2014) or halo (e.g. Maaskant et al. 2013) to reproduce the NIR excess seen in the SED, our choice of the McClure (2009) extinction law increases the intrinsic luminosity of the star and the dereddened stellar photosphere tracks the observed photometry closely out through the NIR. This contrast in the fit to the NIR photometry according to reddening assumption can be seen clearly in Figure 3.9. We adopt a simple uniformly depleted dust gap that extends inward to the sublimation radius in our dust modeling. The surface density profile of our model disk therefore closely follows that used by Andrews et al. (2011), and in particular their Figure 2a.

In all models, we assumed a simple disk structure with two components (large and small grains, with large grains restricted to 20% of the scale height of the small grains to mimic settling of large grains toward the midplane) and a disk rim at 60AU, inside of which the disk has the same general properties, but is depleted by a factor of δ_{gap} . In our best fit model, δ_{gap} =0.004, which is within a factor of a few of $\delta_{gap,dust}$ =0.0009 inferred by Bruderer et al. (2014), although their dust structure is somewhat more complex, being wholly depleted from 1-60AU

with an inner <1AU dust disk depleted by this factor.

The McClure et al. (2010) 6-37 μ m Spitzer IRS spectrum, which traces disk emission through the PAH bands in the NIR-MIR and across the thermal emission bump due to the outer disk rim places significant constraints on dust properties. As noted in Geers et al. (2007), the PAH features in this disk are very strong, pointing to the presence of gas inside of the cavity, thus it cannot be wholly depleted of material. Additionally, Geers et al. (2007) were able to isolate H α emission in this object, providing evidence of ongoing accretion and further lowering the feasibility of a fully depleted central cavity. Our inability to fully remove the stellar residual through halo correction is further anecdotal evidence in favor of material inside the cavity.

Treating the dust mass of $1.6\times10^{-5} M_{\odot}$ derived in Bruderer et al. (2014) from an observed gas mass of $1.5\times10^{-4} M_{\odot}$ (and an assumed 10:1 gas:dust mass ratio) as a fixed parameter, we find that among the large grain dust prescriptions that are standard in the Whitney model, a steep power law grain size distribution (p=3.5) and large maximum grain size (1mm) is best able to reproduce both the Fedele et al. (2013) 55-210 μ m Herschel PACS spectrum and the Andrews & Williams (2007) 850 μ m and 1.3mm SCUBA photometry. A similar model with a smaller maximum grain size (20 μ m) drastically overproduces sub-mm flux unless the dust mass is set to a factor of 3 lower than the Bruderer et al. (2014) value, and this model is also unable to reproduce the correct PACS slope. It is worth noting, however, that the assumption that grain sizes and compositions are uniform throughout the disk is certainly an oversimplification, as discussed in detail in Menu et al. (2014).

With these dust prescriptions and a dust mass of $1.6\times10^{-5}M_{\odot}$, we find that we are best able to match the SED by placing approximately 85% of the dust

mass in large grains, and the remaining 15% in small ISM-like sub-micron grains. More specifically, we find that this 85/15% division of dust mass into large and small grains respectively is best able to reproduce the size and shape of the MIR spectral bump as well as the sub-mm photometry. This is similar to values found by other modelers for both this and other transitional disks (e.g. Andrews et al. 2011).

3.4.3 Modeling of Scattered Light Images

Having narrowed the available parameter space significantly through SED modeling, we also attempted scattered light modeling of the disk with the Whitney code. We reran the Whitney code for our best-fitting SED model with a greater number of photons, as needed for full scattered light image simulations. We iterated on the output until we got a reasonable match to the geometry and brightness distribution of our observed images while maintaining the quality of the SED fit. Although this modeling effort was by no means exhaustive, it does inform some of the likely disk properties.

The presence of a dark lane to the North in the model images shown in Figure 3.10 suggests that grains inside of the disk cavity cause some shadowing of the inner cavity wall to the North. Because of the geometry of the disk's inclination, the dark lane that presumably exists along the Southern rim is not visible. This dark lane is apparent only in the raw model images, and is not evident once they are convolved with the PSF (bottom panels). We believe, therefore, that though a dark lane is not visible in our data, shadowing by the non-negligible amount of material within the cavity contributes to the faintness of the real disk images to the North. Higher resolution imagery of Oph IRS 48 in the future may serve to verify this conclusion.

Radial profiles were also taken through these model images and are shown in

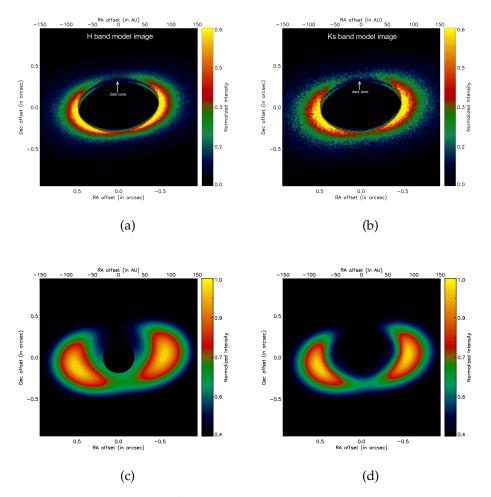


Figure 3.10: Simulated images of the Oph IRS 48 disk created with the Whitney code at H band (left) and K band (right). In the bottom panel, these images have been convolved with the observed PSF at each wavelength.

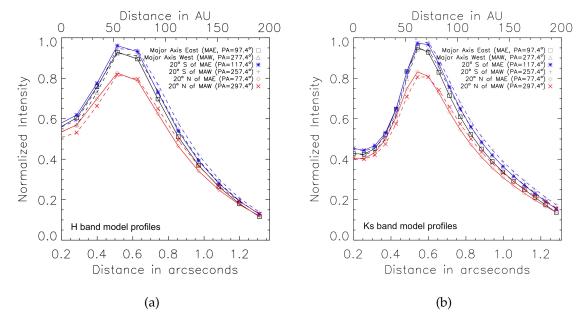


Figure 3.11: Radial profiles taken through the deprojected H (left) and K (right) model images, which are informative in contrast to the true observed radial profiles shown in Figure 3.5a. The model profiles are drastically different from the observations in their East/West symmetry. They do, however provide a reasonable reproduction of the relative shape, location and relative intensity of the peaks outside of the spiral arm.

Figure 3.11 and described in detail in Section 3.4.4.

The primary way in which modeled disk images appear to fall short of the reality of Oph IRS 48 is in their assumption of azimuthal symmetry. In particular, the Southwestern and Western parts of the disk, which host the ALMA excess and spiral arm respectively are significantly less uniform in brightness and smooth in morphology than both the disk models and the Eastern rim would suggest. Non-axisymmetric modeling is beyond the scope of this work.

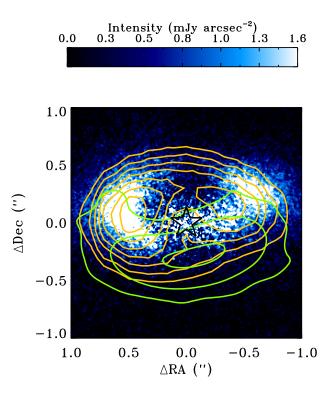


Figure 3.12: Contour map of Geers et al. (2007) $18\mu m$ VLT VISIR data in yellow and van der Marel et al. (2013) 0.44mm continuum ALMA data in green overplotted on our Hi-CIAO median halo-subtracted H band PI image. Note that the $18\mu m$ thermal emission is centered on the star, and that its Eastern lobe is cospatial with ours, whereas the $18\mu m$ Western lobe lies inside of our scattered light lobe. As discussed in the text, this is evidence that the Western arc is fainter simply because it is farther from the central star

3.4.4 Brightness Asymmetries in Real and Modeled Data

The disk of Oph IRS 48 as observed in the NIR contains a number of marked asymmetries. The most immediately apparent is the East/West brightness asymmetry, namely that the Eastern arc appears significantly brighter than the Western at both wavelengths. This is likely a simple result of proximity to the star, as the Western arc is as much as 0′2 more distant from the central star, lying along a non-axisymmetric structure in the West of the disk that we discuss in detail in Section 3.4.6.

This scenario would also explain the offset in the apparent position of the Western arc in our images relative to the $18.7\mu m$ contours of Geers et al. (2007), shown overlain on our PI image in Figure 3.12. For a thermal mechanism, emission should be symmetric about the star, so we have assumed that the star is located at the center of the two lobes of $18\mu m$ emission. While this assumption places the Eastern lobe of the $18\mu m$ emission near the observed scattered light lobe, it places the Western H-band scattered light lobe almost entirely outside of its $18\mu m$ counterpart.

The scattered light models described in Section 3.4.3 are able to reproduce the general shape and relative peak brightnesses of the Eastern disk profiles, although they deviate significantly in the West due to the model's incorrect assumption of azimuthal symmetry. We engage in more thorough discussion of the properties of the models in Sections 3.4.5 and 3.4.6, however we elected to stop our full MCRT simulations at this point, as we have obtained a reasonable facsimile of the general properties of the true disk and full MCRT simulations are computationally expensive.

3.4.5 Offset of Cavity Center from Stellar Residual

As noted in Section 3.3.2, the center of the best fit ellipse to the Oph IRS 48 cavity in both data sets is offset from the stellar location. In the East/West direction, the offsets between the stellar residual and the cavity center are small, and are not uniform between the two datasets. In the Ks band data, the best fit cavity is offset by 5 pixels (0'05) to the East of the star, and in H band it is offset by 6 pixels (0'06) to the West. In both cases, this may be due to the influence of the spiral arm component on the cavity fit. In the H-band data, the asymmetry is not resolved and is therefore not excluded from the cavity fit, though it is clear from the overlay in Figure 3.4 that the Western H-band arc lies mainly along the asymmetry, so this is likely to pull the best fit ellipse center to the West. In the case of the Ks-band data, the exclusion of the region of the asymmetry from the fits and the fact that the Eastern profiles are brighter means that the Eastern arc dominates the fit, and may pull the best fit ellipse center to the East. Given the small magnitude of these offsets and the fact that they are in opposite directions in the two data sets, we do not consider either one to be a true reflection of disk properties.

The large Northward offset of the best fit cavity center from the stellar residual at both wavelengths (15 pixels/0'.14 at H-band and 18 pixels/0'.17 at Ks band) is much larger and requires explanation. Our initial scattered light modeling of the disk, described in detail in Section 3.4.1, was able to reproduce a Northward offset of equal magnitude through simple disk geometry, namely a large scale height at the disk rim ($Z\sim15AU$). Since the Southern disk rim is shielded from our view, this tends to push the apparent Cavity center northward relative to our line of sight.

If we take our best fit model images from the full MCRC modeling described

in Section 3.4.3 and use the same ellipse fitting procedure that we used in Section 3.3.2, then we get similar, though not identical, values for the best fit ellipse parameters. We derive a slightly larger (a=0''58, b=''42) cavity than using the true K-band data (a=0''49, b=0''37), though it is of similar eccentricity (0.69 vs. 0.66) and PA (96.3 vs. 97.2°). We also reproduce a northward offset of the cavity center of between 6 and 9 pixels, depending on the exact parameters for the radial profile fits. This is about half of the observed offset. Our NIR modeling with the Sprout code suggests that simply increasing the scale height of the gap wall further should increase this offset to the observed value, however doing this with the Whitney model resulted in a larger MIR bump at $\sim 20\mu$ m than is observed in this object. We believe both the larger cavity radius and smaller scale height implied by the model are spurious, and are likely a result of the chosen grain prescription and disk geometry, however we leave a full exploration of the effect of these parameters on model images for future efforts.

3.4.6 On the Nature of the Western Extension

The geometry of what we've called the Western arc of the Oph IRS48 disk, and in particular the nearly right angle at which it meets the Southern Ks-band arc, suggest that, at the very least, the disk of Oph IRS48 is not axisymmetric. Rather than tracing the cavity all the way around, emission seems to be concentrated at a significantly greater stellocentric distance along -100<PA<-70, explaining both its relative faintness in this region and the deviation of the location of peak brightness along these PAs from the cavity fit.

One explanation for this deviation is that the emission in this region traces a spiral arm that extends to the Southwest from the Western cavity rim. ALMA gas data reveal that the Western half of the Oph IRS 48 disk is rotating counterclockwise away from Earth (van der Marel et al. 2014). This means that the arm trails

the disk rotation, as is observed in other disk spirals (Grady et al. 2013; Muto et al. 2012). By fitting the radial profile peaks in this region with a line in the deprojected cavity image and comparing the slope of this line to a circle at similar radius, we estimate the pitch of the arm to be $\sim 40^{\circ}$. This value is quite high compared to values seen in other disk spirals (e.g 15° Muto et al. 2012), indicating either a region of high temperature in the disk (H/R \sim 1, consistent perhaps with the large scale height necessary to explain the northward offset of the cavity center from the stellar residual in our modeling) or that the portion of the arm that we are seeing is very close to the planet's corotation radius.

Alternatively, this feature could be a local "divot" in the disk rim at PA~270. Localized surface brightness deficits have been observed in other disks (e.g. Mayama et al. 2012). In terms of a physical explanation for such phenomena, the simulations of Jang-Condell (2009) suggest that a massive planet embedded in the disk could mimic this feature by creating a "shadow" at the location of the planet (due to a rapid change in the disk scale height in this region), and a bright exterior region where the disk surface emerges from shadow.

Regardless of the physical cause of this feature, H and K band profiles along the Western asymmetry, when compared to profiles through the rest of the disk, reveal that grain properties also deviate in this region. A close examination of the deprojected radial profiles shown in Figure 3.5a, reveals that the brightness of scattered light in this region relative to the rest of the disk varies according to wavelength. In particular, the profiles at 20° North of the disk major axis to the East and West (red lines) deviate between the H and K band data. When normalized to the Eastern profile peak, the Western profile is 25% brighter in the H-band data (~80-85%) than in the Ks-band data (~55-60%) even before the latter is smoothed to simulate similar AO correction, which reduces its brightness

relative to the Eastern peak by an additional \sim 5-10%.

Although the same numerical comparison can't be made along the major axis (black lines), since the cavity rim is not clear in the Western profile, the trend appears to hold here as well, and the Western profile remains bright out to a larger radius in the H-band data than in the Ks-band. This is not the case, however, for all other profiles, including all of the Southern (blue lines) and Eastern (solid lines) profiles, which are well matched between the smoothed Ks-band data and the H-band data.

This effect is not present in the model profiles shown in Figure 3.11, which hints at a physical difference in the grains in this region relative to the rest of the disk. Specifically, the grains in the region of the asymmetry appear to have bluer scattering properties, which Mie theory would suggest is an indication that particles here are smaller than elsewhere in the disk. Dynamical perturbations may explain this deviation, as small particles should be well-coupled to the gas in the disk, while large particles have larger stopping times and may lag behind any gas perturbations (Lyra et al. 2009).

Although purely speculative, we note that this feature may somehow be associated with the dust enhancement observed by van der Marel et al. (2013) and shown overplotted on our data in Figure 3.12, which is located at a similar radius, though it lags behind the observed non-axisymmetric scattered light region by at least 30°.

Generally speaking, azimuthal asymmetries such as those observed in Oph IRS 48 are difficult to create through many of the alternative disk clearing mechanisms (photoevaporation, grain growth), and are most consistent with the existence of a perturber (planet or brown dwarf) in the disk. The gravitational influence of this body serves to both excite the observed asymmetries and clear the

central cavity, perhaps with the assistance of other mechanisms. Some disk dynamical processes can create asymmetries, however these are either associated with gap opening (e.g. by a planet, Rossby-wave instability), require a lower gasto-dust ratio than is observed in Oph IRS48 (photoelectric instability) are smaller in spatial scale than observed here (MRI-induced asymmetry) or are associated with immediate planet formation (gravitational instability). Together with the Bruderer et al. (2014) derivation of a two-step decrease in gas surface density inside of the cavity and the van der Marel et al. (2013) observation of a pronounced sub-mm asymmetry in this disk, our data lend additional credence to the assumption that this disk hosts one or more planets.

3.5 Conclusion

We have spatially resolved the circumstellar disk of Oph IRS 48 for the first time at NIR wavelengths. New H and Ks-band scattered light imagery reveal a cleared central cavity with a rim at \sim 60 AU. Fits to the shape of the cavity are consistent with a circular geometry, but with a cavity that is offset to the North relative to the star due to viewing geometry.

The disk hosts a number of interesting asymmetries. The East is brighter than the West, explicable via the fact that the Western disk arc appears to be farther from the star than the Eastern arc. We interpret this feature as either (a) tracing a spiral arm with a pitch of $\sim\!40^\circ$ or (b) a local surface brightness defect at PA=270° potentially caused by a "planet shadow".

The disk is a factor of \sim 2 brighter in absolute intensity at Ks-band than at H-band, however this is likely a simple result of the highly extincted region in which Oph IRS 48 lies. The scattering properties in the region of the spiral arm, however, are markedly blue relative to the rest of the disk, suggesting that grains

in this region are smaller than elsewhere in the disk.

Scattered light modeling using compact grains is able to reproduce a reasonable facsimile of the true disk emission with two important caveats. First, the SED in the region of the 10μ m silicate feature proves to be very difficult to match with most standard grain models. Secondly, the models are not equipped to reproduce azimuthal asymmetries such as are clearly present in the Western half of the Oph IRS 48 disk, and are only able to reproduce the relative shapes and brighnesses of radial brightness profiles along the Eastern disk arc. Neither of these problems is unique to this particular disk, and both suggest future directions for circumstellar disk models.

Taken as a whole, these data on Oph IRS 48 serve to emphasize several key points that are true of high contrast disk imaging in general. First, the extinction assumption that is used to deredden photometry for modeling can have profound effects on conclusions about stellar properties. In our case, we find that our choice of reddening law drastically affects the luminosity and mass derived for the central star. The geometry of this disk at NIR and sub-mm wavelengths also highlights the complications of observed asymmetry in disks. In Oph IRS 48, an azimuthally symmetric disk cannot reproduce the observed data.

Our data are consistent with the existence of at least one planet in the disk, which could function both to clear the central cavity and to incite the observed deviation from axisymmetry. Oph IRS 48 is thus a good candidate for high-contrast adaptive optics imaging in the future.

CHAPTER 4

THE FIRST CIRCUMSTELLAR DISK IMAGED IN SILHOUETTE WITH ADAPTIVE
OPTICS: MAGAO IMAGING OF ORION 218-354

We present high resolution adaptive optics (AO) corrected images of the silhouette disk Orion 218-354 taken with Magellan AO (MagAO) and its visible light camera, VisAO, in simultaneous differential imaging (SDI) mode at H α . This is the first image of a circumstellar disk seen in silhouette with adaptive optics and is among the first visible light adaptive optics results in the literature. We derive the disk extent, geometry, intensity and extinction profiles and find, in contrast with previous work, that the disk is likely optically-thin at $H\alpha$. Our data provide an estimate of the column density in primitive, ISM-like grains as a function of radius in the disk. We estimate that only $\sim 10\%$ of the total sub-mm derived disk mass lies in primitive, unprocessed grains. We use our data, Monte Carlo radiative transfer modeling and previous results from the literature to make the first self-consistent multiwavelength model of Orion 218-354. We find that we are able to reproduce the $1-1000\mu m$ SED with a $\sim 2-540$ AU disk of the size, geometry, small vs. large grain proportion and radial mass profile indicated by our data. This inner radius is a factor of \sim 15 larger than the sublimation radius of the disk, suggesting that it is likely cleared in the very interior.

4.1 Introduction

Silhouette disks were first discovered in 1994 by O'dell & Wen (1994). A single dark silhouette, Orion 183-405, was seen against the bright $H\alpha$ emission of the Orion Nebula in their HST Wide Field Camera (WFC) images of the region. O'dell & Wong (1996) followed this with a Wide Field Planetary Camera 2 (WFPC2) HST

survey of the brightest regions of the nebula and discovered six additional disks. To this day, images of the Orion silhouette disks provide some of the most conclusive evidence for the existence and sizes of dusty circumstellar disks around young stars.

McCaughrean & O'dell (1996) conducted the first detailed analysis and modeling of silhouette disks. The primary conclusion of their modeling of the silhouettes identified in O'dell & Wong (1996) was that all were best-fit by optically-thick opaque disk models with exponential edges. As column density cannot be determined from an optically-thick disk profile, they were only able to place lower limits on the amount of material in these disks. For the disk that is the focus of this work, Orion 218-354, their best-fit models suggests an r=0'54 disk with an inclination of 60° and a total mass of $2.4 \times 10^{-5} \mathrm{M}_{\odot}$.

Later surveys of the Orion region with WFPC2 (Bally et al. 2000) and ACS (Ricci et al. 2008) revealed many more of these disks, bringing the total to 28 known Orion silhouettes. Some have been followed up at other wavelengths, including the sub-mm (e.g. Mann & Williams 2010), thermal infrared (e.g. Hayward & McCaughrean 1997) and x-ray (e.g. Kastner et al. 2005), however except for the very largest silhouette disk (Orion 114-426, r>1"), they have not been imaged from the ground in silhouette until now.

For many years following the pioneering HST observations, ground-based imaging of these disks in silhouette were precluded by their small size ($r \le 1$ "), requiring higher resolutions than were available with seeing limited images. Adaptive optics has long been capable of delivering such resolutions in the infrared, however the bright nebular emission lines where these disks appear in silhouette all lie in the optical regime, blueward of the operating wavelengths of most AO systems. Is is only with the high actuator pitch of modern adaptive secondary

mirrors that correction on the necessary spatial scales for observations at optical wavelengths has been achieved.

The central stars of the silhouette disks are also relatively faint ($R \ge 11$), putting them outside of the working range of most AO wavefront sensors. Pyramid wavefront sensors, however, allow for binning and can achieve correction on fainter stars. MagAO is among the first modern AO systems with the ability to achieve the necessary resolutions for such imaging from the ground, among the first systems with the ability to lock on sufficiently faint natural guide stars, and the first large telescope with an AO-optimized visible light camera capable of imaging at the necessary wavelengths to see disks in silhouette (Kopon et al. 2009, 2012; Close et al. 2012; Males et al. 2012; Follette et al. 2010).

4.2 Observations and Data Reduction

Observations of Orion 218-354 were conducted on December 6, 2012 as part of the commissioning of Magellan's Adaptive Secondary AO System (MagAO). MagAO is a natural guide star (NGS) facility instrument of the 6.5m Magellan Clay Telescope at Las Campanas Observatory. The commissioning performance of the system is detailed in Close et al. 2013, (ApJ, accepted).

The complete data set consists of 72 30 second images of Orion 218-354, however only the best 38 images were used, for a total integration time of 19 minutes. The 1024x1024 pixel CCD47 VisAO camera was used in simultaneous differential imaging (SDI) mode in which a Wollaston prism separates the VisAO beam into two beams of approximately equal brightness (for unpolarized sources). These beams are each passed through a separate narrowband filter, one centered on the spectral line of interest (H α , [OI] or [SII]) and one on the neighboring continuum. For Orion 218-354, the H α SDI filter set was used, with one filter of width 4.6nm

centered on H α at 656.6nm and one of width 6.1nm centered on the continuum at 642.8nm (hereafter "continuum").

Two different image rotator orientations were used, and images were obtained at three different chip locations for each rotator angle in order to distinguish filter artifacts from background structure in the nebula. To create the final images, the raw images for both channels were dark subtracted, flat fielded, rotated to a common orientation, and registered

The median combinations of the $H\alpha$ and continuum channels are shown in Figure 1 in the top left and bottom left panels. Background structure in the nebula is apparent at $H\alpha$, and is lacking in the continuum channel, as expected. A secondary star lies 2" to the SE of Orion 218-354 in both, and the dark silhouette of the disk is visible in $H\alpha$ even before PSF subtraction.

Due to the guide star brightness of R=12.5, requiring wavefront sensor binning to 2x2, even though significant improvement in image FWHM was obtained (FWHM \sim .1" vs. 0.7-1.2" seeing), Strehl ratio was still very low. The PSF therefore consists of a single profile, rather than the typical core+halo profile seen with higher Strehls. Despite the handicaps of short wavelength and dim guide star, the PSF of the central star is exceptionally well measured by the simultaneous acquisition of the continuum channel, and can be robustly removed.

PSF subtraction in MagAO's SDI mode is simpler than "traditional" PSF subtraction because the continuum channel is a simultaneous probe of the PSF at a nearly identical wavelength, and PSF structure is therefore identical. The only non-common optics between the channels are their respective SDI filters, and filter artifacts are easily removed by taking the median over several rotation angles. Because most stellar spectra have real structure (absorption) at $H\alpha$ that varies with spectral type, there is no absolute scaling between the filters for subtraction.

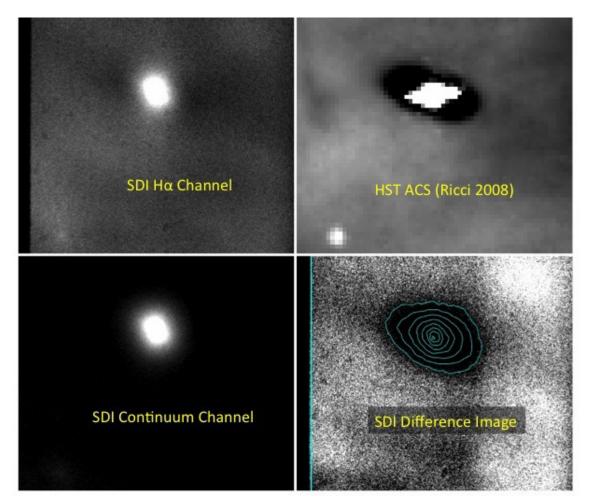


Figure 4.1: Upper left: MagAO H α channel. Lower left: MagAO 643nm continuum channel. Lower right: MagAO difference image created by subtracting the scaled continuum (lower left) from H α (upper left). Upper right: HST ACS image of the same disk (Ricci et al. 2008). All images are North up, East left and have the same physical scale. Note that the disk is visible in all panels except the continuum image, where it shouldn't be visible. Subtraction of the simultaneous PSF provided by the continuum channel effectively removes both the primary star and a secondary star 2" to the SE, isolating H α emission from the background nebula. The disk stands out starkly in silhouette in this (lower right) image, and light blue contours reveal that the background flux from the nebula is attenuated to an increasing degree as column density increases towards the center of this disk. The extent of the contours all the way to the central star is suggestive of an optically-thin disk at H α .

Instead, it is determined on a case-by-case basis from the ratio of the peaks in both the $H\alpha$ and continuum images.

The dominant source of error in the scaling of the PSF for subtraction comes from the 45° tertiary mirror that feeds MagAO, which is made of freshly coated (2012) aluminum with a reflectivity of \sim 97%. Because the \sim 3% of light absorbed by the coating may have a preferential polarization, and the orientation of the Wollaston relative to the tertiary changes as the instrument rotates, the scaling between the channels may change by as much as 3% over the course of observations taken with the rotator on.

In the case of Orion 218-354, the continuum image was scaled by a factor of 1.03 before subtraction from the H α image based on the ratio of the peaks. Both Orion 218-354 and the secondary star in the images are late-type stars (Hillenbrand 1997; Terada et al. 2012) with very little spectral structure in the H α region, which allowed us to use aperture photometry of the secondary star to verify this scaling. We found an identical 1.03 scaling based on this methodology.

The resulting PSF subtracted image is shown in the bottom righthand panel of Figure 1. The error in this subtraction is likely much less than 3% because of the short duration of the observations, however errors derived from a full 3% different scaling of the PSF are shown throughout this letter, and are shown in Figure 2.

The upper righthand panel of Figure 1 shows the HST ACS image of Orion 218-354 from Ricci et al. (2008) taken in the F658N (H α) filter. The scale of the image is the same as the MagAO images to allow for direct comparison of spatial structure. The secondary star and the background structures in the nebula appear in both the MagAO and HST H α images. The primary difference is that the central star in the HST image is heavily saturated, making it difficult to recover

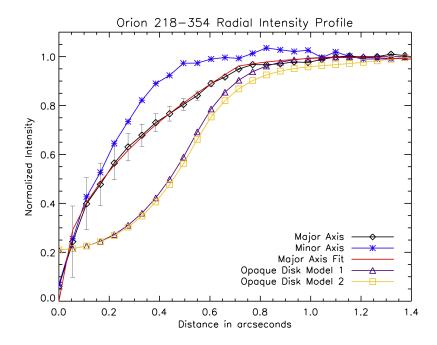


Figure 4.2: Radial intensity profile of the Orion 218-354 disk normalized to the background nebular intensity level. Profiles are shown for the major (black diamonds) and minor (blue asterisks) axes. An opaque optically-thick disk model with the best-fit parameters derived in McCaughrean & O'dell (1996) convolved with both 150x150 and 250x250 MagAO continuum PSFs are shown overplotted (purple triangles, yellow squares). In order to demonstrate that the character of the profiles is the same, the 250x250 pixel model has been normalized to the 150x150 model to remove the pedestal effect caused by "filling in" the inner disk with nebular light. The inconsistency in shape and character of these models with our data is striking, and the steadily decreasing profile we observe suggests an optically-thin disk. The performance of a modern AO system, as well as the simpler methodology of SDI-mode PSF subtraction, gives us a clear advantage over the HST data, in which the central star is heavily saturated. The best-fit to the MagAO profile, a powerlaw with exponent 0.46 and an exponential cutoff beyond ~0.75", is shown overplotted in red. Errors due to a 3% deviant scaling of the continuum PSF for subtraction are also shown.

information about the inner disk

4.3 Results and Analysis

Light blue isophotal contours are overplotted on the continuum-subtracted H α image in Figure 1. They reveal that the brightness of Orion 218-354 at H α falls steadily towards the center of the disk, calling into question it's classification as optically-thick. The radial profile along the major and minor axes is shown in Figure 2. The shape of these profiles remains the same for a range of scaled PSF subtractions. The profiles were binned to 7 pixels (\sim 0'055, half of the measured FWHM of the continuum PSF). The diffraction limit of the MagAO system at H α is \sim 20mas, and the system routinely achieves 6 and 7 pixel FWHM at this wavelength. This drop in performance was due to (a) relatively poor seeing for the site, ranging from 0'7 at the start of the observations to 1'2 at the finish, and (b) the faintness of the guide star, requiring binning of the pyramid pupils to 2x2.

Also overplotted on the radial profile in Figure 2 is the optically-thick best-fit disk model of McCaughrean & O'dell (1996) (a 0'54 60° inclined opaque disk seen against the nebular background) convolved with the Magellan continuum PSF. Several sizes of PSF were used for convolution to investigate the effect of the PSF wings. The PSF reaches the nebular background level at $r\sim75$ pixels, therefore we feel that the 150x150 pixel PSF convolution shown in purple in Figure 2 is the most robust. Enlarging the PSF has a pedestal effect on the convolved profile, effectively "filling in" the central part of the disk with nebular background light, however the shape and character of the profile remain nearly identical, as shown by the yellow 250x250pixel PSF convolution profile (normalized to match the 150x150 profile at the inner and outer edge of the disk). Both convolved model profiles are immediately apparent as entirely different in shape and character

from the observed radial intensity profile.

The MagAO profile is well fit by a powerlaw with exponent 0.46 and an exponential cutoff beyond \sim 0′/75 (\sim 300AU), shown overplotted in red. The flux in the MagAO radial profile reaches the background level at r=1″3, which corresponds to a disk radius of \sim 540 \pm 60 AU at 414pc (Menten et al. 2007), significantly larger than the McCaughrean & O'dell (1996) value of 0″54. Our best-fit to the outer disk suggest an ellipticity of 0.72 \pm 0.08, corresponding to a disk inclination of 44 \pm 5° (slightly smaller than the McCaughrean & O'dell (1996) value of 60°), and a major axis PA=72 \pm 10° East of North.

Intensity in the difference image was translated to extinction by:

$$A_{H\alpha} = -2.5 \log(I/I_B)$$

where I is the intensity in each pixel and I_B is the background intensity in the nebula. We converted this extinction value at $H\alpha$ ($A_{H\alpha}$) to an extinction value at V (A_V) by linear interpolation of extinctions at bracketing wavelengths per Mathis (1990), resulting in a factor of 1.22 increase in the extinction value from $H\alpha$ to V band.

We used the empirical relationship between A_V and hydrogen column density (N_H) derived by Bohlin et al. (1978) $(N_H=1.87\times10^{21}\times A_V\ cm^{-2})$ to convert A_V to N_H . We multiplied this value for N_H by the physical size of a MagAO pixel at 414pc, and by the mass of a hydrogen atom to get a disk dust mass estimate of $2.3\pm1\times10^{-5}M_{\odot}$, or approximately $7.5M_{earth}$. Errors were estimated from an equivalent conversion of the difference image with a 3% different PSF scale factor.

It is important to note that the A_V to N_H conversion value that we've employed is appropriate only for an extinction curve slope R_V =3.1, which corresponds to the diffuse interstellar medium. Larger grains have poorer reddening efficiency and are therefore "missing" in this approximation. In other words, our

mass estimate probes only primitive ISM-like grains in the disk.

Assuming a 100:1 gas:dust mass ratio, we convert this value to a total disk mass of $\sim 2.3\pm 1\times 10^{-3} M_{\odot}$. As a probe of the total disk mass, including large grains, the sub-mm disk mass estimate of $0.0237 M_{\odot}$ (Mann & Williams 2010) is much more robust. The order of magnitude difference between our mass estimate and the sub-mm estimate suggests that just $\sim 10\%$ of the disk mass lies in primitive grains that absorb efficiently at $H\alpha$.

We believe this value to be robust for several reasons. First, any foreground $H\alpha$ emission should be relatively uniform across the disk, and will not contribute to the differential extinction measurement. Secondly, scattered light from grey grains in the disk should be virtually identical between the two SDI filters, and therefore should be removed by PSF subtraction. Finally, because the radial profile of our best PSF subtraction approaches zero in the interior, there is little room to achieve a higher integrated extinction.

This small grain dust mass, as well as the shape of the radial intensity profile shown in Figure 2, are suggestive of a disk that is optically-thin at $H\alpha$ for $r\gtrsim 25$ AU (our innermost resolution element), a surprising result given previous work on silhouette disks. This low optical depth may be due to any number of factors, including small grain blowout, blowout due to ionizing radiation from the Trapezium, or grain growth in the disk.

An extinction/column density profile derived following the same procedure described above is shown in Figure 3 in both linear and log (inset) space. The inner disk is well fit by a powerlaw with exponent -1.43. An exponential disk edge reproduces the deviation from this powerlaw at large radii, as shown by the overplotted best-fit intensity profile. The primary source of potential error in this and other measurements of the observed radial profile is that they have been

effectively convolved with the instrumental PSF and may deviate somewhat from the "true" disk profile. We leave investigation of this effect for future work.

This extinction profile is a probe of the integrated mass density profile of ISM grains in the disk. Total mass in a given column scales with both the density at the midplane and the scale height of the disk at that radius (M(r) $\propto Z \times \rho_{midplane} \propto r^{\beta} \times r^{-\alpha}$). Assuming a geometrically flat disk where scale height $Z \propto r^{1}$ (β =1), the extinction powerlaw gives us an estimate of the midplane density profile: $\rho_{midplane} \propto r^{-2.43\pm0.3}$ (α =2.43±0.3). This is well within the range of values commonly assumed for the midplane density distribution. If the disk is flared (β >1), this is instead a lower limit (α >2.43±0.3).

To support our conclusions about the geometry and mass distribution in the disk, we've used our derived parameters and a series of simple assumptions as inputs to the Whitney 3D Monte Carlo Radiative Transfer Code (Whitney et al. 2003b,a, Whitney et al. 2013, ApJS, accepted) to generate a 0.1- 1000μ m SED. Input parameters are given in Table 1, and the generated SED is shown in Figure 4.

Photometric points from the literature are overplotted on the SED. They show that emission in Orion 218-354 is at or only slightly in excess of photospheric emission for all wavelengths L' and shortward. Only the sub-mm photometry (and, of course, the observed silhouette) strongly suggest the presence of a circumstellar disk.

The SED also suggests the presence of a small inner gap in the disk, as models that extend inward to the sublimation radius drastically overproduce NIR flux. A model with $r_{disk,in}=r_{sub}$ is shown overplotted as a dashed line on Figure 4 to demonstrate this. In order to reproduce the L' photometry of Muench et al. (2002) and Terada & Tokunaga (2012), the disk cannot extend farther inward than \sim 15 r_{sub} , or about 2AU. Therefore, the innermost 2AU of the Orion 218-354 disk is

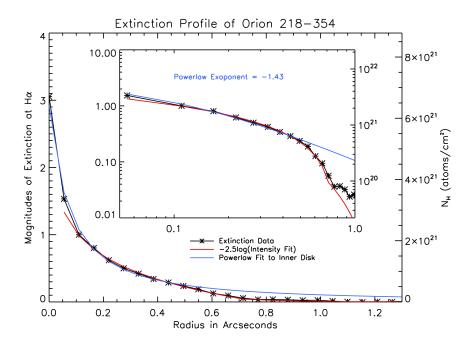


Figure 4.3: Radial extinction/column density profile of the Orion 218-354 disk in linear and log (inset) space. Overplotted on the data (black asterisks) in both plots are the Figure 2 intensity fit translated to units of extinction (red line) and a best-fit powerlaw to the extinction profile in the inner disk (blue line), which has an exponent of -1.43. This value is used to estimate the radial mass distribution of the disk ($\rho \propto r^{-\alpha}$), under assumptions described in detail in the text. This extinction should be considered to represent the column density of primitive, ISM-like material in the disk, and not of the disk as whole.

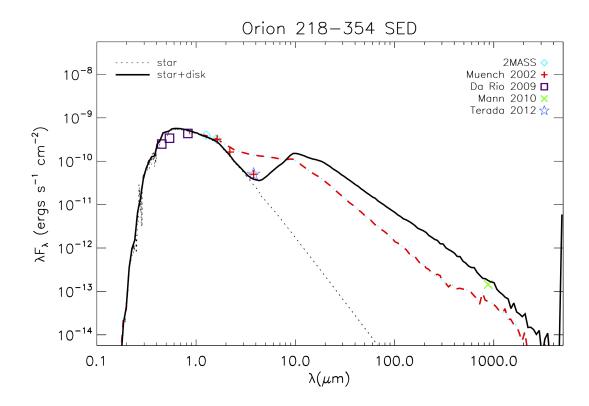


Figure 4.4: Multiwavelength photometry of Orion 218-354 (Muench et al. 2002; Da Rio et al. 2009; Mann & Williams 2010; Terada et al. 2012) overplotted on our best-fitting Whitney model output. The stellar spectrum (dashed black line) and best-fit model (solid black line, Table 1) are shown. The dashed red line corresponds to the same disk model with $r_{disk,in}=r_{sub}$. The poor fit suggests that Orion 218-354 contains an inner \sim 2AU clearing.

likely cleared of material.

4.4 Conclusion

In this study we have presented the first ground-based adaptive optics images of a circumstellar disk seen in silhouette. We derived geometric parameters of Orion 218-354 (r \sim 540AU, $i\sim$ 46°, PA \sim 72°) that suggest it is more extended and less inclined than previous observations would suggest. The radial intensity profile of our data is inconsistent with the opaque, optically-thick disk suggested by earlier modelers, as it shows a steady increase with radius in the amount of background nebular flux that passes through the disk. This suggests that the column density in small grains decreases steadily with radius. The fact that the amount of absorption does not plateau at small radii, as convolution of optically-thick disk models with our observed PSF suggest it would have, suggests that the disk is optically-thin at H α , at least as far inward as our innermost resolution element. Using the extinction/column density profile, we estimate the mass as a function of radius at the midplane in the disk assuming that it is geometrically flat, and find that it is proportional to $r^{-2.43}$.

The integrated column density of the disk translates to a mass of $\sim 2.3\pm 1\times 10^{-5} M_{\odot}$ in primitive ISM-like grains, and a total disk mass of $\sim 2.3\pm 1\times 10^{-3} M_{\odot}$ (assuming a 100:1 gas to dust mass ratio). The $880\mu m$ sub-mm continuum flux for this disk, on the other hand, suggests a disk mass that is higher by an order of magnitude $(0.0237 M_{\odot})$. This suggests that the majority of the mass in this disk (90%) may lie in grains that have grown beyond ISM-like sizes.

We model the $1-1000\mu m$ SED of Orion 218-354, showing that the sub-mm mass estimate and our derived parameters are consistent with published multiwavelength photometry of the disk. The NIR region of the SED suggests that the inner

disk contains a small gap, which we estimate at 15rsub or \sim 2AU.

This study demonstrates the power of modern adaptive secondary AO systems to achieve atmospheric correction into the visible wavelength regime, the ability of pyramid wavefront sensors to achieve excellent correction on faint guide stars, and the power of simultaneous differential imaging. Due to these complimentary technologies, the future of circumstellar disk imaging from the ground at visible wavelengths is bright.

Best-fit Model Input Parameters

Model Input(s) 1.9M _☉
1.9M _O
$2.4 \mathrm{R}_{\odot}$
5272K
1.51
$0.02373\mathrm{M}_{\odot}$
$15r_{sub}$
538AU
46°
0.9
7.5AU
2.43
1.0

Table 4.1: f_d represents the fraction of the disk's dust mass in large grains, Z_{100} is the scale height at 100 AU, α is the exponent for the radial midplane density, and β is the exponent for the radial scale height. The stellar and extinction parameters are from Hillenbrand (1997); the disk mass is from Mann & Williams (2010). All other parameters were derived from our observations as described in the text with the exception of Z_{100} , for which the model was found to be insensitive to a range of reasonable parameters (values 5-10AU), and $r_{disk,in}$, which was determined iteratively from the fit to existing J-L' photometry. The large grain dust prescription is Model 1 of Wood et al. 2002, a mixture of amorphous carbon and astronomical silicates with a maximum grain size of $1000\mu m$. The small grain dust prescription is the ISM grain model of Kim et al. 1994, a mixture of silicate and graphite with a maximum grain size of $0.28\mu m$.

CHAPTER 5

CONCLUSIONS AND FUTURE WORK

This chapter summarizes how the work highlighted in this dissertation addresses some of the open questions in planet formation theory, and gap clearing mechanisms in particular, followed by three projects I have initiated that are natural follow-ons to this work. These projects are not complete, and therefore are not included in the body of this dissertation, but they represent a large proportion of my graduate work and so are included here as highlights of works-in-progress. I'll finish by describing how and where I plan to continue this work as a postdoctoral fellow.

5.1 Summary

The theme that unites the work presented in this thesis is that of high-contrast adaptive optics imaging of circumstellar disks in general, and transitional disks in particular. Transition disks are new and complex enough that the majority of disk studies to date revolve around the characterization of single objects, and I've presented three such examples - SR21, Oph IRS 48 and Orion 218-354.

Even focusing on one object at a time, contrasting each observation with previous observations of similar disks reveals both patterns and oddities. Broader disk demographics is a very new field, but my work fits in with several patterns that are beginning to emerge.

The transitional disk SR21, which is described in detail in Chapter 2, was the first transitional disk whose sub-mm resolved cavity was demonstrably absent in NIR imagery. Other SEEDS results in the publication pipeline show that SR21 is not the only disk exhibiting this phenomenon. A satisfying theoretical inter-

pretation - that of dust filtration at planet-induced gap edges - has been put forward (Zhu et al. 2012). SR21, therefore, is one of a number of targets observed during the last five years where planet formation seems like the most feasible gap-clearing mechanism.

Oph IRS 48 is another object whose NIR imagery, as reported in Chapter 3, reveals asymmetries that are most consistent with a planet-induced formation mechanism. The blue scattering properties of the observed asymmetry provide an interesting theoretical puzzle for dynamicists, and the exact morphology of the asymmetry may be revealed by multiwavelength and higher resolution followup.

Orion 218-354, as reported on in Chapter 4, is in the very early stages of disk dispersal, yet it too informs the debate over disk clearing mechanisms. In particular, the estimated age of disks in the region makes the small proportion of $H\alpha$ absorption in Orion 218-354 puzzling. The simplest explanation for the optically thin disk in Orion 218-354 is that its grains have grown to large enough sizes that they are no longer efficient absorbers of $H\alpha$ photons. Grain growth is not the only explanation for this observation, but it is the most compelling one.

As even my own small body of work reveals, it's likely that all three mechanisms - planet clearing, photoevaporation and grain growth - contribute to gap formation, and each may dominate in certain types of systems. Observations such as those described in this thesis are important not just in providing constraints for theoretical planet formation models, but also in vetting targets for future high-contrast adaptive optics instruments. In particular, establishing planet formation as the likely gap clearing mechanism in objects such as SR21 and Oph IRS 48 makes them potentially profitable targets for future planet searches. If and when planets are found in these disks, the multiwavelength picture provided by previous disk imaging will be crucial in interpreting the age and nature of these

planets.

In the rest of this chapter, I'll describe ongoing projects with the goals of (a) illuminating the photoevaporation process through high-contrast imaging, (b) directly imaging forming protoplanets by isolating their H α accretion luminosity and (c) illuminating radial and azimuthal variation in grain properties, and their ice content, as a means of probing the mechanics of planet formation.

5.2 Resolving Photoevaporative Disk Surfaces

The unprecedented spatial resolution in the optical provided by VisAO and the ability to do differential imaging at the very short wavelength of the [OI] forbidden emission line should have opened up an exciting window into the process of photoevaporation in transitional disks, which has only been "resolvable" in a spectral sense to date. In my role as developer of science cases for the MagAO project, and as part of several proposals to build an optical fiber-fed integral field spectrograph (MagIFS) for MagAO (e.g. Follette et al. 2010), I completed a number of modeling and on-sky experiments to investigate the feasibility of resolving photoevaporating disk atmospheres with the system. The results of the modeling and on-sky experiments are inconsistent and puzzling, and both are described briefly below.

5.2.1 Modeling and Test Cases

Many groups have taken advantage of the presence of forbidden line emission in the spectra of Young Stellar Objects (YSOs) to isolate disk emission from stellar emission. As it cannot be generated in the relatively high-density region of the stellar photosphere, any forbidden line emission observed must originate from the lower density circumstellar region. Whether this emission comes from a stellar wind, a disk, or a jet, however, is often unclear and heavily disputed. Very

high-resolution spectra, combined with models of possible disk geometries, have helped to inform the debate. However, there are a multitude of unanswered questions, mostly owing to the relatively small number of disks with full geometric models and the degeneracy of explanations for features in the high-resolution spectra. The ability to spatially resolve forbidden line emission in YSOs is a unique capability of the MagAO system, and this variety of disk imaging was one of the driving science cases behind the development of the VisAO SDI mode. A wide range of key disk features can be mapped with 20mas images, including asymmetries, spiral arms, cavities and gaps, all of which are indicative of planet formation and/or clearing processes.

Some YSOs show broad symmetric forbidden line profiles centered on the stellar radial velocity, and these are generally interpreted as signatures of rotating Keplerian disks with extents of up to 100AU (e.g. van der Plas et al. 2008; Acke & van den Ancker 2006). However, some groups have put forward alternative explanations, including that of a spherically symmetric stellar wind (e.g. Boehm & Hirth 1997). Some YSOs show complex forbidden line profiles with multiple components, including in many cases a strongly blueshifted component. These have been interpreted as the signature of an outflow whose redshifted component is obscured by disk material (e.g. Corcoran & Ray 1997). Others show a low velocity blueshifted component interpreted as a stellar or disk wind (e.g. Rigliaco et al. 2013). Still other YSOs, and in particular those of later spectral type, show photospheric absorption features in the same regions of their spectra as the forbidden lines, which masks any other emission at these wavelengths (Acke et al. 2005). Their geometries remain a mystery. Spatially resolved images of YSOs in forbidden lines, such as those made possible by MagAO's SDI mode, will help to answer these morphological questions by directly distinguishing emission from disks, extended disk winds and jets (see Figure 5.2).

Because emission from the star in such systems is simple continuum, scaling and subtraction of the neighboring continuum filter with MagAO SDI imaging results in nearly perfect removal of residual starlight, as demonstrated with our $H\alpha$ imaging of the Orion 218-354 transition disk described in Chapter 4.

Of the many forbidden lines observed in YSO spectra, two of the strongest and most well studied are [OI] at 6300A and [SII] at 6716/6731A. Both can originate in jets through collisional ionization and shock photoionization (Shang et al. 2007), yet both may also originate from photodissociation of simple molecules in disk atmospheres (Ercolano & Owen 2010). Spatially resolved forbidden line emission should therefore trace the extent and geometry of jets where present, as well as the geometry of disks in cases where photodissociation is occurring. Since the MagAO system is the first system capable of resolving emission at these wavelengths, it may be the first to unequivocally isolate the origin of this emission. If a disk wind is present, SDI imaging should be able to isolate the region or regions from which it is launched, as well as its extent, informing the debate on whether FUV, EUV or x-ray photons are most responsible for launching the wind, as described in detail in Chapter 1.

Among the sample of NGS-accessible transition disks, at least nine have already been observed spectroscopically at [OI]. Of these, three show evidence of Keplerian rotation in their line profiles, three show strong evidence for a disk wind component, and three lie in systems where forbidden line emission is confused with photospheric absorption so nothing is known about system geometry (Acke et al. 2005, E. Rigliaco and I. Pascucci private communication). The number of NGS accessible disks that have been observed spectroscopically at [SII] 6717/6731 is comparatively small, and the only reported detection is in TW Hya

(Pascucci et al. 2011). Data recently acquired by others has also revealed that extended foridden line emission may originate from accreting brown dwarfs and protoplanets (A. Kraus, private communication), which would make observations at this wavelength complementary to $H\alpha$ imaging campaigns.

While the MagAO system was still being built, we developed a number of science cases to take advantage of the unique capabilities of the system, including imaging of disks at [OI]. I investigated the suitability of VisAO SDI to resolve conceivable disk morphologies by simulating images of a protoplanetary disk around a typical Herbig Ae/Be star. I placed the simulated disk at the 120 pc distance of the ρ -Oph star forming region and derived the total [OI] emission from observations of the Herbig Ae/Be star HD100546 conducted by van der Plas et al. (2008). I spread the total line emission from this YSO into three potential disk morphologies: a disk, a disk with a gap from 5-10AU and a jet. Figure 5.1 shows simulated [OI] images of these three cases (assuming a conservative 10% Strehl 0.63 μ m VisAO PSF). The stellar continuum (and PSF) is easily isolated and subtracted through the continuum filter image. This PSF is then used for deconvolution, and even with a forced 20% PSF mismatch. On-sky VisAO performance has shown this estimate to be very conservative both in the sense of absolute Strehl ratio and quality of PSF match.

Anticipating cases in which a young star might exhibit forbidden line emission from both a circumstellar disk and a jet (particularly at the [SII] 6716, 6731 Angstrom jet lines), we also created a hybrid disk/jet model with a variety of brightness scalings between the two components. Shown in Figure 5.2 is one such recovered spatial map of a simulated system in which the jet is 10 times brighter than the disk. Even with such contrast, the disk is still recoverable.

This preliminary study led us to conclude that photoevaporating disk atmo-

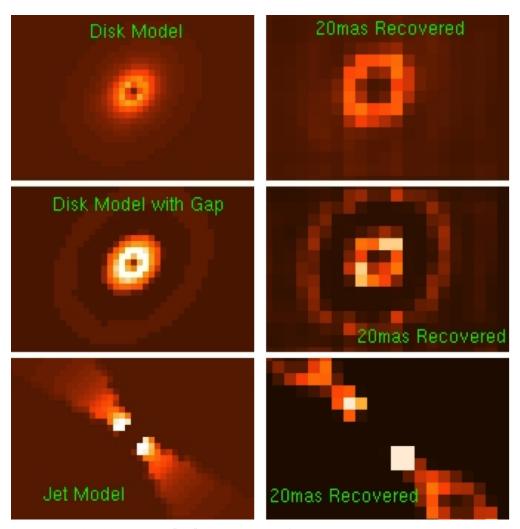


Figure 5.1: Simulated MagAO [OI] SDI images of HD100546 with three plausible origins: A continuous disk, A disk with a gap from 5-10AU and a jet. In all cases, despite a conservative Strehl PSF and an implausible PSF mismatch, the morphology of the system is well recovered.

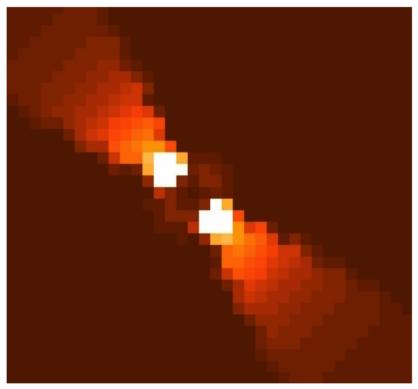


Figure 5.2: Model (top) and recovered (bottom) images of a disk+jet system, where the jet is an order of magnitude brighter than the disk.

spheres should be resolvable with MagAO if the current models are correct, and we chose to incorporate both [OI] and [SII] filters into the design of the VisAO SDI system.

5.2.2 Case Study: HD100546

We completed an on-sky experiment in this methodology in the Spring of 2013 when we obtained 90 minutes of closed loop imaging of the transition disk HD100546 in excellent seeing. We chose this target because it exhibits broad, symmetric [OI] line emission that peaks at >20% of the continuum (Acke et al. 2005), making the [OI]/Continuum emission ratio quite favorable for detection relative to other [OI] emitting disks.

Acke & van den Ancker (2006) deconstructed the HD100546 [OI] line profile into Keplerian rotational components and predicted the radial emission profile of [OI] in the disk. Compensating for the MagAO [OI] filter width, we predict that the [OI] line excess should amount to an additional 0.3% above the continuum according to their model. This is a small difference, however since [OI] emission from the disk atmosphere and stellar continuum emission don't follow the same spatial distribution, the contrast between continuum and [OI] should be more favorable in some places than in others. Figure 5.3 shows the predicted contrast between the Acke & van den Ancker (2006) [OI] disk emission model and a Gaussian approximation to an Airy pattern (chosen for its simplicity) of the functional form:

$$I = e^{\frac{-x^2}{2FWHM^2}} (5.1)$$

where FWHM is the observed (6.4 pixel) full width at half-maximum of the [OI] SDI Continuum channel. At separations under \sim 30 pixels, Continuum emission should dominate, but [OI]/Continuum contrast remains above 10^{-4} at all radii. Outside of \sim 30 pixels, [OI] emission should dominate, though the absolute

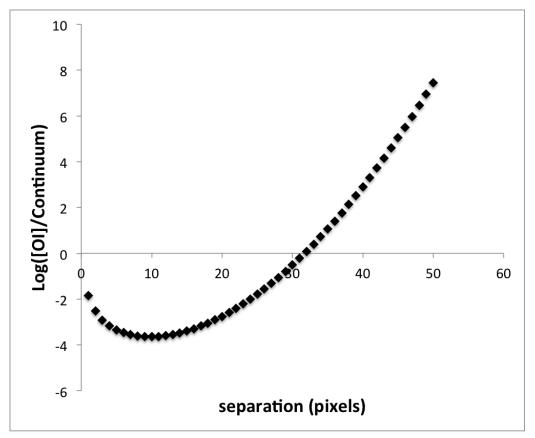


Figure 5.3: Predicted contrast between [OI] and continuum emission in HD100546 as a function of radius in the disk (plotted in VisAO pixels)

amount of emission at these radii will rapidly approach the noise level.

Figure 5.4 shows a preliminary continuum-subtracted image, which was taken at three different rotator positions to mitigate instrumental artifacts. Despite this, several survive, including filter ghosts and a strong residual from the ASM control radius, described in detail in Chapter 1. This image shows no obvious contribution from a photoevaporative disk atmosphere, and the noise level at a given radius is sufficiently small that we believe we should have seen something if the disk emission follows the Acke & van den Ancker (2006) prescription. A more rigorous modeling effort that includes inserting a simulated [OI] disk into the observed continuum images and attempting to extract it using the same post-

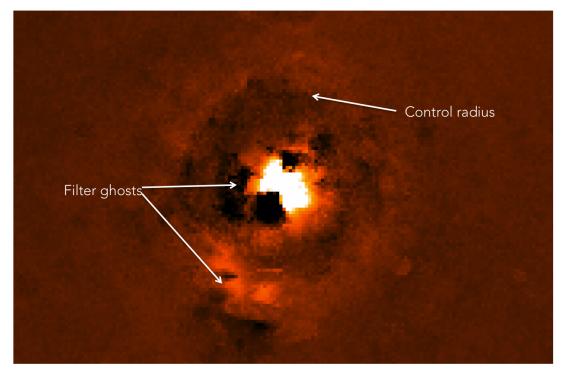


Figure 5.4: Continuum subtracted MagAO [OI] image of HD100546. Aside from various repeatable artifacts (labeled), no obvious [OI] emission is visible.

processing techniques as were used here is underway, and will help to answer whether the lack of [OI] emission in this image is because of the poor performance of the filters, an unreasonable assumption in the back-of-the-envelope calculation described above, or because the Acke & van den Ancker (2006) model of a continuously-emitting rotating Keplerian disk as the source of [OI] emission in HD100546 is incorrect.

5.2.3 Future Work

The ability to spatially resolve features such as gaps, jets and flaring would aid greatly in our understanding of disk structure and its connection to the process of planet formation. However, it would seem that the theoretical and physical realities described in the previous two subsections are in conflict, at least in the case of HD100546. Based on these results, it stands to reason that either (a) our

current understanding of the [OI] emission mechanism, (b) the inferences of [OI] disk modelers, and/or (c) our assumptions about the [OI] sensitivity of VisAO are faulty in some way. Perhaps [OI] emission in this system does not originate in the rotating Keplerian disk, but is broadened by some other turbulent mechanism. Alternatively, perhaps there are uncharacterized systematics that are affecting our ability to resolve [OI] emission with the MagAO system.

Either (a) repetition of the experiment on another disk for which evidence of photoevaporation has been put forward or (b) upgrading the VisAO [OI] filters (as has already been done with our new single substrate $H\alpha$ filters) to remove some of the filter artifacts that plague ASDI reductions is warranted in order to completely rule out the possibility of resolving [OI] emission with MagAO. A full reduction and interpretation of the HD100546 data that provides encouragement as to the feasibility of this endeavor may be necessary, however, in order to justify either test. The preliminary results of this study were the impetus for the depriotitization of the [OI] filter, which was removed from the filter wheel in favor of another coronagraph in Spring, 2014. Whether or not the filter is reinstalled, we realize that the results of this experiment are important in constraining photoevaporation models, and we will publish the data on HD100546 either as a separate result or as a part of the broader GAPlanetS initial results paper.

5.3 MagAO Giant Accreting Protoplanet Survey

I recently embarked on a survey of 15 nearby transition disks with the 6.5m Magellan Adaptive Optics system (MagAO) called the Giant Accreting Protoplanet Survey (GAPlanetS) that is the culmination of my efforts as a member of the MagAO team for the past six years. Although none of the results have yet been submitted for publication, I would be remiss not to include them in this dissertation

as they represent a significant portion of my graduate work.

GAPlanetS takes advantage of MagAO's revolutionary performance at visible wavelengths to image young stars at the short wavelength of the H α emission line (Follette et al. 2013a; Close et al. 2014). This takes advantage of both intrinsically higher resolution at visible wavelengths and the contrast advantage provided by actively accreting objects, which emit strongly at H α . Transitional disks are ideal targets for an H α direct imaging census because (a) the clearing of material in their inner regions means that relatively little intervening circumstellar material is present to attenuate emission from companions, and (b) many transitional disks are still actively accreting, making the probability for accretion streamers and H α -emitting objects high.

As described in Chapter 4, H α SDI mode has been extensively tested on sky, and several cutting-edge results have already emerged. Key aspects of our onsky demonstrations include (a) that even with faint (R \sim 12) guide stars, the performance of MagAO gives up to three orders of magnitude improvement over seeing-limited imaging in the stellar FWHM and (b) that SDI mode PSF subtraction offers significant improvement over traditional PSF subtraction, allowing us to probe much closer to the star (0'05) than was possible with other imagers, including the Hubble Space Telescope.

5.3.1 Case Study: HD142527

We have already demonstrated the power of this observing mode for isolating accreting objects inside of disk clearings in our published observations of the transition disk HD142527 (Close, Follette et al. 2014). Through these observations, we were able to verify the existence of a previously disputed companion (mass still unknown, though likely a low mass star or brown dwarf) embedded inside the disk gap at just 86 ± 3 mas (~13 AU). This companion was previously

inferred interferometricaly (Biller et al. 2012), but MagAO was able to directly image the object. The object-star contrast through MagAO's H α filter is nearly three times that in the continuum (δ_{mag} =6.33 \pm 0.2 at H α vs δ_{mag} =7.50 \pm 0.25 at continuum), demonstrating that accreting objects have more favorable planet/star contrast ratios at H α than at other wavelengths.

The scientific substance of these point source detections will be limited without (a) model predictions of H-alpha luminosities and equivalent widths for forming protoplanets (for which we have begun a collaboration with Zhaohuan Zhu at Princeton) and (b) multiwavelength followup. The circumsecondary morphology of forming protoplanets is poorly understood and is likely to involve significant contributions from coorbital dust in the form of accretion streamers and/or an accretion disk. Therefore, broadband photometry alone may be misleading in characterizing accreting protoplanets, as some broadband light is likely to originate from dusty circumsecondary structures.

A combination of polarized and total intensity imaging of these point sources will aid greatly in disentangling these contributions, and the Gemini Planet Imager (GPI) is the ideal system to do this. We demonstrated this fact with GPI followup of HD142527B (Rodigas, Follette et al. 2014), which we were able to resolve at Y band in both total and polarized intensity emission despite its tight separation. Intriguingly, the total intensity and polarized intensity sources appear to be offset by ~20mas (3AU), though at just 2 sigma confidence, so this is a result that requires followup with both systems. In any case, the fact that a polarized intensity source was observed at all indicates that scattered light from circumsecondary material contributes to the brightness of HD142527B, and broadband photometry should thus be considered simply an upper limit. Further multiwavelength imaging in both total and polarized intensity if therefore needed to

characterize the object and obtain a robust mass estimate.

Since publication of the initial HD142527 MagAO result, I have improved our Angular Simultaneous Differential Imaging (ASDI) pipeline, integrated PCA (Soummer et al. 2012) algorithms for more sophisticated PSF removal, and obtained a second epoch of MagAO data on this object. We also replaced the original set of $H\alpha$ SDI filters with a new single substrate filter pair, significantly reducing filter ghosts and therefore image artifacts. In light of the GPI result, I have reprocessed our 2014A MagAO data and find an extension in the direction of the polarized source in many PCA reductions, shown in Figure 5.5. I am preparing this and several other preliminary GAPlanetS results for journal submission in 2015.

5.3.2 2014A Candidates and Future Work

Based on (a) the performance of the system, (b) the accretion luminosity of HD142527, (c) theoretical predictions that gas is funneled through transitional disk gaps in accretion streamers that interact with planets (Dodson-Robinson & Salyk 2011), and (d) recent observations supporting the existence of these streamers (Casassus et al. 2013), we estimate that MagAOs H α SDI imaging mode should be able to detect accreting protoplanets with masses as low as $0.5 M_{Jupiter}$ and at separations as tight as 80mas (Close et al. 2014).

This mass and separation regime is interesting because it corresponds to planets similar to the gas giants we see in our own solar system, a regime that has been inaccessible to the direct imaging technique until now. It also allows us to isolate objects at a key time in the coevolution of planets and disks, as disk material is being swept up and accreted onto forming protoplanets. Only one planet has previously been detected (and not directly) inside of a transitional disk gap (LkCa15b, Kraus & Ireland 2012), but if the prevailing theories for the formation

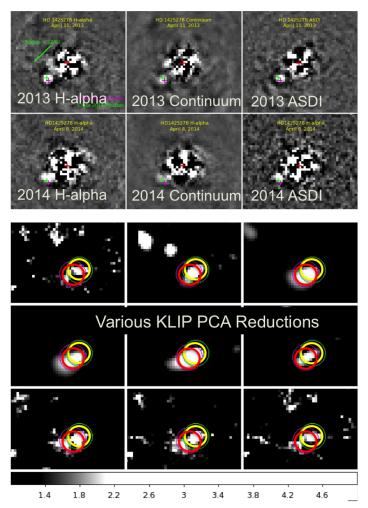


Figure 5.5: SNR Maps of MagAO images of the 86mas (~13AU) companion HD142527B. Top panel: Improved ASDI processing of the 2013 data and 2014 followup observations (Follette et al. 2014). Orbital motion of the companion is apparent. Bottom panel: Various zoomed-in PCA reductions of the 2014 data. The location of the ASDI source, which we assume is the Rodigas, Follette et al. 2014 total intensity source, is marked with a yellow circle. The red circle shows the offset of the polarized intensity source in the GPI data. An elongation in this direction is often present in PCA reduced images, indicating a possible extension (circumsecondary material?) in this direction.

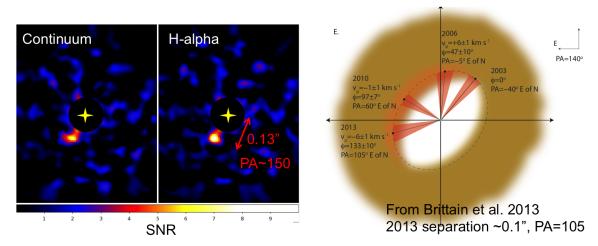


Figure 5.6: MagAO GAPlanetS PCA reduced images of HD100546 revealing an apparent H α point source at 0.13" and PA \sim 150 that is bright (\sim 8 sigma) at H α and \sim 20-50% fainter in the continuum. This location corresponds roughly with the April 2014 position of a planet candidate predicted by Brittain et al. 2014 based on high resolution CO spectroastrometry over a 10 year baseline (right). Reproduced from Follette et al. 2015 (in prep).

of these gaps is correct, there should be many more for us to find. No other currently operating AO systems are capable of imaging planets at such low masses and at such tight separations. This is due to the resolution advantage offered by a visible wavelength AO system and the additional H α luminosity of accreting objects, which makes the planet-star contrast at this wavelength much (\sim 1000x) more favorable for detection.

Of the eight disks already observed through GAPlanetS, we have identified three more accreting point source candidates in addition to HD142527. One of these candidates is highlighted in Figure 5.6, and all three are slated for followup in 2015A. If these GAPlanetS point sources are verified, they may prove to be the tightest separation directly imaged giant planet candidates to date.

5.4 Future Work

I have accepted a postdoctoral position at Stanford University with the Gemini Planet Imager Exoplanet Survey (GPIES) Team to begin January 2015.

GPI is in many ways the best available AO system to follow up GAPlanetS candidates at other wavelengths. Although we are able to operate MagAO's NIR camera Clio2 simultaneously during our GAPlanetS observations, GPI is much better suited to probe into the regions where the candidates lie for several regions. Clio2 is an engineering grade detector and is plagued by artifacts that make very high contrast imaging difficult. Although plans are being made to upgrade it to a science quality detector, these are not likely to be implemented until late 2015 at the earliest. Furthermore, GPI is optimized for high contrast coronagraphic imaging, has polarimetric capability (which Clio2 does not) and is equipped with non-redundant aperture masks to aid in resolving very tight companions.

5.4.1 Multiwavelength Polarized Intensity Disk Imagery

In addition to resolving GAPlanetS point sources in polarized intensity, GPI is uniquely suited to conduct a campaign to probe the broader planet-forming environment of the disks in which they are embedded and to image the asymmetries (spiral arms, divots, warps, accretion streamers) excited by their presence. Polarized Intensity imaging is a long-standing and highly successful method for imaging circumstellar disks. With the exception of sub-mm disks, most ground-based resolved disk images today are obtained through this technique. All of the transition disks in nearby star forming regions that are sufficiently bright for AO imagery have been resolved in at least one NIR wavelength, however ground-based AO imaging of these disks still has much to contribute to our understanding of these systems. In particular, three improvements can be made over already

obtained images.

- 1. Tighter inner working angle. Saturation effects, instrumental polarization (which leaves some stellar residuals at the center of the image) and the speckles induced by non-common path errors limit the inner working angle of polarized intensity observations. GPI implements a novel "integral field polarimetry" technique (Perrin et al. 2008, 2010, 2014) that eliminates non-common path wavefront error between the two channels and allows for significant improvements in inner working angle. Although the architecture of the GPI system also induces some unique complications to polarimetric data reduction, these have been well characterized and calibrated (Wiktorowicz et al. 2014), and improvements in inner working angle over previous space and ground-based imaging campaigns have already been demonstrated with on-sky data. For example, Perrin et al. (2014) recently reported on GPI polarimetric imaging of the HR4796A debris disk ring, which they were able to resolve at all azimuths, including those that appear closest to the star in projection.
- 2. Higher resolution and contrast. GPI has also demonstrated significant improvements in Strehl Ratio over other currently operating high contrast imagers, and this translates to improvement in both resolution and contrast. For example, GPI was able to achieve an H-band Strehl ratio of 0.89 during commissioning on the star Beta Pictoris (Macintosh et al. 2014). In contrast, most ground-based polarized intensity disk images to date were created with either the Subaru telescope's AO188 system, which achieves H-band Strehl ratios in the 0.2-0.3 range (Grady et al. 2014, in prep), or the VLT NaCo Instrument, whose top performance is ~40% at H-band Girard et al. (2012) and, as of this writing, is no longer functioning.

3. Multiwavelength imagery My recent work on the transition disk Oph IRS 48 (Follette et al. 2014) highlights the power of obtaining polarized intensity imagery of the same disk at multiple wavelengths, which GPI is capable of doing with better contrast, resolution and inner working angle than most other currently available instruments. In particular, this work demonstrates two important advantages of multiwavelength polarized imaging. First, resolving the same disk at multiple wavelengths allows for verification of the reality of structures seen at one wavelength. Second, it allows us to identify regions in the disk with unique grain properties. Oph IRS 48 highlights in particular that grain properties can deviate not just radially within a disk, as has long been suspected, but also azimuthally. This information is important for classifying the asymmetries that we observe in these disks and in understanding the physical processes (e.g. planet formation) that drive their creation and will ultimately help to inform planet formation theory.

5.4.2 Total Intensity Disk Imagery

Polarized intensity imagery is powerful in and of itself, but it has the added advantage that total intensity imagery is recovered at the same time (through adding rather than subtracting the polarization states). While total intensity disk imagery should in principle be a very powerful complementary technique, in practice, very few successful total intensity imaging campaigns have been conducted from the ground. This is because total intensity imagery requires robust removal of stellar light in order to isolate disk emission from stellar emission (whereas disk scattered light and starlight are easily separated via polarization). This is done either through so-called "classical" PSF subtraction, whereby a separate star is observed to serve as a disk/planet free PSF, or through post-processing algorithms (e.g. ADI, LOCI and PCA), which are designed to estimate and re-

move starlight using observations of the science target itself. I will describe the potential in each technique, as well as how I propose to investigate them, below.

From the ground, without the point spread function stability of a space-based telescope, classical PSF subtraction is a particularly difficult endeavor. In the case of adaptive optics observations, it requires opening the AO loop and relocking it on a PSF star. A variety of parameters affect the quality of PSF match between a comparison star and science target. These include airmass, magnitude at both the wavelength of the wavefront sensor (generally optical) and the science camera (generally NIR), variations in seeing, instrumental flexure, and the amount of time between observations of the two targets. Some AO observers elect to alternate rapidly between the target and PSF star, and others observe each in succession. Despite numerous attempts and a limited number of successes (e.g. Honda et al. 2009; Hinkley et al. 2009), a systematic investigation of techniques for AO PSF selection and observation in order to optimize total intensity disk imagery has never been conducted, nor are the techniques that don't work generally reported on. I myself have simply included a statement to the effect of: "we observed a PSF star and attempted classical PSF subtraction to obtain total intensity imagery, but were unsuccessful due to poor PSF match" in both of my own polarized imaging papers (Follette et al. 2013b, 2014). In fact, in the entire SEEDS campaign, despite 15 resolved polarized intensity images of protoplanetary and transitional disks, none have been successfully PSF calibrated to yield a total intensity image.

I have proposed to complete a systematic investigation of optimal observation strategies for AO PSF subtraction by (a) utilizing the ever-growing GPI database of disk observations to analyze the quality of PSF subtraction according to the parameters described above and (b) completing an on-sky experiment in the use

of AO telemetry to optimize PSF estimation, a technique that has been demonstrated in theory (e.g. Veran et al. 1997; Poyneer et al. 2009) but has not been extensively tested on-sky. Membership on the GPI team will aid significantly in both endeavors, as I will have access to the extensive campaign database, will be able to participate in data collection for the campaign, and will have access to the world's leading experts on AO telemetry (e.g. Jean-Pierre Veran, Lisa Poyneer) for collaboration in designing on-sky experiments.

Perhaps equally important will be a systematic exploration of the feasibility of various post-processing techniques for obtaining total intensity disk imagery. In recent years, many important advancements have been made in the field of post-processing techniques for high contrast imaging with adaptive optics, however most of these techniques, which include Angular Differential Imaging (ADI, Marois et al. 2006), Locally Optimized Combination of Images (LOCI, Lafrenière et al. 2007) and Principal Component Analysis (PCA, Soummer et al. 2012), are optimized for point source detection and do not translate well to disk imaging. The primary reason for this is that these techniques all rely on on-sky rotation to isolate a planet signal from stellar speckles, which are fixed in the detector frame. Because disk emission does not vary azimuthally in the same way, it is much more difficult to isolate it through rotation. Some successes have been made in applying these techniques to disk imagery, but only in the case of edge-on disks (e.g. Rodigas et al. 2012, 2014) or ring-like geometries (e.g. Perrin et al. 2014), and these still suffer from so-called "self-subtraction", whereby some disk light is present in the PSF and is subtracted from the science image. I will also complete a systematic investigation of the usefulness of these techniques for isolating disk emission and attempt to identify the regions of disk morphology parameter space (inclination, extent, etc.) in which each is viable. The GPI team includes experts on all three techniques, and so is the perfect environment in which to complete and publish such a systematic investigation.

5.4.3 Polarization Mapping

A number of imaging campaigns have succeeded in comparing polarized and total intensity emission from the same disk and have demonstrated the power of the two in disentangling the degeneracies between grain size, porosity and composition. These successes have either been (a) entirely space-based (e.g. Perrin et al. 2009; Graham et al. 2007), in which case only the largest and brightest disks can be imaged due to the limitations of the HST aperture, or (b) have relied on non-contemporaneous space-based total intensity data and ground-based polarized intensity data to glean the polarization fraction (e.g. Muto et al. 2012; Hinkley et al. 2009; Hales et al. 2006). GPI has already demonstrated, the ability to do both simultaneously in the case of a ring-like disk geometry through PCA post-processing (Perrin et al. 2014), which bodes well for future GPI successes in this area.

My personal goal as a member of the GPI team will be to optimize total intensity imagery through the techniques outlined above, to use that information to identify disks whose geometry is favorable for both polarized and total intensity imaging, and to image those disks at multiple NIR wavelengths in polarized and total intensity, focusing on those with identified planet candidates. Such an imaging campaign will provide unparalleled constraints on grain properties such as size, porosity and composition in these disks, and therefore constraints on the materials from which planets form, as well as radial and azimuthal variations in those properties. This study will thus help to inform both planet formation theories and the debate on disk clearing mechanisms.

CHAPTER 6

AN INFORMED APPROACH TO IMPROVING QUANTITATIVE LITERACY AND MITIGATING MATH ANXIETY IN UNDERGRADUATES THROUGH INTRODUCTORY SCIENCE COURSES

6.1 Background and Literature Review

As technology, politics, medicine and even ethics adapt to the challenges of modern life, educated people are faced with a continuous barrage of statistics, projections and graphical representations of trends intended to inform their decisions, and yet some estimate that as few as one-half of American adults have mastered the skills necessary to interpret such data. Innumeracy can have dire consequences for the American adult, as lower numeracy skills are postulated to result in lower quality of life and more limited employment opportunities (Kirsch et al. 1993; Charette & Meng 1998).

In its most traditional sense, literacy is founded on the "3 R's" of primary education: reading, writing and arithmetic. Indeed, the vast majority of quantitative data that people are confronted with in everyday life takes the form of arithmetic, yet we find that increasingly, students in introductory college science courses lack the skills required to interpret the quantitative data in which our society is afloat. In the current education literature, the skills required to achieve such proficiency are termed "quantitative literacy", "quantitative reasoning" or "numeracy", and I will use these terms interchangeably throughout. Like reading and writing, quantitative literacy is nothing without context, and in this sense it is fundamentally different from pure mathematics. In the same sense that a student's ability to decipher the meaning of Shakespeare or Poe does not necessarily support them in understanding the average news article or rental agreement, quantitative lit-

eracy is nothing without real world context. As rather eloquently put by Hallett "Quantitative literacy describes a habit of mind rather than a set of topics or a list of skills. . . . [It] insists on understanding" (Madison & Steen 2003).

The average citizen needs to be quantitatively literate in order to make informed decisions about many situations encountered in daily life, including but certainly not limited to: interpretation of political debates, medical statistics, and price comparisons. In fact, studies suggest that innumerate adults are less successful at managing their own health (Schwartz et al. 1997; Apter et al. 2009), understanding the informed consent process (Couper & Singer 2009) and comprehending other forms of risk-benefit analysis (Peters et al. 2007).

While it is clear that quantitative literacy is distinct from the ability to perform arithmetic out of context, it is also different from mathematics. There is a "historical dichotomy of mathematics as academic and numeracy as commercial" (Steen 2001), but today the two have become almost entirely separate pursuits. School curricula have diverged from societal needs in mathematics perhaps more than in any other discipline. What are the goals of a traditional education if not to prepare our students for their roles in society? Is it any wonder that even Honor's College undergraduates in our non-major astronomy classes state that "Math is only relevant to pass at test in math class?".

Certain mathematical skills, such as the ability to complete simple arithmetic operations, are necessary for the achievement of quantitative literacy, however the current structure of mathematics education is largely vertical, building abstracted mathematical skills on top of increasingly more abstract concepts. While mathematics advances vertically, numeracy proceeds sideways, applying the same arithmetic skills in broader and broader contexts. "Unlike mathematics, numeracy does not so much lead upward in an ascending pursuit of abstraction as it

moves outward toward an ever richer engagement with life's diverse contexts and situations (Madison & Steen 2003). The difference between mathematics and quantitative literacy, Ewell explains in Mathematics and Democracy (Steen 2001) "may be in part because early exposure to mathematics presents it as a distinctly different activity from natural forms of communication. Reading and writing thus appear to be expected extensions of everyday life in ways that are not necessarily true of mathematical concepts." Over the past two decades, numeracy has been increasingly recognized as an essential component of literacy as a whole. It has been measured as one of three fundamental areas of literacy, along with prose and document literacy, on the Department of Educations National Assessment of Adult Literacy since 1985 (called the Young Adult Literacy Assessment in 1985). The 2003 results indicated that 55% of adults had either below basic or basic quantitative literacy skills, with only 13% performing at the level of proficient (US Department of Education, 2003).

American adults also perform poorly relative to adults in other countries, as revealed in the 2013 Organization of Economic Cooperation and Development (OECD) report on the results of the Programme for the International Assessment of Adult Competencies (PIAA). American adults ranked 21st among the 23 countries measured in the area of numeracy, with 67% of adults 16-65 surveyed falling into the lowest two (of five) proficiency levels.

Innumeracy is a common complaint of businesses and universities (Steen 1999). They malign the fact that their students and employees do not enter higher education and the work force with the skills necessary to function in those settings. Many college professors lament the necessity to remediate, reinforcing quantitative reasoning skills that were in theory part of the standard high school, junior high or even elementary school curriculum, but that many students leave the sec-

ondary school system without being able to apply. Businesses invest billions of dollars per year in training programs, many of which are designed to help employees gain mastery of basic skills that also should have been learned in high school or earlier (Rosen et al. 2003). In order to correct this problem, both the pedagogy and the psychology of quantitative reasoning need to be carefully addressed.

The literature, and particularly the compilations by the National Council on Education and the Disciplines: Mathematics and Democracy: The Case for Quantitative Literacy of 2001 and Quantitative Literacy: Why Numeracy Matters for Schools and Colleges of 2003, (there are referred to throughout as references [1] and [2]) diagnose a number of "problems" with the current approach to education in the United States. Among them are the willingness of teachers to show their students the "method" when confronted with a difficult application problem, a general discomfort with estimation and preference for precision, a tendency to fuel explanations with mathematical proofs and definitions rather than a concrete application that will give students an intuitive understanding of the topic, and a general level of abstraction resulting in a failure to provide real world contexts for mathematical reasoning. Together, these lead to a phenomenon that so many of us are familiar with - "math anxiety.

Math Panic and Quantitative Illiteracy are complex and deeply rooted problems in the modern educational system. Concrete suggestions for steps that need to be taken in order to reverse the trend of innumeracy in educated adults abound in the literature (notably [1] and [2]) and fall broadly into four categories: mathematics curricular reform, cross-curricular application, faculty development and training, and revision of admissions standards. In the areas of mathematics curricular reform and revision of admission standards, we, as educators in the sciences and non-administrators, can only be aware of the recommendations in the literature and support our administrators and mathematics colleagues in addressing them, however in the areas of cross-curricular reform and faculty development and training, we feel that introductory science courses are an indispensible forum to investigate and inform solutions to this problem. I will address the suggestions of the literature in regards to these two areas below, and move on to describe my study in sections 2 and 3.

6.1.1 Cross-Curricular and Continuous Application

We recently received an application to Astronomy Camp from a 14 year old female honors student that included the statement I have always heard a lot about science and math being related, but I have never actually used them together. I have finished an entire year of algebra, but there has never been any science in it. The same holds true in my science class. Frankly, I have never seen any connection between the two of them. This heartbreaking statement is compelling evidence as to the importance of cross-curricular application of mathematics.

Literacy in all forms is by its very nature cross-curricular. A literate citizen needs to have developed reading, writing and quantitative reasoning skills to the extent that they are an integral part of the way they view the world. This means that quantitative reasoning, like reading and writing skills, cannot be confined to the study of math, but must be applied in all disciplines, a conclusion that is supported by many across the curriculum:

• "teachers must encourage [students] to see and use mathematics in everything they do Only by encountering the elements and expressions of numeracy in real contexts that are meaningful to them will students develop the habits of mind of a numerate citizen. Like literacy, numeracy is everyone's responsibility" ([1] The Case for Quantitative Literacy),

- "Although the mathematical foundation of quantitative literacy is laid in middle school, literacy can only be developed by a continuous, coordinated effort throughout high school and college" ([2] Hallett)
- "Mathematicians have a lot of work to do to convince students that they are teaching something useful. Having faculty outside mathematics include quantitative problems in their own courses is extremely important. These problems are much more likely to be considered realistic" ([2] Hallett)

These themes are addressed over and over again in the literature, yet innumeracy is often perceived to be a problem only for mathematics and mathematics educators. Widespread "math panic" in the general population and among teachers in every discipline seem to have led to a certain comfort with its compartmentalization.

Perhaps most importantly, we believe that those of us who were educated in traditional mathematics and acquired quantitative literacy through disciplinary necessity need to step back and make a conscious effort not to formulate our curricula according to currently held mathematical norms.

We all, regardless of discipline, need to become crusaders for quantitative literacy as an essential skill to be instilled in our students. To do any less would be to do them a disservice. Many scholars cite the rising popularity of statistics courses as evidence of a readiness to change ([2] Scheaffer). Furthermore, the case is made in the literature that "secondary school teachers, who historically have been obedient foot soldiers on the precalculus drill team, have now developed an appetite for relevant mathematics" ([1] Kennedy). The atmosphere is primed for change. We as educators now need to embrace it, and measuring and teaching applied mathematics in the science classroom is one step in a coordinated effort

to support individuals in developing the quantitative literacy skills they need for life by using practical mathematics in context.

6.1.2 Faculty Development and Training

"One [difficulty in cross-curricular implementation of QL] is the notion of 'special expertise' A major challenge in implementing writing across the curriculum, for instance, is the fact that faculty do not all know automatically how to coach or assess writing effectively, so substantial efforts at faculty development are generally required I believe the same level of effort is required for quantitative literacy." ([1] Ewell).

Faculty training and continued development are necessary counterparts to curricular reform, both to teach faculty in the quantitative disciplines to effectively and consciously address quantitative literacy in their classrooms and to teach faculty in non-quantitative disciplines how to introduce quantitative examples and analysis tools into their classrooms. Among physics and astronomy faculty teaching introductory college science courses, instructional development workshops, particularly those conducted by fellow STEM faculty, have been shown to be effective at raising awareness of new teaching techniques and encouraging changes in teaching practice (Henderson & Dancy 2008; Dokter 2008). Faculty development and training should therefore be an integral component to any attempt to reverse the trends of innumeracy and math panic in the population.

My study of the cross-curricular application of numeracy in the context of science is described in Section 2, and a faculty development workshop addressing quantitative literacy tips and techniques that I have been developing with Don McCarthy is described Section 3.

6.2 Investigatory Study Component

We have embarked on a study to inform whether the inclusion of numerical skills into the curricula of introductory science courses for non-majors improves students comfort with and ability to manipulate and reason with numbers (i.e. their quantitative literacy). This study revolves around development and administration of an assessment instrument that we have called the Quantitative Reasoning for College Science (QuaRCS) Assessment. We hope that the results will ultimately reveal whether introductory science courses for non-majors are a good forum for developing numerical skills and improving attitudes towards mathematics.

6.2.1 Study Motivation

This project was born from a puzzling observation that I made during my first semester of teaching at Pima Community College (Fall, 2009). I noticed that my students shut down immediately when I wrote an equation on the board, whether I intended to have them manipulate it or not. I was puzzled and disturbed by this, having never encountered it in my high school classroom, where it was expected that mathematics would be a key part of the curriculum. I attended the Astronomical Society of the Pacific's Cosmos in the Classroom conference on undergraduate astronomy instruction the following summer, and one of my principal goals was to try to find out what caused this problem and what I could do about it as an instructor.

Instead of encountering solutions, I found myself faced with a generally antiquantitative attitude and rhetoric in the astronomy education community. I heard many people lament that it "was not our job to remediate arithmetic skills that our students should have mastered in elementary school. The push to emphasize purely qualitative astronomy disturbed me for two reasons. First, it was in precisely a class like my own (an undergraduate astronomy course for non-majors) that I fell in love with science for the first time, and a large component of that was the revolutionary experience of seeing practical applied mathematics for the first time. The second reason was that the simple quantitative skills I wish to emphasize in my classroom graph reading, order of magnitude estimation, proportionality, simple arithmetic, error analysis, etc. are, I believe, absolutely essential in life. Such skills are not just important to the understanding of science, but are the key to making people savvy consumers, discerning voters, conscientious citizens and recognizers of pseudoscience in the media. If my class is their terminal science class in life, then arent I doing them a disservice if I excise the quantitative from my curriculum?

Science faculty have been trained to manipulate and interpret numerical data, and quantitative literacy is fundamentally applied mathematics, so the sciences are a natural place to begin emphasizing such skills. A majority of college students are required to complete one course in science before graduation in order to fulfill distribution requirements. Data compiled by the American Institute of Physics (2010) reveal that introductory astronomy enrollments have remained in the 180,000-190,000 range since 2004 (Nicholson & Mulvey 2010). This figure does not include community colleges, where an estimated 100,000 students take Astronomy 101 in departments not covered by the AIP survey (Fraknoi 2001). According to a Bureau of Labor Statistics report entitled College Enrollment and Work Activity of 2010 High School Graduates, in Fall 2010, 2.2 million people were enrolled in college in the U.S. This means that more than 10% of college students eventually pass through the door of an Astronomy 101 course in college. This large population is why we feel that introductory astronomy courses

for non-science majors are a logical place to begin a study of the efficacy of science courses in improving quantitative literacy in the general college population.

While simple mathematical skills are essential to the understanding of the basics of virtually every scientific discipline, the increasing discomfort and lack of ability of college students to think and reason numerically has led many instructors to remove nearly all of the mathematical reasoning from their non-major science courses. Examples of mathematical misconceptions that I and my colleagues have encountered in our classrooms (a very small subsample) are given in Table 1. We have also observed a rapid change in the rhetoric of and attitudes expressed toward mathematics in the sciences during teaching conferences over the course of the last decade. The attitude towards mathematics in the introductory science classroom appears to be moving in the direction of conscious avoidance. We have often heard instructors expressing sentiments such as my students cant do it, and I dont have time to remediate, so why bother?

Table 6.1: Common mathematical misconceptions encountered frequently in our classrooms.

Operation	Common Incorrect Answer
1 ÷ 5	0.5
0.5=	5%
How many seconds in an hour?60sec/min + 60min/hr	120sec
$10^2 =$	20
$4.3 \times 10^6 =$	4.3000000

Following the ASP conference, Dr. McCarthy and I discussed why we found the trend toward purging mathematics from intro science curricula very disturbing, and we turned to the literature in search of concrete data supporting the need for the demonstration of practical, applied mathematics in the college science classroom. Unfortunately, we found that while math anxiety and quantitative illiteracy are well-documented phenomena, there is a decided lack of educational research into how to alleviate their effects at the college level outside of mathematics courses. This is the fundamental motivation for this study. We hope to answer some of the key questions about quantitative literacy in the science classroom through data rather than anecdote. This includes assessing whether students can take the quantitative skills they may learn in Astronomy 101 and apply them to quantitative information they encounter in everyday life, as well as whether their attitudes towards mathematics can be improved. We will use these results to inform the development and conduction of a faculty development workshop series to support science faculty and graduate students in learning about evidence-based mathematics teaching practices that can be incorporated into the introductory science classroom. Obtaining such data is essential to inform this debate. This research will be important in providing a first quantitative look at whether and where basic quantitative illiteracies exist among college students enrolled in general education astronomy courses and whether they can be addressed through these courses.

6.2.2 Regarding Important At Risk and Influential populations

Another key aspect of this study will be to inform on how certain pedagogical approaches affect members of populations that are particularly at risk for Math Anxiety and Quantitative Illiteracy. Several of these at risk groups have been revealed in the literature, most notably/comprehensively by the results of the National Assessment of Adult Literacy, which surveyed 26,000 American adults in the areas of prose, document and quantitative literacy in 1992, with a follow-up in 2003. This NALS defined quantitative literacy as "the knowledge and skills required to apply arithmetic operations, either alone or sequentially, using numbers embedded in printed materials; for example, balancing a checkbook, figur-

Table 6.2: Instrument Development Timeline

	N Students		
Semester	Pre/[Mid]/Post	Instrument Description	Summary of changes from previous version
	(Matched)		
	Instructor 1: 32	Format: Scantron	
		Administration: In class. No credit.	
		Pre: Mid-semester	
Fall, 2010		Post: None	N/A
		Questions:	
		a) 3 Multiple Choice quantitative questions	
		b) 1 Essay	
		Format: Scantron	
Spring, 2011	Instructor 1: 70/60	Administration: In class. No credit.	
		Pre: First week of semester	(1) Add dam careable and attitude assertions
		Post: Two weeks from end of semester	(1) Add demographic and attitude questions
		Questions:	(2) Expanded quantitative question bank to 22 (3) Added confidence ranking after EACH quantitative question
		a) 8 Demographic + 5 attitude/skill self-assessment	
		b) 10 ¹ Multiple choice quantitative questions	(4) Pre AND post assessment
		c) Confidence ranking after each quantitative question	
		d) 3 questions reflecting on assessment	

¹22 total quantitative questions split into 3 versions, first 4 repeated for all students

		Format: Online	
		Administration: Out of Class. Participation credit.	(1) Correct for attrition and allow matched data by collecting names
Fall, 2011	Instructor 1: 156/163 (103) Instructor 2: 22/22 (10) Instructor 3: 29/6 (5)	Pre: First two weeks of semester Post: Last two weeks of semester Questions: (1) 11 demographic + 5 attitude/skill self-assessment (2) 22 Multiple choice quantitative questions (3) Confidence ranking after each quantitative question (4) 3 questions reflecting on assessment	(2) Several questions reworded for clarity based on focus group sessions (3) Encourage participation by assigning for participation credit (must offer alternate assignment per IRB) (4) Recruited instructors request online format to free up class time
Spring, 2012	Instructor 1: 67/56 (40) Instructor 2: 23/17 (11) Instructor 4: 77/37 (22) Instructor 5: 438/539 (278)	Format: Online Administration: Out of Class. Participation credit. Pre: First two weeks of semester Post: Last two weeks of semester Questions: a) 12 demographic ² + 5 attitude/skill self-assessment b) 5 questions about course ³ c) 22 Multiple choice quantitative questions d) Confidence ranking after each quantitative question e) 3 questions reflecting on assessment	(1) Add questions about course and whether and how quantitative skills were emphasized to post assessment (2) Removed duplicate demographic question and added two questions about previous science coursework

²only name, course, age and instructor collected in post assessment ³post assessment only

		Format: Online	(1) To address question of whether late semester apathy is contribut-
	T 1 111 /01 /00		ing to low post scores:
	Instructor 1: 111/91/23	Administration: Out of Class. Participation credit.	a. Add mid semester (post midterm) assessment
	(21)	Pre: First two weeks of semester	b. Add question asking students to quantify their effort on the as-
	Instructor 6: 112/79/61	Mid: Week 11 of 16 week semester	sessment
	(50)	Post: Last two weeks of semester Questions: a) Course, Instructor and Name	(2) Major question rewordings and addition of table reading and
Fall, 2013	Instructor 7: 30/17/7		area/volume skills
	(4)		(3) Moved all demographic questions to end of assessment to miti-
	Instructor 8:	b) 25 Multiple choice quantitative questions	gate stereotype threat
	157/102/83 (58)	c) Confidence ranking after each quantitative question	(4) Removed several demographic questions to make room for:
	Instructor 9: 129/0/0	d) 4 questions reflecting on assessment (incl. calculator usage)	a. 5 Likert scale attitude questions
	(0)	e) 8 demographic + 8 attitude	-
		f) Effort question	b. question about calculator usage on assessment
	Instructor 2: 20/9 (6)		c. question about why chose major (1) Revised key math attitude question to Stapel scale format, which
			forces students to place themselves on a scale between two opposite
Spring,		same as Fall, 2013 except no mid-semester assessment	adjectives
2014			(2) Addition of statistics as a skill category based on analysis of in-
			structor surveys
	Instructor 6: 163	Format: Online	
		Administration: Out of Class. Participation credit. Weeks 4-5	(1) Revised questions and distractors to reflect analysis of F13 and
		Questions:	F14 free response administrations
		a) Course, Instructor and Name	(2) Addition/revision of several new questions to fill out skill cate-
Fall, 2014		b) 25 Multiple choice quantitative questions	gories, and removal of several redundant or low-performing ques-
		c) Confidence ranking after each quantitative question	tions to make room
		d) 3 questions reflecting on assessment (incl. calculator usage)	(3) Added questions about ethnicity, disability status, and future
		e) 8 demographic + 8 attitude	coursework in mathematics
		f) Effort question	

	Instructor 1: N=30-72	Format: Free Response Questions		
Spring, 2012		Administration: In Class. Participation Credit. Throughout	(1) Generated new authentic distractors based on misconceptions	
		Semester.	(2) Reworded for clarity	
		Questions: (3) Revised in cases where students arrived at correct a		
		5 Open-Ended Free Response Versions of Questions from the Assess-	incorrect reasoning.	
		ment		
		Format: Free Response Questions		
	Instructor 1: N=73-106	Administration: In Class. Participation Credit. Throughout	(1) Generated new authentic distractors based on misconceptions	
E II 2012		Semester.	(2) Reworded for clarity	
Fall, 2013		Questions:	(3) Revised in cases where students arrived at correct answer through	
		11 Open-Ended Free Response Versions of Questions from the As-	incorrect reasoning	
		sessment		
		Format: Free Response Questions		
Fall, 2014	Instructor 1: N=35-68	Administration: In Class. Participation Credit. Throughout	(1) Generated new authentic distractors based on misconceptions	
		Semester.	(2) Reworded for clarity	
		Questions:	(3) Revised in cases where students arrived at correct answer through	
		21 Open-Ended Free Response Versions of Questions from the As-	incorrect reasoning	
		sessment		

ing out a tip, completing an order form, or determining the amount of interest from a loan advertisement, and included among the key results were the following conclusions:

- 1. more than 20% of American adults performed at the lowest level of quantitative proficiency, and an additional 25% of adults demonstrated the next highest degree of proficiency. Both levels indicate an inability to perform more than a single operation or use more than one number from a document to determine information.
- The quantitative proficiency of younger adults was lower than that of older adults, suggesting that the general numerical proficiency of the US population is declining with time.
- Individuals with the highest numerical proficiencies reported weekly wages
 2-3 times higher than individuals in the lowest levels and were more likely to be employed.
- 4. Black, American Indian/Alaskan Native, Hispanic, and Asian/Pacific Islander adults were more likely to perform in the lowest two levels than White adults.
- 5. The quantitative proficiencies of men were somewhat higher than those of women.
- 6. Adults with physical, mental or other health conditions were more likely to perform in the lowest two levels of proficiency.

Conclusions (4) through (6) regarding at risk populations have also been revealed in other studies. Particularly relevant are results suggesting that: women are more susceptible to math anxiety (Betz 1978), minorities are more likely to

leave college with deficits in numeracy (Raudenbush & Kasim 1998; Ogbu 1990), and that students with learning difficulties are particularly likely to struggle with applied mathematics (Chan & Dally 2001; Rousselle & Noël 2007). We are also acutely aware that many (up to 15% according to initial survey responses) of our non-major students are pre-service teachers, and we feel that this is another particularly important classroom demographic to study. The literature suggests that many people learn math anxiety from their teachers (Beilock et al. 2010; Jackson & Leffingwell 1999), so leaving this group of students with, at the least, an improved attitude toward mathematics is a very important step towards mitigating math anxiety and improving quantitative literacy in future generations.

6.3 Faculty Development Component

We have also launched a workshop series on Techniques, Tools and Tips for Improving Quantitative Literacy in the Science Classroom. We see the goals of such a program as threefold

- (a) To raise awareness of the problem of serious deficits incoming college students have in the area of mathematical reasoning
- (b) To share evidence-based tips, tricks, language, and curricular materials that we have collected and developed, which will be increasingly informed by the results of our study. Our experience shows these methods can be integrated into the existing science curriculum to augment, not subtract from, science content and understanding. This overall approach could inspire an even broader impact on the inclusion of numerical reasoning and expression in non-science subjects.
- (c) To recruit instructors interested in further investigation to participate in our

study

Regarding (a) Data verifying the problems of innumeracy and math anxiety abound in the literature, but many science instructors lack the time and/or inclination to keep up on the broader educational literature, and rely on workshops to boil it down to its important points. This has been very successful in the area of learner-centered teaching techniques, and we hope that these workshops will be met with similar enthusiasm by the community.

Regarding (b) We have presented the following materials and techniques at our workshops over the past three years.

- Daily quantitative brainteasers
- Quantitative supplements to questions in the very popular Lecture Tutorials in Introductory Astronomy workbook
- Homework and test questions involving simple arithmetic reasoning
- Math Skills activities that include both astronomical and real world contextual problems
- Labs in which student evaluate the reasonability of numbers presented in popular media
- Labs in which students examine selection effects and how to recognize them in data
- "Real life analogs to astronomical formulae and graphical relationships
- a variation of the very popular science v. pseudoscience curricula espoused by experienced educators such as Doug Duncan (CU Boulder) with the additional component of how to recognize good mathematical data vs. bad

Regarding (c) Workshop participants will be asked to participate in Phase II of the study, and we hope that the workshop will convince them that whether and how numeracy should be emphasized in the science classroom is an open question worth pursuing.

6.4 Conclusion

The dual problems of math anxiety and quantitative illiteracy are widely acknowledged and lamented in the educational literature, and a resounding call for action has been issued. We propose to rise to this call and begin studying how numeracy issues can be resolved or mitigated using introductory sciences, which we feel are uniquely suited to this purpose. Introductory science courses for non-majors, as they are often a students last opportunity to see applied mathematics in a formal educational setting, are an essential forum for promoting quantitative literacy.

Results from the literature further suggest that neither math anxiety nor innumeracy are correlated with math ability (Ashcraft 2002). Both are strongly correlated, however, with math avoidance. This means that both innumeracy and
math anxiety are significant sources of leaks in the STEM pipeline. In fact, the relationship between low self-efficacy expectations in mathematics and a tendency
to avoid science-based college majors was revealed more than 25 years ago (Betz
& Hackett 1983), yet very few studies have since investigated how to alleviate
it, and none through the vehicle of college-level introductory science. We believe
that our study is an important first step in revealing whether introductory science
courses for non-majors can be used to reverse negative attitudes towards mathematics, alleviate math anxiety and develop the numerical skills that are essential
for our students success in myriad aspects of modern life.

CHAPTER 7

THE QUANTITATIVE REASONING FOR COLLEGE SCIENCE (QUARCS) ASSESSMENT I: DEVELOPMENT AND VALIDATION

7.1 Abstract

This article reports on the development and validation of the Quantititive Reasoning for College Science (QuaRCS) Assessment, a numeracy assessment instrument designed for college-level general education science courses. In particular, it outlines the selection of specific quantitative skills to be assessed on the QuaRCS, question development, and preliminary results from the initial administration. These preliminary results include performance statistics for individual items and the instrument as a whole as well as student response distributions. We include statistics on attrition and apathy that we believe have wide-ranging implications for the pre-post assessment instruments used in much of educational research. We also report on a survey of science and math educators used to inform instrument development, which reveals dichotomies between their opinions regarding the utility of certain quantitative skills.

7.2 Introduction

The academic literature contains a plethora of compelling arguments for why numerical skills are critical to success in everything from managing one's personal health (Schwartz et al. 1997; Apter et al. 2009) to understanding the informed consent process (Couper & Singer 2009). In fact, the issue of rampant innumeracy among American adults has gained so much attention in recent years that it frequently grabs national news headlines (e.g. Green 2014; Perez-Pena 2013) and has been the subject of a number of works of popular literature (e.g. Seife 2010;

Bennett 2014; Schneps & Colmez 2013). We ourselves have written extensively on the subject of why numerical skills are critical to science literacy (Follette & McCarthy 2012; Follette & McCarthy 2012, 2014; McCarthy & Follette 2013), and similar arguments for its importance have been put forward in many other fields. As the importance of numerical skills for success in modern life has been argued in many other venues, we will not belabor it here, but will focus instead on next steps in the effort to combat innumeracy.

Missing in much of the literature on numeracy ¹ are specific, research-validated strategies for mitigating innumeracy and improving the numerical skills of American students and adults. This is a fundamentally difficult problem, and tackling it requires not just proven instructional strategies or techniques, but also assessments capable of measuring learners' quantitative abilities and improvements in them. In this article, we present one such assessment specifically targeted toward students in general-education college science courses.

7.2.1 Why Should Science Educators Join the Numeracy Crusade?

Despite the natural disciplinary entwinement of science and mathematics, instructors of general education science courses (taken predominantly by non-majors) often move away from an emphasis on quantitative skills in their courses, due primarily to the well-documented innumeracy of American adults ². In our experience leading workshops for science educators on numeracy and presenting at science education conferences, the laments on this issue boil down to three main complaints/concerns.

 $^{^{1}}$ We will use this term as well as quantitative literacy and quantitative reasoning interchangeably throughout this paper

²This has been demonstrated in both national and international surveys including the National Assessment of Adult Literacy (2003), National Adult Literacy Survey (1993) and the Organization for Economic Cooperation and Development's Program for International Student Assessment (2013)

(a) It's too late for them.

College science educators often express frustration at the numerical deficits of their students, but they also express the belief that if a student's entire K12 education has failed to instill the necessary skills to understand science at a quantitative level, then its unlikely that in a single course they will be able to significantly increase students' quantitative abilities.

This concern was the principle motivation for the development of the QuaRCS, whose refinement and validation as reported on in this article constitutes Phase I of our study. The goal of Phase I has been to devlop an instrument capable of measuring quantiative abilities in specific areas deemed most important for the understanding of science content and for success in life (selection of these skills is described in Section 7.3.1.

The next step (Phase II) will be to implement the instrument in general education college science classrooms across the country. We will endeavor to address the questions of whether (a) the QuaRCS is well-suited to measuring change in quantitative abilities and, if so, (b) which college science educators are making the largest difference in their students' quantitative abilities. We hope that identification of these individuals will help to inform which pedagogical techniques, class formats, etc. are most successful in fostering improvement.

(b) It's not my job to teach math.

Probing comments like this at our workshops, it appears that this concern is twofold. First, science educators worry about having the pedagogical wherewithal to remediate mathematical skills. Second, they express concern that engaging in such remediation will reduce the amount of time that they can spend on science in their course, and their students will therefore develop lower levels of understanding of disciplinary content.

Regarding the first concern, science is by its very nature quantitative, and scientists are experts at interpreting quantitative information. The commonly-held incorrect belief that appears to prevent science educators from emphasizing mathematics in their courses is that they will need to remediate mathematics as a separate entity. This belief is pervasive, and is revealed, for example, in one educator response to the skills survey that we report on in Section 7.3.1.

in general, I am currently going the opposite direction and trying to reduce/simplify the math that is expected of my students. . . The main reason I would consider adding math is to improve my students' numeracy, not to improve their understanding of astronomy, which is why I haven't done it. I think it's a good goal, but my main goal is to increase students' self-efficacy and help them to have a growth mentality about learning, and adding more math would make that more difficult"

Beliefs such as those reflected in the statement above are troubling simply because quantitative literacy is fundamentally mathematics in context, and it includes not just calculations, but reasoning with numbers, visual representations of data, and spatial resoning. The context in the case of a science classroom is the science. To excise the quantitative from science curricula is to engage in the shortsighted belief that you can teach real science without engaging learners in quantitative tasks and reasoning.

The second concern has been well addressed by others, most notably (Hester et al. 2014). This group showed that students in a Biology course where remediation of numerical skills was emphasized scored equally well on general Biology content questions at the end of a semester of instruction as courses that didn't emphasize numerical reasoning. At the same time, the intervention course showed substantially higher gains on so-called "BioMath" questions (Biology questions housed in a quantitative context). Thus, this group demonstrated that adding

numerical emphasis to a course does not necessarily result in reduced discipline content gains. Introductory science students can improve in their ability to answer quantitative questions in science contexts.

Given this work, we set out to answer a separate, though related, question with our assessment. Are students able to apply numerical skills when they encounter "real life" problems that require similar reasoning to problems that they might encounter in a quantitatively-rich science course? In other words, when quantitative skills are emphasized in context in a science classroom, are skill improvements transferable to other contexts? This is a fundamentally more difficult problem, but we believe an important one. In our Phase II implementation of the QuaRCS, we will administer the instrument widely and attempt to address the question of whether it is even possible to instill transferable quantitative skills over the course of a single semester of science instruction.

(c) **My good students will be bored.** Although on average the quantitative skills of nonmajor college students are poor, the distribution is wide, as evidenced in Figure 7.1, which shows the histogram of scores on the Fall, 2013 QuaRCS, which was administered to 518 students in five general education astronomy courses.

While anecdotal, we often turn to the following quote from a very gifted middle school student who we worked with through Astronomy Camp to combat the belief that emphasizing simple mathematics will disadvantage more skilled students:

"I have always heard a lot about science and math being related, but I have never actually used them together. I have finished an entire year of algebra, but there has never been any science in it. The same holds true in my science class. Frankly, I have never seen any connection between the two of them.

Similar statements have been made by other gifted junior high and high school campers attending out camps, and we suspect that even gifted students have rarely seen mathematics applied in science contexts. If this is the case, then they are unlikely to be bored when rich numerical examples are used in science courses.

Furthermore, much of the dialogue in the science education community, particularly in the context of college-level general education science courses, is focused on the instillation of "science literacy" (Rutherford & Ahlgren 1991; DeBoer 2000). The skills that comprise science literacy, such as the ability to interpret and understand scientific information in the media, are often quantitative in nature. From graphs and tables to percentages and error bars, the information reported to laypeople about science in the media is awash in the numerical. Thus, we would argue that numeracy and science literacy are inseparable.

7.2.2 Why General Education College Science Courses?

We have focused our instrument development efforts on a very specific student population: those enrolled in general education introductory college science courses. Notwithstanding the fact that we ourselves are college science educators, this population was chosen for several specific reasons. First, the goal of this project is ultimately to answer whether a difference can be made in student skills at the college level or whether the concern of many educators, that by the time students have reached college it is "too late" to learn basic quantitative skills, is warranted.

As for the focus on science, we suspected that for some of our students, a general education science course is the final quantitative course of their formal education. With the intention of probing this further, we added several questions to the Fall 2014 version of the assessment to better understand why students chose to take the course and whether they intended to take any further

Fall 2013 Pre Semester Assessment Scores (N=518) **Number of Students** Fall 2013 QuARCS Score

Figure 7.1: Distribution of scores (out of 25 total possible) on the 2013 Pre-Semester Assessment. The range of student abilities is very wide, and the average ability level is low. Mean = 14.8 (59%), Median=15 (60%), Standard Deviation=4.8 (19%).

science and math courses in college. In our Fall 2014 cohort of 437 students, 18.5% (N=81) reported that they did not plan to take another science course, and 36.4% (N=159) planned on taking only one more ³. When asked how many math courses they intend to take in college, 10.5% (N=29 ⁴) reported that they did not intent to take any math courses, and 23.6% (N=65) intend to take just one mathematics course. When asked why they chose to take the course, 67.1% (N=294) indicated that it was to fulfill a general education requirement. As we suspected, a non-negligible portion of the students taking these courses are expecting this to be the final STEM course of their college careers.

Even if students do take more quantitative courses than they anticipate (or, more likely, if they do not anticipate that courses in disciplines such as Business and Social Science will require quantitative skills), then an emphasis on numerical skills in a general education science course may still serve them well. In particular, since the majority of students in these courses are Freshmen or Sophomores ($\sim 50\%$ Freshmen and $\sim 30\%$ Sophomores according to our data), a quantitative emphasis in general education science has the potential to set students on a path toward making quantitative reasoning a part of their "academic toolbox" for the remainder of their college career.

Many universities and colleges require that their students take at least one science course before graduating, so these courses represent a broad cross-section of the college population. Students pursuing STEM majors are less likely to enroll in general education courses. In fact, in our Fall 2013 and Fall 2014 cohorts (N=981), only 4.9% (N=48) reported that they would major in science and 6.0% (N=59) in

³We suspect that this is due, at least in part, to the requirement at the University of Arizona that students take two science courses. In schools without a science requirement or where only one science course is required, students reporting that they do not plan to take any further science courses may be much higher.

⁴this question did not appear on Fall 2014 v1 of the assessment, and was only present for the 275 students who took v2

engineering, math or computer science.

In contrast, 31.9% (N=313) of the students reported that they would major in business and 21.0% (N=206) in social science. Although these disciplines certainly can and should emphasize numerical reasoning skills, students are much less likely to perceive them as mathematical. We present evidence for this claim in table 7.1 below, which shows the number of students who reported choosing their major "to avoid writing" or "to avoid math". These statistics are reported for the entire cohort, as well as divided according to STEM and non-STEM majors. Whereas STEM students are \sim 5 times more likely to report that they chose their major to avoid writing, these statistics are reversed among non-STEM majors, who are \sim 4 times more likely to chose their major to avoid math.

Table 7.1: Reason for Choosing Major by Major

	to avoid writing	to avoid math
All Students (N=981)	52 (5.3%)	128 (13.0%)
STEM majors (N=103)	19 (18.4%)	4 (3.9%)
Non-STEM majors (N=899)	32 (3.6%)	128 (14.2%)
Education majors (N=51)	0	10 (19.6%)

Additionally, 5% (N=51) of the students in our sample are future educators (education majors), and national studies of general education science courses suggest that this proportion is often much higher (\sim 40%, Lawrenz et al. 2005). This is another important population to reach, as poor attitudes of teachers toward mathematics has the potential to carry over to their students and perpetuate the problem of innumeracy (Beilock et al. 2010; Jackson & Leffingwell 1999).

7.2.3 Why Astronomy 101?

We have further chosen to focus on "Astronomy 101" and equivalent courses for our validation sample. As the questions in the instrument, described in detail in Section 7.3.2, are not couched in science contexts, but rather "real life" contexts, we believe that the instrument could be used with any similar population of students. In particular, we would encourage its use in other popular general education science courses such as Geology and Biology. The instrument may also be valid for assessing other populations, such as students in courses for science majors and students in mathematics or numeracy courses, however we have not yet investigated its validity for these populations.

Although it is a single discipline, there is reason to believe that Astronomy 101 students are a key population for numeracy advocacy. Data compiled by the American Institute of Physics (2010) reveal that "introductory astronomy enrollments have remained in the 180,000-190,000 range since 2004 (Nicholson & Mulvey 2010). This figure does not include community colleges, where "an estimated 100,000 students take Astronomy 101 in departments not covered by the AIP survey (Fraknoi 2001).

According to a Bureau of Labor Statistics report entitled College Enrollment and Work Activity of 2010 High School Graduates, in Fall 2010, 2.2 million people were enrolled in college in the U.S. This means that at least 10% of college students eventually pass through the door of an Astronomy 101 course in college, and this number could be as high as 50% depending on assumptions made about repeat enrollments and the number of years that students are in college. Regardless of the assumptions made, there is no arguing that a large percentage (>10%) of college-educated adults have taken an introductory astronomy course. Therefore pedagogical changes related to emphasis on numeracy and numerical skills, even if applied to this population alone, have the potential to reach a large portion of the college educated population.

7.3 Assessment Development

The QuaRCS was developed over the course of seven semesters between 2010 and 2014 and has been administered to over 2000 students in more than twenty courses, all of which were first or second tier general education Astronomy courses. The majority of the courses were large (100+) lecture courses at a large research university in the Southwest, however six of the courses were at a minority-serving community college, also in the Southwest. Table 6.2 describes changes made to the instrument between administrations.

7.3.1 Skill Selection

There are numerous and sometimes conflicting opinions regarding what specific skills constitute numeracy, therefore we began the design of our instrument by selecting a subset of skills for assessment. An initial list of potential skills was drafted by a panel of five science educators. However, we suspected that opinions regarding which numerical skills were most important in science and in "real life" would vary from instructor to instructor, as well as among disciplines, nor could we possibly design a short assessment capable of evaluating all of these skills. To this end, we designed a survey for math and science instructors in order to validate our choice of skill set. The survey asked these educators to rank a number of numerical skills according to their importance in various contexts, and to assess their students' current and desired level of proficiency at each skill.

The survey was intentionally populated with a mixture of (a) skills that appear frequently in the numeracy literature (e.g. proportional reasoning, arithmetic), (b) skills that are frequently invoked in the context of science literacy (e.g. graph reading, estimation), (c) "traditional" science skills whose utility in introductory courses is debated (e.g. scientific notation, algebra) and (d) advanced

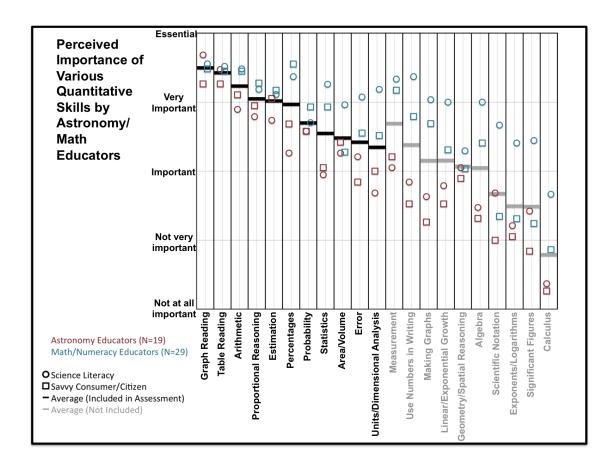


Figure 7.2: The perceived importance of various quantitative skills to science literacy (circles) and being a savvy consumer/citizen ("real life", squares) among Astronomy Educators (red symbols) and Math/Numeracy educators (blue symbols). The skills are ranked in order of their perceived importance from left to right, with those skills included in the final assessment in bold black (left) and those that were excluded in grey (right). Three skills measurement, using numbers in writing and making graphs were ranked higher than several skills included in the assessment, but were not added because of the difficulty in assessing them in multiple choice format. Average ranking for all contexts are shown as horizontal bars. Histograms showing the full breakdown of educator responses by discipline and situation are given in Appendix D

numerical skills that scientists use as a matter of course (e.g. calculus, exponents, logarithms).

The survey was pilot tested in paper form at two science education workshops in 2013 (N=34, 42), and the question wording and choice of skills were modified based on analysis of these pilot surveys. The final survey was administered online to several groups of both Astronomy (N=19) and Mathematics/Numeracy (N=29) educators. The list of skills and statistics regarding their perceived importance in various contexts are shown for both math and science educators in Figure 7.2.

Several interesting trends emerge from this survey data. First, mathematics educators consistently rank all skills as more important than science educators in every context, including science literacy. Second, there are just a handful of skills on the list that both math and science educators rank, on average, at or below the halfway mark ("Important") in our importance ranking. These skills are: (1) Calculus, (2) Exponents/Logarithms, (3) Significant Figures, and (4) Scientific Notation. None of these skills are included in our assessment, and we feel that this result validates the decision not to include them.

There are also a number of skills that math educators rank a full category or more higher than science educators on average. These are (for science literacy in decreasing order of the magnitude of the difference in ranking between math and science educators): (1) Algebra, (2) Using Numbers in Writing, (3)Unit Conversions, (4) Making Graphs and (5) Statistics. Regarding the importance of quantitative skills in life, the magnitude of the difference between math and science educators' opinions was smaller on average, with only two standouts: (1) Making Graphs, and (2) Using Numbers in Writing. Because all of the listed skills are numerical in nature, it is perhaps not surprising that mathematics educators

tend to rank them more highly. However, the fact that the differences in ranking are unique to certain contexts and vary according to the skill in question is interesting, and worthy of followup, which we will pursue in later publications.

There are five skills in the list that stand out as being the most highly and consistently ranked by both math and science educators both for their importance to science literacy and to being a savvy consumer/citizen. These are, in order of perceived importance: Graph Reading, Table Reading, Arithmetic, Proportional Reasoning and Estimation. All five of these skills are included in our assessment.

We have also chosen to draw from certain additional skills that lie in the "Very Important" region. The skills in this region, again in order of their perceived importance are: Percentages, Measurement, Probability, Statistics, Area/Volume, Error, Using Numbers in Writing, and Dimensional Analysis/Unit Conversions. We include all of these skills in the QuaRCS except for Measurement and Using Numbers in Writing. These are excluded because, though important and ranked highly, we felt that they would be more difficult to assess in a multiple choice format.

We thus selected the eleven most highly ranked and easily testable skills from this educator survey and focused on these in our assessment. We split unit conversions and dimensional analysis into two separate categories for the purposes of the assessment, for a total of twelve quantitative skills.

In the same educator survey, we also asked participants to rate (a) the degree of skill that their students have coming into their courses and (b) the skill level that their students would need in order to teach their course at the level that they would like. The differences between these real and desired ability levels are shown in Figure 7.3 On the whole, instructors in both disciplines would like for student skills to be higher in all areas, and believe that this would have a signif-

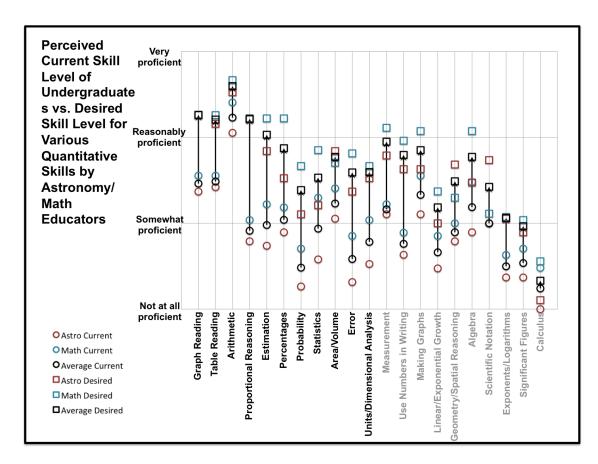


Figure 7.3: Desired (squares) and current (circles) student proficiency levels for various quantitative skills among math (blue) and science (red) educators.

icant effect on their course structure. Science educator responses to the followup question "How would your class be different if your students came in with this level of proficiency" included:

- "We would spend more time discussing scientific implications"
- "We could explore not only deeper into topics, and a greater breadth of topics, but we could explore topics more relevant to their everyday and future lives, and ones that students find more interesting"
- "They would have a better grasp of how science is done, including what information to trust and what evidence is good evidence"
- "I would be better able to help them use equations as a guide to thinking"
- "The 'I'm not good at math and science' hump likely wouldn't exist and they wouldn't shut down before they even started"
- "More students would describe their nonmajor science class as 'useful'"

 The skills that we assess in our instrument, in order of the amount of improvement that educators believe is needed (from greatest to least), are: Proportional Reasoning, Estimation, Error Analysis, Probability, Percentages, Dimensional Analysis, Graph Reading, Table Reading, Statistics, Area/Volume and Arithmetic. Particularly notable among the skills that are not included in our instrument is the desired increase in proficiency in Using Numbers in Writing.

7.3.2 Question Wording

Questions were developed over the course of five semesters by a panel of five science educators. They were designed to reflect situations that students might reasonably encounter in the course of their daily lives (e.g. bills, cooking, election polls, home repair), and every effort was made to word them as concisely as possible to reduce cognitive load. The final question set is given in Table 7.4. An initial vetting of question wording and answer choices was completed on a volunteer basis by ten college-educated adults who were not involved in survey development. This resulted in a number of tweaks to the initial wording of questions for clarity.

Figure 7.4 shows histograms of student responses from two semesters (Fall 2013 and Fall 2014) to the questions: (a) "Overall, how difficult were the questions in this survey" and (b) "In your everyday life, how frequently do you encounter situations similar to the problems in this survey". It reveals that students generally find the questions to be of easy to moderate difficulty (86% find it of moderate difficulty or easier), and that they generally believe that the questions reflect situations similar to what they encounter in real life - if not daily, then at least weekly or monthly (15, 29 and 29% respectively, 73% total).

All quantitative questions were piloted in both multiple choice and "free response" forms in large Astronomy 101 courses, and results of the free response administrations are described in detail in Section 7.4.1.

7.3.3 Development of Demographic and Attitudinal Questions

In addition to the 25 quantitative questions in the assessment, we also ask students a series of 25 low cognitive load demographic and attitudinal questions. The purpose of demographic questions is to assess whether variables such as age, race/ethnicity, gender, major and mathematical background are correlated with student performance on the assessment. Ba-

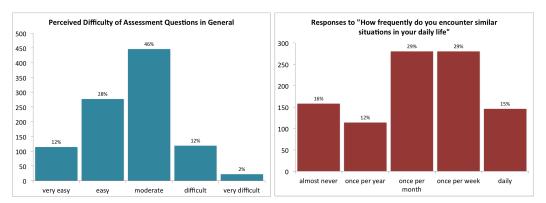


Figure 7.4: Histograms of student responses regarding the perceived difficulty of questions in the QuaRCS (left) and how frequently they feel they encounter similar situations in their daily lives (right). These figures represent a conglomeration of student responses from the Fall 2013 and Fall 2014 administrations.

sic demographic questions changed very little during development of the instrument except for minor wording changes to reflect frequent write-in responses in the "other" category and the addition of questions asking respondents to specify race/ethnicity and disability status in the Fall, 2014 administration. These questions were added because the literature suggests that these populations are particularly underserved in the area of quantitative literacy (Raudenbush & Kasim 1998; Ogbu 1990; Chan & Dally 2001; Rousselle & Noël 2007).

We moved all demographic questions to the end of the assessment in Fall, 2013 based on a concern of inducing stereotype threat (Steele & Aronson 1995) by responding to demographic questions before the quantitative questions, particularly in regard to the belief that women are not as adept at mathematics as men (Spencer et al. 1999).

A robust measure of student attitudes and changes within them is particularly important to the key questions we wish to address with this study, and so attitudinal questions have been included since the initial QuaRCS administration and have been heavily analyzed and refined. Questions in this category probe student attitudes toward mathematics, their view of the utility of numerical skills in both science and in their daily lives, and their confidence and satisfaction with their current level of numerical ability. Attitudinal questions are asked at each assessment, however in the post-assessment, some of the demographic questions are foregone in lieu of questions that ask students to reflect on the types of quantitative skills that were emphasized in their course, how these skills were emphasized (homework, lecture, discussion, etc.), and whether students feel that they improved their skills in given areas over the course of the semester. These questions will be used to assess student perspectives on pedagogical techniques and quantitative emphases in the broader Phase II implementation of the QuaRCS.

We added and subtracted both demographic and attitudinal questions throughout instrument development based on data-driven questions that arose during analyses of earlier versions of the instrument and during our review of the literature on quantitative literacy. For example, we added a question asking students to specify the reason(s) why they chose or will choose a major. We were particularly interested in this case in probing whether students choose majors based on an aversion to mathematics or writing, with interesting preliminary results as presented in Table 7.1.

Particular effort was devoted to the development of a question asking students to choose adjectives to describe numerical skills from among a list of opposites: Useful and useless, interesting and boring, easy and hard, and fun and scary. We asked this question in the same way (as a "choose all that apply" question) through the Fall, 2013 assessment. In Fall of 2014, however, we hit upon the idea of forcing students to take a stand on each of these scales between two opposite adjectives. In particular, we noted that students generally chose just two or three adjectives and were particularly unlikely to take a stand on the fun vs. scary scale, which is particularly interesting in the context of math anxiety. We investigated two alternatives by adding them to free-response question sheets that students were already completing as part of survey development.

On the same day, half of the students in a single large lecture class received a version of the question that allowed them to choose a neutral option (3) on a five point scale, and the other half received a version that did not allow for a neutral option (four point scale). As shown in Figure 7.5, it appears that when forced to choose between the two adjectives, the 20-40% of students who would select a neutral option will move in both directions, though in each case they are more likely to move one way than the other. All of these shifts are toward the positive adjective (Useful, Interesting, Easy) except in the case of the Fun vs. Scary scale, where students tend to move toward Scary when the neutral option is eliminated. This is also the scale where the largest proportion of students will choose the neutral option (41%) when available. We interpret both of these results as subtle effects of math anxiety leading to student reluctance to take a stand on this scale, as well as their tendency when pressed to move in the direction of "scary"

The current bank of demographic and attitude questions, and the statistics of student responses on the Fall 2013 and Fall 2014 versions of the instrument are given in Appendix B. Connections between these variables and student performance on the assessment will be investigated in detail in

Phase II.

7.3.4 Inclusion of Confidence Rankings

We were also interested in understanding whether students are aware of their own numerical deficits or whether some mathematical misconceptions are so deeply rooted that students make them without being aware of their mistake. In light of this, we chose to ask students to rate their confidence in their answer to each quantitative question. An initial analysis of students' awareness of their own quantitative abilities was done on the 2011 cohort of students, and included both pre and post assessments, for a total of 523 completed assessments and more than 11,000 individual questions. These results are shown in Figure 7.6.

We found that students who answered questions correctly follow an expected distribution in confidence levels, with the majority either confident (27%) or very confident (56%) in their answers. Somewhat surprisingly, however, approximately half of respondents fall into the same confidence regime (21% very confident, 30% confident) when they answer *incorrectly* as well. This suggests that the instrument is capable of measuring student awareness (or lack thereof) in their own quantitative abilities, and suggests that many may be "unconsciously incompotent". These very interesting results will be expanded upon in Phase II of the QuaRCS study.

7.3.5 Format

After piloting the instrument in paper form for two semesters (Fall 2010 and Spring 2011), we elected to move to an online format. This choice was motivated by several factors. First, instructors who we attempted to recruit were

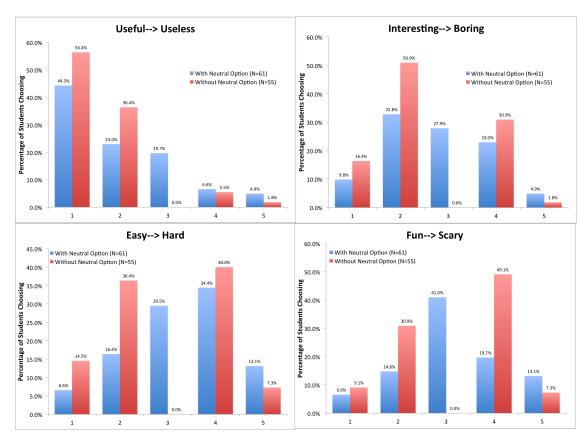
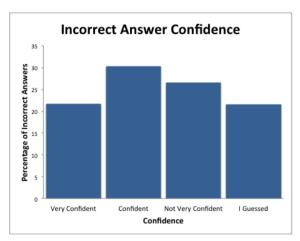


Figure 7.5: Normalized distribution of student responses placing "doing mathematics" on four scaled between two opposite adjectives. Half of the students were provided with a "neutral" option halfway in between the extremes (blue histograms), and the other half were provided a four point scale that forced them to prioritize one extreme over the other (red histograms). When forced to commit to one side or the other, students tended to move toward the positive option ("easy", "useful" or "interesting") except in the case of fun vs. scary, where they were more likely to choose scary.



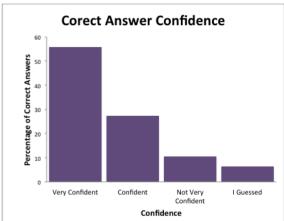


Figure 7.6: Histograms of student confidence in correct answers (right) and in incorrect answers (left). More than half of students answering questions incorrectly report that they are "confident" or "very confident" in their answers.

reluctant to devote class time (~30min) to the instrument, which is required for a paper administration. Second, because the distribution in student abilities is so varied, the amount of time that they take to complete the QuaRCS is also widely varied (from <10minutes to more than 1 hour). If the assessment was given at the beginning of the class period, many students who finished quickly would sit idle until a sufficient portion of the slower finishers had completed it, and the instructor chose to collect the surveys. If the assessment was given at the end of the class period, desire to leave became a motivator for students to hurry through the assessment, and slower assessment takers were not able to finish. A final obstacle was the overhead in entering data into a spreadsheet by hand or reading in scantron forms for analysis.

In light of these obstacles, we elected to move to an online format. There are a number of advantages to this method of administration. Instructors may assign the instrument for completion outside of class time, and we found that this made them much more willing to implement it. Students are also able to take as much time as needed to complete the assessment, and the amount of time that they take to complete it is recorded by the system and can be analyzed.

Online administration is not without disadvantages, however. Direct oversight is foregone, so we are unable to prevent collaboration and usage of outside resources (internet, mathematics textbooks). The long tail of the time distribution shown in Figure 7.15 also makes it clear that some students are not actively engaging with the instrument throughout. They may be taking it in more than one sitting or simply leaving it open on their computers for long periods of time before beginning. As the data collection software (DatStat Illume) has no way of measuring active engagement, we can't precisely quantify these effects. Just 13% of students on the Fall 2013 pre assessment took longer than 60 minutes to complete it, and this number dropped with each subsequent administration during that semester. The median number of correct responses among these students taking longer than 60 minutes on the QuaRCS was just one point higher (16, 64%) than among the entire sample, perhaps due to the fact that they are $\sim 15\%$ more likely to self-report devoting effort to the assessment (a description of how we measure effort is included in Section 7.3.6) if they fall into this tail of the time distribution.

Whereas we were originally able to forbid the use of calculators on the assessment, the online format does not allow us to control this. We elected therefore to include the following statement in the introduction to the assessment: "These questions were designed to be answerable without a cal-

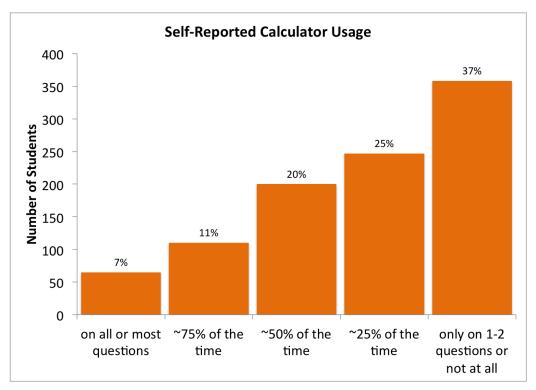


Figure 7.7: A histogram of student responses from the Fall 2013 and Fall 2014 administrations reflecting the self-reported usage of calculators. More than half of students report that they use a calculator on 25% or fewer of the questions.

culator, but you are welcome to use one if you choose". We also ask students to self-report their usage after taking the quantitative questions on the assessment. The distribution of responses to the caclculator usage question is shown in Figure 7.7 below. Although it varies significantly, a majority eschew calculators on 75% or more of the questions (37% using one on just 1-2 questions or not at all, and 25% using a calculator 25% of the time or less).

Given the relatively low proportions of (a) students clearly dividing their attentions between the instrument and other activities and (b) heavy calculator usage on the assessment, we believe that the benefits of the online format outweigh the problems, and will continue to use this format in the

future. As further validation of this choice, we note that the average score on the QuaRCS dropped by just $\sim\!\!3\%$ between the Spring 2011 paper format assessment and the Fall 2011 online version, although the majority of the questions were substantively unchanged. ⁵

7.3.6 Self-Reporting of Effort

The final multiple choice question on the QuaRCS from the Fall, 2013 assessment onward has been as follows.

Knowing that this survey is being used for research to try to improve courses like yours and that your answer to this question will not be shared with your instructor, please honestly describe the amount of effort that you put into this survey.

- (a) I just clicked through and chose randomly to get the participation credit
- (b) I didnt try very hard
- (c) I tried for a while and then got bored
- (d) I tried pretty hard
- (e) I tried my best on most of the questions

We added this question to the assessment primarily because we suspected that some of our early results were being heavily affected by increasing student apathy toward the end of the semester, and analysis to this effect is described in Section 7.5.1.

The distribution of student responses to this question is shown in Figure 7.8 and correlations between self-reported effort and various other assessment

⁵We note, however, that individual students were asked only 10 questions each from a bank of 22 total questions on the Spring 2011 assessment, while they were asked to answer all 22 questions for the Fall 2011 assessment. In the end, this longer assessment may be more responsible for the small drop in average score than the change in format

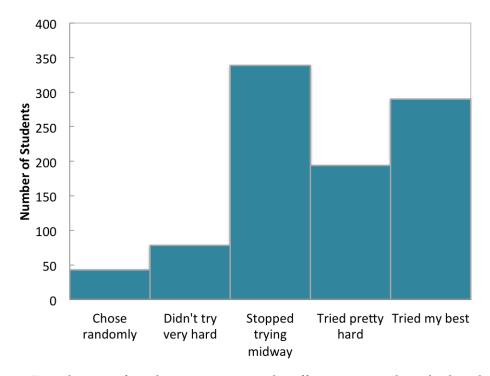


Figure 7.8: Distribution of student responses to the effort question described in this section. These responses (N=944) include all three administrations (Fall 2013 and Fall 2014 v1 and v2) in which the question was asked. Variations between administrations in students choosing each effort ranking are <10% between administrations.

variables is addressed in detail in Section 7.5.1. Throughout the remainder of this paper, we will refer to the top two effort categories ("I tried pretty hard" and "I tried my best on most of the questions") as students who "devote effort" to the assessment. We will refer to students in the third effort category ("I tried for a while and then got bored") as students whose effort waned midway through the assessment.

7.3.7 Assessment Length

Measuring twelve quantitative skills (selected as described in Section 7.3.1) in a single assessment is an ambitious endeavor, and test fatigue (AERA

2005, PISA 2000) was certainly a concern. We took advantage of overlap between skill categories, and each of our questions assesses multiple skills simultaneously, though in varying combinations. The final question stems, their categorization and various statistics about difficulty, discrimination and reliability are shown in Table 7.4. We aimed to test each individual skill through at least four separate questions to mitigate the effects of question wording, placement within the assessment, and pairing with other skills on measurement of student knowledge in a particular area.

Each skill is also assessed with questions of varying difficulty ranging from easy to moderate to difficult defined as difficulty values of <0.33%, 0.33% <X<0.57% and >0.57% respectively. These cutoffs correspond to the top, middle and bottom third of overall percentage correct among items on the assessment, so these rankings are only relevant relative to one another. Every effort was made to include at least one question of each difficulty level in the assessment, however certain skills proved more difficult for students than others and this was not possible in the case of probability and error, for which we were unable to write a question that students answered correctly more than 58%/37% of the time respectively. Statistics regarding the distribution of difficulty by skill are listed in Table 7.2.

We quantified the effect of fatigue on assessment takers in two ways, both involving analysis of the effort question outlined in Section 7.3.6.

To study the effect of assessment length and fatigue on student responses, we administered a reverse-ordered version to half of the students in one large lecture class ("Class 1") in the Fall 2013 semester. The pre, mid and post-semester instruments were completed by 115, 92 and 24 students respectively, and approximately half of these students were assigned the re-

Table 7.2: Number of Questions by Skill and Difficulty

	Number of	Number of	Number of	
Skill	Easy	Intermediate	Difficult	
SKIII	(PC>84%)	(70% <pc<84%)< td=""><td>(PC<70%)</td></pc<84%)<>	(PC<70%)	
	Questions	Questions	Questions	
Graph Reading	1	1	3	
Table Reading	1	1	1	
Arithmetic	6	4	1	
Proportional Reasoning	2	1	2	
Estimation	1	2	1	
Percentages	1	1	3	
Probability	0	1	3	
Statistics	1	2	2	
Area/Volume	1	4	1	
Error	0	0	4	
Unit Conversions	3	1	2	
Dimensional Analysis	5	3	2	

verse ordered instrument in each case (N=56, 43, 11 ⁶).

As a first measure of the effects of test fatigue, we compared the results for students in Class 1 taking the forward and reverse ordered assessments. We selected only those students who completed the instrument and reported that they either "tried my best on most of the questions" or "tried pretty hard" on the assessment. We then completed an analysis of variance test on student scores for the "first" 11 questions, the middle 3 questions and the "last" 11 questions in the assessment. Here we place "first" and "last" in quotations because the students taking the reverse-ordered assessment answered these questions in the opposite order. The middle three questions appeared in the same place for all students taking the assessment.

We found that students in this high effort group were equally likely to answer a question correctly whether it appeared in the first or second half of the assessment. This leads us to conclude that fatigue is not likely to figure prominently into item difficulties *among students who expend effort on the assessment*.

We note, however, that we've excluded all students who reported a lower degree of effort than "I tried pretty hard". Particularly relevant in the determination of optimum assessment length was the effort choice "I tried for a while, and then got bored" (N_{pre} =15/15 and N_{mid} =11/9 for the forward/reverse ordered assessment).

Among these students student who self-report that their effort waned mid-

⁶The small number of post-semester completions for this course is due to the instrument having been assigned for extra credit rather than being required during this administration. These data were neglected in this analysis due to the relatively small number of completions and likelihood that the subset of students who elected to complete the post instrument may not be the same as the original population.

way through the assessment, the mean score is lower for the same set of questions when it was encountered later in the assessment, and these differences are statistically significant in several cases. This is little more than a validation that students have correctly assessed their own effort, as indeed they themselves reported that their effort waned toward the end of the assessment.

More relevant in the determination of the appropriate instrument length is the relative *number* of students in this group (N_{pre} =30, N_{mid} =20) as a proportion of the entire class (N_{pre} =115, N_{mid} =92). For both the pre and mid semester instruments, students who "gave up" midway through the assessment amount to roughly one-quarter of the population. On some level, this means that approximately one in four students find the instrument too long to maintain effort throughout.

Were we to lengthen the instrument to more than 25 questions in order to assess more skills, it is reasonable to expect this proportion to increase. During Phase II of this project, we will pilot a shortened version of our instrument drawing only from the most highly ranked skills from our educator survey in an effort to reduce the proportion of students in this category. This analysis reveals that in its full form, the assessment reported on in this paper is considered too long by approximately 25% of students who take it. Those who do manage to maintain effort throughout, however, are no more or less likely to get a particular question right based on its placement within the exam.

7.3.8 Matched Data

Both the Spring, 2012 and Fall 2013 QuaRCS assessments included pre and post administrations (plus a mid semester administration in Fall, 2013, as will be described in section 7.5.1). Pre-post data will be the norm in future semesters and will be essential for studying the effectiveness of various interventions during the broader Phase II QuaRCS study.

In order to match data, we asked students to provide names at each administration and used these to match pre, mid and post semester results automatically. We then combed through similar names by hand, correcting capitalization, spacing and spelling differences between administrations for responses that were clearly the same person (e.g. MollyT homas \rightarrow Molly Thomas, Susan B. Anthony \rightarrow Susan Anthony, Jim Mcleod \rightarrow Jim McLeod, Tom Smith \rightarrow Thomas Smith). Per our IRB protocol, names and corresponding scores were kept confidential and instructors were only provided lists of students who completed the instrument, not individual results.

The Fall 2014 implementation of the assessment was focused on refinement of the questions based on the analyses described in Section 7.4, including pre-post question comparisons as described in section 7.4.3. During this semester, the instrument was administered twice (to different groups of students). The first administration (v1) took place \sim 6 weeks into the semester (so later than a traditional "pre" semester administration), in order to allow time for question rewordings to be vetted in free response form before multiple choice administration. V1 was administered to only two classes, and analysis of the results resulted in the identification of several additional problematic questions.

For this reason, another round of question revisions, including free response implementation, was completed and a revised multiple choice assessment was given to five classes of students 11-13 weeks into a 16 week semester. This version of the assessment (Fall 2014 V2) is used for the majority of our final item analysis. Because it was administered so late in the semester, no post data were taken in Fall 2014, so all pre-post analysis in this paper is done ont he next most recent assessment (Fall 2013).

7.4 Item Analysis

7.4.1 Free Response Items

All questions in the QuaRCS have been vetted in "free response" form. Free response questions were formatted as open-ended versions of the multiple choice question stems, with students doing no more than two problems on a given day and submitting their answers for participation credit. This was done over the course of three semesters (Spring 2012, Fall 2013 and Fall 2014) as the questions were drafted and revised. The minimum number of "free response" forms received per question is 31, the maximum is 106 and the average is 61 per question.

The free response questions served several purposes for question and instrument validation. First, asking the questions in an open-ended format allowed us to generate compelling distractors for the multiple choice assessment based on common mathematical misconceptions and mistakes. In some cases, these mistakes and misconceptions were anticipated (e.g. students calculating perimeter instead of area), and in others, they were unexpected. For example, in a question regarding a poll reporting $60\pm5\%$ of

votes for the victor and a sample size of 400, 7 of 72 (\sim 10%) students calculated that this corresponded to 240 \pm 12 votes (0.05 \times 0.6 \times 400=12) rather than 240 \pm 20 votes (0.05 \times 400=20). We were thus able to add a more authentic distractor (240 \pm 12) to the assessment based on a misconception that had not originally occurred to us.

Careful analysis of student responses also revealed in several cases instances of words or phrasing that were unclear, imprecise, or could be misinterpreted. For example, when asked to interpret a table compiling numbers for various types of injuries at a ski resort by ability level (novice, intermediate, expert), the phrasing "if you were a random skier visiting this resort", led 5 of 53 students to chose the ability level most appropriate to them personally, rather than for a generic randomly-selected skier. To clarify, the question was reworded to ask "who was the most likely" to sustain a severe injury.

Where only minor wording changes were required and a relatively small number of people misinterpreted the wording in the first place, the minor revisions were not retested. Any substantial wording changes, as well as minor wording changes that were misinterpreted by more than 5% of respondents, however, were given again in free response form to ensure that the problem was corrected.

A final benefit of the free response testing was ensuring in each case the students were reaching the correct answer through legitimate quantitative reasoning and not stumbling upon it through error. For several questions, open-ended responses revealed that students could arrive at the correct answer via incorrect reasoning. In all cases where more than one student in the sample arrived at the correct answer in this way, numbers or wording

were tweaked to correct the problem.

For example, in the table reading question bank, one question originally asked students to compare the likelihood of minor injuries among novice (N=374 injuries) and intermediate (N=97 injuries) skiers. Students were also given the total numbers of each type of skier (12,016 and 9,640 respectively) and were expected to use this information in computing relative probabilities. As expected, many students ignored the relative numbers of each type of skier and simply divided 374 by 97. In using long division to divide these numbers, several students (N=3 of 106) stopped after the first digit (3) even though the full answer is much closer to 4 (\sim 400/100). In this case, 3 was the correct answer ($374/12016\approx400/12000\approx3.3\%$ and $97/9640\approx100/10000\approx1\%$), though these students arrived at it incorrectly. To prevent this, we simply tweaked the numbers slightly to 384 and 96 to make them more easily divisible.

In all, thirty five questions were developed and tested as free-response. Of these, ten were rejected because of inability to clarify the wording sufficiently, topical redundancy or low discrimination and four were substantially revised.

7.4.2 Expert Analysis

We also utilized the input of outside experts in validating the instrument. The Fall, 2013 version of the assessment was administered to a group of thirty four experts a mixture of astronomy (N=17) and numeracy (N=17) educators - who completed it as though they were students. Their mean score on the assessment among experts was 22.8 (90%) and the standard deviation was 1.9 (9%). This distribution is markedly different from that of

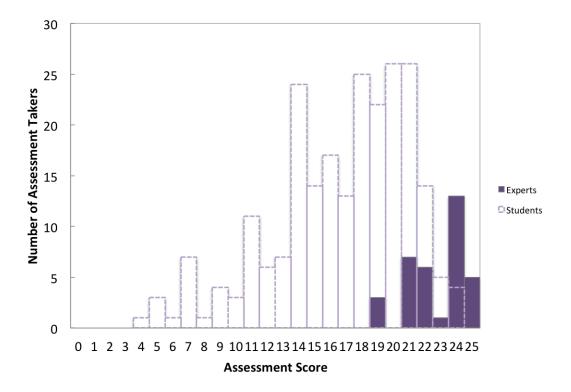


Figure 7.9: Histogram of scores among "experts" taking the QuaRCS (N=34, filled purple). A similar histogram for students (N=251, dashed purple) taking the same version of the assessment. Only the students who devoted effort to the assessment were selected here, because this group has the highest mean score among students. Even so, the two distributions are discrepant at the 0.000 significance level.

students on the same assessment. Even if we keep only those students who survive the effort and time filters (the student distribution with the highest mean score, N=251), the mean is only 17.1 (68%) and the standard deviation is 4.1 (16%). A comparison of these two distributions is shown in Figure 7.9. Using an independent samples T-test, we find that these populations are different at the 0.000 significance level. Thus, the instrument is capable of distinguishing experts (the "quantitatively literate") from novices.

Administration of the instrument to this group of experts also helped us

to identify three problematic questions. A comparison of student and expert correct response rates for individual questions is shown in Figure 7.10. In two cases, fewer than half of experts answered the question correctly, whereas all of the other nonproblematic questions on the assessment ranged from 82% to 100% (N=8) correct. By interviewing experts, it became clear that the problem in one of the questions was the term "margin of error", which has a more specific meaning in the mind of statistics experts than what we were intending to convey. The problem with the second question was not easily identified, and so we simply removed it from the assessment. The third problematic question was one that experts answered correctly 82% of the time, while students answered it correctly 86% of the time. In this case, we determined that the answer choices were too closely spaced for a non-calculator user to distinguish quickly between the correct and incorrect response. As the question was intended to assess estimation, we simply spaced the distractors farther from the correct answer to encourage this further.

If the three problematic items are removed for comparison of experts and students, then the average score on the assessment among experts becomes 20.9/22 (95%) and the standard deviation falls to 1.3 (6%). Among students, the average on the same 22 questions is 15.3 (70%), and the standard deviation is 4.2 (19%). The expert sample therefore shows that the quantitatively literate (experts) score significantly higher on the assessment, and the instrument is therefore a robust measure of quantitative ability.

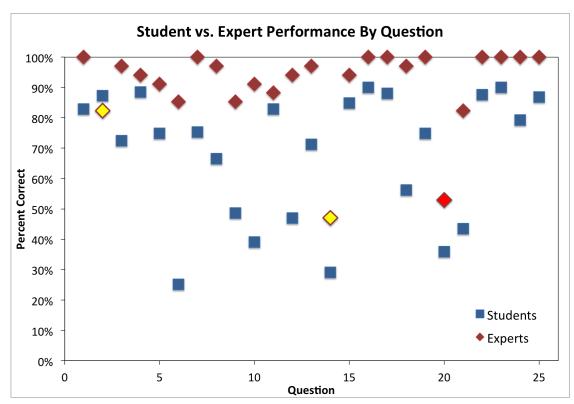


Figure 7.10: Average correct response rates for students (blue squares, N=251) and experts (red diamonds, N=34) on the Fall 2013 pre assessment. Experts perform substantially better than students, indicating that the test is effective at measuring quantitative ability. Three problematic questions were identified through this analysis. One (item 20, red diamond) was removed from the assessment, and the other two (items 2 and 14, yellow diamonds) were substantially revised for clarity, as described in the text.

7.4.3 Item Statistics

Throughout development, we evaluated the quality of the instrument using Classical Test Theory (CTT), for which it is assumed that a student's score on the assessment reflects their (unmeasurable) true score plus a random error (Allen & Yen 1979). The evolution of instrument statistics such as mean, standard deviation, and Cronbach's α (a measure of internal consistency of an assessment and a number between 0 and 1, where high values indicate high reliability) during development are reported in Table 7.3.

Table 7.3: Instrument Evolution

Form	n students	n questions	Mean (%)	Standard Deviation (%)	Cronbach's α
Spring 2011	68	10	66	22	0.692 7
Fall 2011	190	22	63	19	0.767
Spring 2012	574	22	58	18	0.749
Fall 2013	518	25	59	19	0.801
Fall 2013 ⁸	343	25	63	20	0.825
Fall 2014	276	25	55	23	0.865
Fall 2014 ⁹	166	25	62	24	0.885

Evolution of instrument statistics from Spring 2011 through Fall 2014

The mean score on all versions of the assessment was in the 55-65% range and the standard deviation was very high (15-20%), indicating a low overall ability and a wide range of abilities for the population of students assessed. Cronbach's alpha has remained at or above 0.7, considered in the acceptable range for a multiple choice assessment (George & Mallery 2003), since initial administration, and reaches 0.89 in the final version of the assessment.

Reported statistics include "item difficulty" for the questions in both free response and multiple choice format. This value is computed by subtracting the proportion of correct answers from 1, so that items with high difficulty are answered correctly by proportionally fewer students. In general, difficulty values for multiple choice assessments are considered acceptable when they lie between 0.2 and 0.8 (indicating 80% and 20% of students answering correctly, respectively; Bardar et al. 2006; Schlingman et al. 2012), however we have chosen to include several questions with difficulty values below 0.2. This was a conscious decision, as we felt it important to include several very easy questions as a baseline and to help alleviate math anxiety. Discrimination values for each question are also given in the table. Conventionally accepted values for this parameter lie between 0.3 and 0.7 (Bardar et al. 2006). The majority of items on our assessment lie in this range, however two of the five graph reading question have values in the 0.2-0.3 range. It is not entirely clear why this is the case. Perhaps graph reading is different enough from the other skills assessed on the QuaRCS that student scores in this are are not correlated well with their overall score, or perhaps these two questions are problematic in some way that our initial analysis

⁸Without students who reported that they quit midway through the instrument

⁹see footnote 8

Table 7.4: Item Statistics

Question Text	Free Response Difficulty	Multiple Choice Difficulty Unfil- tered/Filtered	Discrimination Unfil- tered/Filtered	Skills Addressed
You have a rectangular fish tank that's 10 inches tall, 20 inches wide, and 15 inches		(N=276/166)		
deep. If the volume of one gallon of water is 231 cubic inches, then how many gallons	0.30 (N=56)	0.32/0.25	0.40/0.44	AV, SA, UC
are required to fill the tank?				
A college that typically has 50,000 students experiences an increase in enrollment to 55,000 students. By what percentage did enrollment increase?	0.23 (N=93)	0.37/0.30	0.43/0.44	PC, SA
According to the graph, what was the approximate population of City X in 1980?	0.09 (N=59)	0.16/0.13	0.29/0.31	GR
If the current population growth rate continues, which is the best estimate for the population of City X in the year 2050?	0.37 (N=102)	0.56/0.54	0.22/0.21	GR, ES
Based on this graph, compare the population growth rates (i.e. increase in number of people per year) before and after 1970	0.64 (N=88)	0.68/0.66	0.18/0.19	GR
Imagine you have already filled a measuring cup (like the one shown above) with the amount of peanut butter in the recipe and you want to add the correct amount of shortening on top of it. Which line on the measuring cup should you fill to with shortening?	0.17 (N=51)	0.28/0.27	0.39/0.50	PR, SA, DA

 $^{^{10}}$ the filtered dataset here represents the exclusion only of student who reported that they stopped trying midway through the assessment

If your measuring cup has ounces on the side instead of cups, which line should you	0.13 (N=35)	0.26/0.20	0.43/0.50	UC, SA
fill to when measuring the flour? There are 8 ounces in 1 cup.	V/1 0 F0			
	V1: 0.59			
You have only a Tablespoon measuring spoon. How much should you fill it to get	(N=33),	0.69/0.65	0.44/0.50	PR, UC
the correct amount of baking soda? There are 3 teaspoons in 1 tablespoon.	V2:0.31	0.09/0.03	0.44/0.30	T K, UC
	(N=39)			
How many total injuries (including deaths) were sustained at Resort Y during this	0.11 (N=55)	0.27/0.22	0.46/0.58	TR, SA, ST
time period?	0.11 (11-33)	0.27 / 0.22	0.40/0.50	1K, 5A, 51
YAZIn at a year the alternation of a way depole colored alting acceptaining an injury of any bind	v1: 0.37			DD TD EC
What were the chances of a randomly-selected skier sustaining an injury of any kind	(N=49), V2:	0.42/0.37	0.33/0.36	PB, TR, ES,
(minor, severe or death) while at Resort Y during this time period?	0.24 (N=50)			ST
What proportion of severely injured skiers at Resort Y during this time period were	, ,	0.50 /0.51	0.47/0.50	TD DD DD
intermediate skiers?	0.43 (N=79)	0.59/0.51	0.47/0.50	TR, PR, PB
The graph above shows the predicted viewership of three television shows in two	V1: 0.74			
cities based on a poll of a small number of residents in each city. The poll has a re-				GR, ER, PC,
ported error of 25%, shown as vertical error bars. Which of the following statements	(N=50), V2:	0.68/0.64	0.32/0.38	ST, PB
	0.56* (N=50)			31,10
about the predicted viewers of Show A is most accurate?				
Which of the following predictions can be made based on the information (including				
errors) shown in the graph? Prediction 1: In City 2, more people will watch Show	V1:0.79			CD ED CT
B than Show C Prediction 2: In City 1, Show C will have the smallest viewership	(N=42), V2:	0.74/0.67	0.32/0.38	GR, ER, ST,
Prediction 3: None of the three shows (A, B or C) will be equally popular in Cities 1	0.76* (N=50)			РВ
and 2				

Homeowner's Association requires that solar panels not be visible from the road. You decide to put solar panels on the roof of a shed in your backyard instead. The shed has a flat 5 foot by 5 foot roof. Complete the following sentence: To produce the same amount of power as your original design, you need to buy panels that produce times more power per unit area than your original panels. If you cover the shed with your original panels, how many more sheds would you have to put up in your backyard in order to fit the rest of the panels? Your cable bill is \$36 per month from January 1 through September 30 and then doubles to \$72 per month starting October 1. What is your average monthly bill over the course of the entire calendar year (January-December)? If you place \$10 under your mattress every day for the next 40 years, approximately how much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60% of the votes with 5% error)? ER, PC 10.11 (N=37) 10.32 (N=8) 10.41/0.32 0.42/0.44 AV, SA AV, SA 21. DA, PR, UC 10. DA	You purchased 100 square feet of solar panels for your roof. However, your local				
shed has a flat 5 foot by 5 foot roof. Complete the following sentence: To produce the same amount of power as your original design, you need to buy panels that produce times more power per unit area than your original panels. If you cover the shed with your original panels, how many more sheds would you have to put up in your backyard in order to fit the rest of the panels? Your cable bill is \$36 per month from January 1 through September 30 and then doubles to \$72 per month starting October 1. What is your average monthly bill over the course of the entire calendar year (January-December)? If you place \$10 under your mattress every day for the next 40 years, approximately how much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60 % of the votes with 5% error)? Several days later, the newspaper conducts another survey with 400 participants, how many people can report that they will vote for Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)?	Homeowner's Association requires that solar panels not be visible from the road.				
shed has a flat 5 foot by 5 foot roof. Complete the following sentence: To produce the same amount of power as your original design, you need to buy panels that produce times more power per unit area than your original panels. If you cover the shed with your original panels, how many more sheds would you have to put up in your backyard in order to fit the rest of the panels? Your cable bill is \$36 per month from January 1 through September 30 and then doubles to \$72 per month starting October 1. What is your average monthly bill over the course of the entire calendar year (January-December)? If you place \$10 under your mattress every day for the next 40 years, approximately how much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60 % of the votes. The newspaper estimates the error in this prediction to be 5%. If the newspaper repeats the survey with 400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? EVALUATE:	You decide to put solar panels on the roof of a shed in your backyard instead. The	0.40 (0.1. 50)	0.41/0.32	0.42/0.44	AV, SA
itimes more power per unit area than your original panels. If you cover the shed with your original panels, how many more sheds would you have to put up in your backyard in order to fit the rest of the panels? Your cable bill is \$36 per month from January 1 through September 30 and then doubles to \$72 per month starting October 1. What is your average monthly bill over the course of the entire calendar year (January-December)? If you place \$10 under your mattress every day for the next 40 years, approximately how much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60 % of the votes. The newspaper estimates the error in this prediction to be 5%. If the newspaper repeats the survey with 400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? A vill receive 60% of the votes with 5% error)? A vill receive 60% of the votes with 5% error)? A vill receive 60% of the votes with 5% error)?	shed has a flat 5 foot by 5 foot roof. Complete the following sentence: To produce the	0.18 (N=56)			
If you cover the shed with your original panels, how many more sheds would you have to put up in your backyard in order to fit the rest of the panels? Your cable bill is \$36 per month from January 1 through September 30 and then doubles to \$72 per month starting October 1. What is your average monthly bill over the course of the entire calendar year (January-December)? If you place \$10 under your mattress every day for the next 40 years, approximately how much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60 % of the votes with 5% error)? Several days later, the newspaper conducts another survey with the original prediction (that Candidate A will receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? A V. S. A. V.	same amount of power as your original design, you need to buy panels that produce				
Now much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60% of the votes with 5% error)? Str. DA, PR, UC, ES, D.32 (N=8) O.32 (N=8) O.39/0.36 O.39/0.37 O.39/0.47 Str. DA, PR, UC O.11 (N=37) O.32/0.23 O.39/0.59 O.39/0.59 O.39/0.59 O.39/0.50 O.39/0.47 Str. DA, PR, UC O.11 (N=37) O.32/0.23 O.44/0.50 UC, ES O.36 (N=72) O.36 (N=72) O.36 (N=72) O.36 (N=72) O.36 (N=72) O.37/0.59 O.39/0.47 O.39/0.4					
have to put up in your backyard in order to fit the rest of the panels? Your cable bill is \$36 per month from January 1 through September 30 and then doubles to \$72 per month starting October 1. What is your average monthly bill over the course of the entire calendar year (January-December)? If you place \$10 under your mattress every day for the next 40 years, approximately how much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60 % of the votes. The newspaper estimates the error in this prediction to be 5%. If the newspaper repeats the survey with 400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? ER, PC O.20 (N=82) O.63/0.52 O.63/0.52 O.53/0.58 ER, PC	If you cover the shed with your original panels, how many more sheds would you		0.54/0.42	0.56 /0.59	AMCA
doubles to \$72 per month starting October 1. What is your average monthly bill over the course of the entire calendar year (January-December)? If you place \$10 under your mattress every day for the next 40 years, approximately how much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60 % of the votes. The newspaper estimates the error in this prediction to be 5%. If the newspaper repeats the survey with 400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)?			0.34/0.42	0.36/0.36	AV, SA
doubles to \$72 per month starting October 1. What is your average monthly bill over the course of the entire calendar year (January-December)? If you place \$10 under your mattress every day for the next 40 years, approximately how much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60 % of the votes. The newspaper estimates the error in this prediction to be 5%. If the newspaper repeats the survey with 400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? UC, ES 0.39/0.36 0.39/0.36 0.39/0.47 UC UC, ES 0.30/0.59 0.60/0.59 ER, PC ER, PC ER, PC ER, PC	Your cable bill is \$36 per month from January 1 through September 30 and then				CT DA PP
the course of the entire calendar year (January-December)? If you place \$10 under your mattress every day for the next 40 years, approximately how much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60 % of the votes. The newspaper estimates the error in this prediction to be 5%. If the newspaper repeats the survey with 400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)?	doubles to \$72 per month starting October 1. What is your average monthly bill over	0.22 (N=68)	0.39/0.36	0.39/0.47	
If you place \$10 under your mattress every day for the next 40 years, approximately how much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60 % of the votes. The newspaper estimates the error in this prediction to be 5%. If the newspaper repeats the survey with 400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? O.32/0.23 O.44/0.50 UC, ES O.69/0.59 O.60/0.59 ER, PC O.60/0.59 ER, PC O.53/0.58 ER, PC	the course of the entire calendar year (January-December)?				UC
how much money will you have? A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60 % of the votes. The newspaper estimates the error in this prediction to be 5%. If the newspaper repeats the survey with 400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? ER, PC 0.69/0.59 0.60/0.59 ER, PC 0.63/0.52 0.53/0.58 ER, PC					
A newspaper conducts a survey and predicts that in the local election between Candidates A and B, Candidate A will receive 60 % of the votes. The newspaper estimates the error in this prediction to be 5%. If the newspaper repeats the survey with 400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? ER, PC 0.69/0.59 0.60/0.59 ER, PC 0.20 (N=82) 0.63/0.52 0.53/0.58 ER, PC	how much money will you have?	0.11 (N=37)	0.32/0.23	0.44/0.50	UC, ES
mates the error in this prediction to be 5%. If the newspaper repeats the survey with 400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? O.36 (N=72) O.69/0.59 ER, PC O.63/0.52 O.63/0.52 O.53/0.58 ER, PC					
400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? O.36 (N=72) O.69/0.59 ER, PC O.60/0.59 ER, PC	didates A and B, Candidate A will receive 60 % of the votes. The newspaper esti-				
400 participants, how many people can report that they will vote for Candidate A for the result to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? ER, PC will receive 60% of the votes with 5% error)?	mates the error in this prediction to be 5%. If the newspaper repeats the survey with	0.26 (N. 72)	0.60./0.50	0.60.40.50	ED DC
60% of the votes with 5% error)? Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? ER, PC O.20 (N=82) O.63/0.52 O.53/0.58	400 participants, how many people can report that they will vote for Candidate A for	0.36 (N=72)	0.69/0.59	0.60/0.59	ER, PC
Several days later, the newspaper conducts another survey with 300 new participants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? ER, PC O.20 (N=82) O.63/0.52 O.53/0.58	the result to be consistent with the original prediction (that Candidate A will receive				
pants. What is the minimum number of votes that Candidate A can receive in this new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)? ER, PC O.20 (N=82) O.63/0.52 O.53/0.58 ER, PC					
new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)?	Several days later, the newspaper conducts another survey with 300 new partici-				
new survey in order to be consistent with the original prediction (that Candidate A will receive 60% of the votes with 5% error)?	pants. What is the minimum number of votes that Candidate A can receive in this	0.20 (NI_92)	0.62 /0.52	0.52 /0.58	ED DC
	new survey in order to be consistent with the original prediction (that Candidate A	0.20 (IN=02)	0.03/0.32	0.55/0.56	EK, FC
250	will receive 60% of the votes with 5% error)?				
					250

You want to carpet a 15 foot by 20 foot room. You have two carpet options to choose from. One is \$1.50 per square foot and the other is \$3.00 per square foot. How much more will your total bill be if you choose the more expensive carpet rather than the	0.09 (N=102)	0.38/0.25	0.60/0.59	AV, SA, DA
cheaper one?				
To carpet your 15 foot by 20 foot room and a hallway that is 4 feet by 12 feet, about how much total carpet do you need?	0.05 (N=88)	0.39/0.25	0.60/0.59	AV, SA, DA
A sweater that was originally \$100 is on sale for 30% off. Which of the following coupons should you use to get the lowest final price?	0.21 (N=51)	0.31/0.22	0.41/0.39	PC, SA
You drove 200 miles on 11 gallons of gas. Which of these is closest to the number of miles per gallon that you got?	V1: 0.37 (N=46), V2: 0.05 (N=93)	0.32/0.21	0.47/0.57	SA, ES, DA
If one scoop of lemonade powder is needed for every 12 ounces of water, then how many scoops should you add to 3 gallons of water to make it into lemonade? (16 ounces = 1 Pint, 2 Pints = 1 Quart, 4 Quarts = 1 Gallon)		0.62/0.50	0.47/0.47	UC, SA, AV
Your grocery store has a 20 ounce jar of peanut butter for \$4.00, and a 45 ounce jar for \$9.00. Which purchase will get you the best price per ounce?	0.07 (N=61)	0.28/0.22	0.41/0.45	PR, SA, DA

did not identify. We will investigate this further in future work.

Difficulty and discrimination values in Table 7.4 are reported for both the complete unfiltered Fall 2014 dataset and with students who reported that they stopped trying midway through the assessment filtered out. The statistics for the full sample allow for a more direct comparison with other multiple choice assessments, while the subsample highlights the fact that including students whose effort wanes midway through the assessment results in an overall decrease in Cronbach's α (see Table 7.3), and a decrease in item discrimination. On average, the difficulty of questions decreases by 0.07 when this filter is applied, and the discrimination increases by 0.04.

The difficulty values for free response questions are, on average, 0.15 lower than for the same question in the unfiltered multiple choice dataset and 0.08 lower than the effort-filtered dataset. This is perhaps to be expected as the free response questions were administered one at a time during class, so motivation and time spent answering the question was likely higher and fatigue was negligible. Lack of multiple choice options in this format also likely makes guessing less frequent.

Although the majority of the courses in which the assessment was admin-

istered were "non-intervention" classes in the sense that instructors did not attempt direct remediation of quantitative skills, we examined pre/post statistics throughout instrument development as an additional indicator of item performance. In particular, if an item decreased significantly in discrimination or increased significantly in difficulty from pre to post instruction, we either revised or reworded it.

For example, Figure 7.11 shows pre vs. post difficulty and discrimination values for all of the questions on the Fall 2013 version of the assessment. These statistics are shown both for the entire sample, and for only those students who devoted effort to the assessment.

Fall, 2013 was also the semester that we completed most of the free response analysis, so every item but 9 was revised in some way, though in several cases (yellow points), only the distractors were revised. Particular attention was paid during revision to questions that increased in difficulty (e.g. 14, 20, 13) or decreased in discrimination (e.g. 6, 24, 15, 1) from pre to post, and all questions that fell into these categories were either cut or revised.

In general, the results of the preliminary pre-post analysis are encouraging. Even in courses that don't explicitly emphasize numerical skills, difficulty and discrimination show the proper trends toward improvement. In both the filtered and unfiltered datasets, two-thirds of the questions increase in discrimination from pre to post. When corrected for effort, nearly half of the questions also decrease in difficulty, though nearly all increase in difficulty if this filter is not applied (a result that will be discussed in further detail in Section 7.5.1.

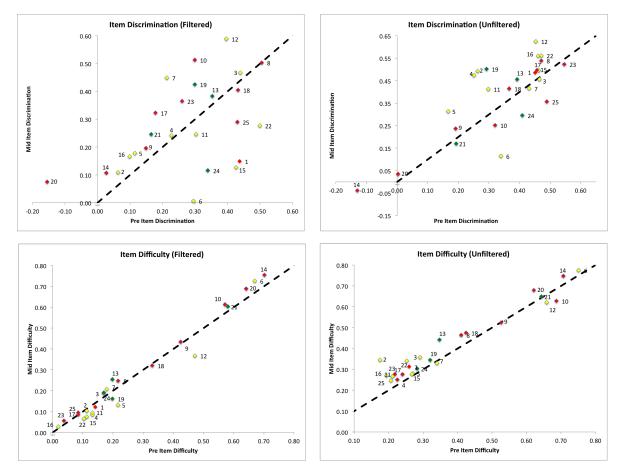


Figure 7.11: Pre/Mid semester changes in item difficulty (bottom panels) and discrimination (top panels) for the Fall 2013 instrument. The mid administration was choen in this case both because more students completed it and, as will be discussed in Section 7.5.1, a larger proportion of them devoted effort to the assessment. The data are shown both unfiltered (right panels) and filtered to exclude only students whose effort waned midway through the assessment (left panels).

7.5 Implementation and Lessons Learned

7.5.1 Attrition and Apathy

Across all semesters of development, attrition was severe. This is partly an effect of students dropping courses between the pre and post assessments (one reason why matched data are critical in pre-post surveys), and partly an effect of student participation rates for individual assignments decreasing over the course of the semester. For the purposes of this study, we define "attrition" as students dropping out of the sample by not completing later assessments.

We also see the effect of what we will refer to as "apathy" in the sample. For our purposes, this means a waning of effort from early semester assessments to late semester assessments, as measured by our question asking students to self-report their own degree of effort.

We made two changes to the instrument in Fall 2013 in an attempt to quantify the effects of apathy and attrition. First, we added the effort question described in detail in Section 7.3.6. Second, we suspected that post results from previous administrations were low in part due to a waning of student effort and focus from early to late in the semester, particularly as the instrument is assigned for participation credit and points are not assigned to students based on their performance or effort, only completion. In order to quantify this effect, we added a mid semester administration of the same instrument, in addition to the pre and post administrations, to allow us to test whether student apathy dropped steadily throughout the semester or precipitously in the last few weeks when students are often overwhelmed with other duties and are perhaps less likely to devote effort to an ungraded

activity.

The "mid" administration was done approximately 75% of the way through the semester (between the 11th and 13th weeks of a 16 week semester). Table 7.5 summarizes the number of complete assessments at each administration, as well as the number of matching assessments between various administrations. The effects of both attrition and apathy are clear. As the semester progressed, fewer and fewer students completed the assessments as assigned. At the same time, fewer and fewer of the students completing the assessment reported that they devoted effort ("tried pretty hard" or "tried my best on all/most") on each successive assessment, dropping from 52% of completed assessments at the pre to 45% at the mid to 40% at the post.

Table 7.5: Completion Statistics

	Pre	Mid	Post
Completed	299 ¹¹	249	182
Completed with Effort	158 (52%)	113 (45%)	73 (40%)

Because we collected student names, we were also able to match pre, mid and post assessments taken by the same individuals. The relative proportions of students who completed each assessment or combination of assessments is shown in Figure 7.12. Of the 371 unique individuals included in the sample, 33% completed only one of the assessments (19% pre only, 10% mid

¹¹Two classes have been excluded for the purposes of this analysis. One in which the instructor did not assign the mid or post instruments and one in which the post assessment was offered as an optional extra credit activity and not assigned for participation credit. The remaining three courses were all large lecture classes where all three assessments pre, mid and post-were required for participation credit.

only, 5% post only), 36% completed two of the assessments (22% pre+mid, 10% pre+post, 5% mid+post) and 30% completed all three assessments.

Among students who completed one, two or three assessments, we've also indicated in Figure 7.12 the number who completed the assessment(s) in question with a sufficient effort level to be included in the final sample. Among students who completed the assessment only once and those who completed it twice, about 50% survive the effort filter. However, among students completing all three assessments, only 29% put a sufficient amount of effort into all three to be included in the final sample.

Because our data are matched, we are also able to quantify the movement of individual students among effort categories. The results of this analysis are shown in Figure 7.13. As the purpose of administering the assessment three times during the semester was to determine whether earlier administration of the "post" assessment would mitigate the effects of apathy and attrition, we include only comparisons between the pre/mid and pre/post effort levels.

The results indicate that, as suspected, student apathy does increase in a measurable way toward the end of the semester. In particular, in students who matched across the assessments (which corrects for the effect of attrition throughout the semester), the proportions who maintained and increased their effort are higher for the mid assessment than for the post. Similarly, while 32% of the matching students exerted less effort on the midsemester assessment than on the pre-semester, this number was markedly higher (42%) on the post assessment 3-4 weeks later. Similarly, the proportion of students who "survive" the effort cuts for both samples is lower in the pre-post matched sample (35%) than the pre-mid matched sample

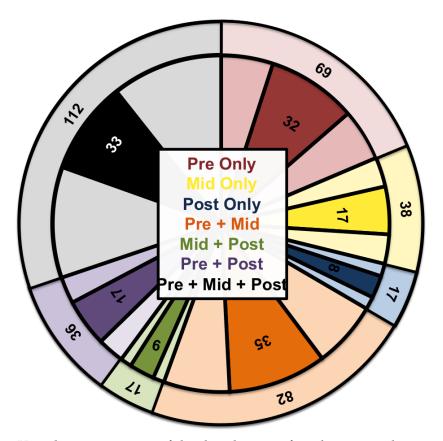


Figure 7.12: Visual representation of the distribution of students completing just one assessment (red, yellow, blue wedges), those completing two assessments (orange, green, purple wedges) and those completing all three (black wedge) in classes where all three were assigned for participation credit. Among the 371 unique students in this sample (represented by the outer circle), roughly one-third fall into each category. The dark wedges along the inner circle represent the students within each category who completed the assessment with effort. These students represent 50% of the population in each lightly shaded wedge, except in the region representing students who completed all three assessments (black region), where the proportion falls to 30%.

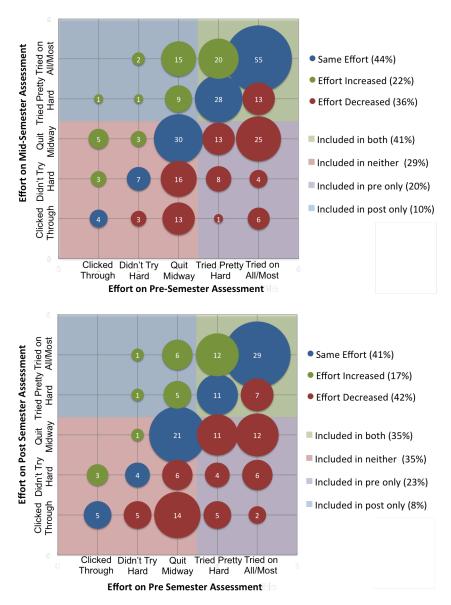


Figure 7.13: Charts representing the movement of students among effort categories on pre-mid (top) and pre-post (bottom) matched assessments. Effort on the pre-semester assessment is along the x-axis in each case and effort on the mid/post is along the y-axis. The size of each bubble in the grid corresponds to the number of students falling into that category. Bubbles are colored according to whether students improve (green), maintain (blue) or reduce (red) their effort level. The background of the plot is divided according to regions where students have always devoted sufficient effort to persist into the final filtered sample (green), students who never devote enough effort to be included in a sample (red) and those who fall out of the sample at the pre (blue) or mid/post (purple) assessments. The number of students in the purple region increases throughout the semester, indicating increasing apathy.

(41%).

The magnitude of the change in students' quantitative abilities will likely rely on the amount of time available to an instructor to remediate, an effect that would tend to argue for late-semester administration. At the same time, improvements will be increasingly difficult to measure as student apathy increases and a larger proportion of the sample must be excluded. It appears from our data that administration of the "post" assessment 65-75% of the way through the semester may result in a larger sample of assessments for analysis, and so we will advocate for this in future administrations.

7.5.2 Effects of Effort Filter on Distribution

The inclusion of an effort filter also substantially affects conclusions about student performance throughout the semester. During the 2011 and 2012 administrations, when we were administering the assessment during the last two weeks of the semester, we often saw a decrease in student scores from pre to post semester. This is shown, for example, in the lefthand histograms of Figure 7.14, which show a population of students evidently moving into the lowest-scoring (5-10) range on the post assessment. As the effort-filtered (only students devoting effort to the assessment are kept) histograms at right reveal, this effect disappears when students who are not devoting effort to the assessments are filtered out.

As described in Section 7.3.6, this was the reason that we added the effort filter in the first place - because we suspected that late semester apathy was contributing to low post scores and that general education science courses were unlikely to *hurt* students' quantitative abilities, though as the similarity of the filtered pre, mid and post distributions reveal, they are also

unlikely to help it (at least in the case of non-intervention classes).

Because Phase I of the QuaRCS study did focus on intervention, statistics on student improvement over the course of a semester are not reported in this paper, and are the focus of the broader Phase II study ¹².

7.5.3 Additional Filters

In addition to filtering student responses by reported effort, we've also implemented completion and time filters for the sample. Students completing fewer than 95% of the questions are excluded from analysis as are those who complete the assessment in under 10 minutes. This cutoff was determined by analyzing the data from the sample of 34 experts (Astronomy and Numeracy Educators) who completed the assessment. Among experts, the fastest time to completion was 11 minutes. Given this, we believe that students taking the assessment in 10 minutes or less are highly unlikely to be expending effort, regardless of what they report regarding their effort level. This results in the exclusion of 8, 14 and 5 students from the pre, mid and post instruments respectively, which amounts to under 1% of the remaining students in each case.

With these students excluded, we are able to analyze completion time statistics for students who expend effort. Unsurprisingly, the medians for completion time in the filtered sample for each of the three assessments are 5 minutes or more longer than in the sample as a whole.

Completion times also decrease by 5 minutes between each assessment in both the filtered and unfiltered samples. This could be due to any number of

¹²Instructors interested in participating in this study should contact Kate Follette at kfollette@as.arizona.edu

Table 7.6: Filter Statistics

Assessment	Total Com- pleted	Median Time	Total Surviving Effort Filter	Total Surviving Effort + Time Filters	Median Time with Filters
Pre	542	26	289	281	31
Mid	360	19.5	180	166	27
Post	211	15	91	86	21

Statistics for completed Pre, Mid and Post Semester assessments in the Fall, 2013 semester.

factors, including late-semester apathy, the smaller number of demographic questions included in the mid and post assessments 13 , student familiarity with the questions in the assessment, and familiarity with the assessment format. The higher completion rate and longer completion time of the midsemester assessment over the post-semester assessment supports our conclusion that future "post" assessments should be given $\sim 75\%$ of the way through a course rather than in the final two weeks, however we cannot rule out the effect of simply being asked to take the same assessment three times during a single semester as contributing to attrition and apathy in this data set.

7.5.4 Demographics of Low Effort Sample

In an effort to better understand the students who do not expend effort on the assessment, we analyzed the demographics of the full sample accord-

¹³we consider this unlikely since the demographic questions involve a very low cognitive load, and because the trend increases into the post assessment, which included an equal number of questions about the nature of the course

ing to effort expended. Many demographic variables proved not to be a factor in whether or not students devoted effort to the assessment. These included gender, age, major, class year, and amount of math background. Table 7.7 shows a selection of demographic and attitudinal variables that do correlate with effort level on the assessment. We find that the students who do not expend effort on the assessment are more likely to find the questions difficult, have low confidence in their answers, fail to see the utility of numerical skills in their daily life, and are less likely to use positive adjectives to describe mathematics ¹⁴. Interestingly, it appears that students who do expend effort on the assessment are more likely to be forward-thinking in choosing their major, as they are more likely to answer both that they chose/will choose it in order to get an enjoyable or well-paying job after graduation. All other choices related to why students chose/will choose their major proved to be uncorrelated with effort level, including whether students chose it "to avoid math".

7.6 Conclusions

On the whole, we have demonstrated that the QuaRCS is an effective assessment of numerical abilities and attitudes appropriate for the college general education science population. Refinements of the instrument as described in detail herein include:

(a) The QuaRCS Assessment is focused on 12 quantitative skills that were deemed important both for science literacy and for general educated citizenship by the 48 science and math educators completing our skills

 $^{^{14}}$ Though, interestingly, they are no more likely to use negative adjectives or to take a stance on the scale of Fun \rightarrow Scary

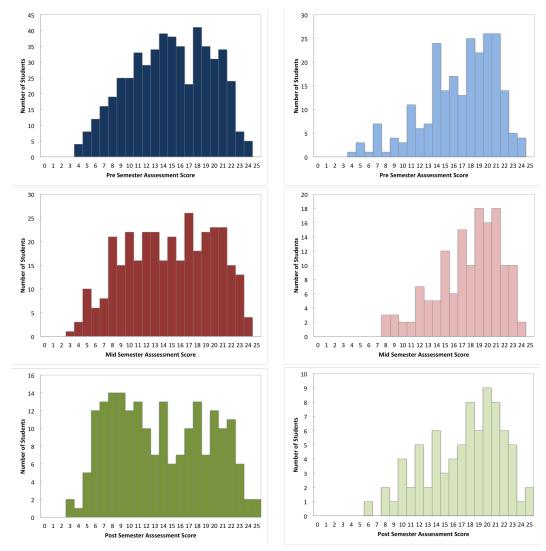


Figure 7.14: Histograms of students scores on the three Fall 2013 assessments - pre (top), mid (middle) and post (bottom) semester. The lefthand column represents the raw samples (N=518, 343, 202), and the righthand column represents the effort and time filtered sample (N=282, 175, 91). The mean of the filtered samples (pre=16.7, mid=17.6, post=17.6) are systematically higher than those of the unfiltered sample (pre=14.8, mid=14.7, post=13.7). This filter becomes particularly important when comparing the pre and post-semester distributions. The raw post population appears to score significantly lower on the assessment than the pre semester sample, but the large population of students in the 5-10 score range that appears in the post-assessment histogram all but disappears when filtered for self-reported effort. We believe this is a reflection of increased apathy toward the end of the semester, as discussed in detail in the text. Note also that the number of students completing each assessment drops significantly as the semester progresses.

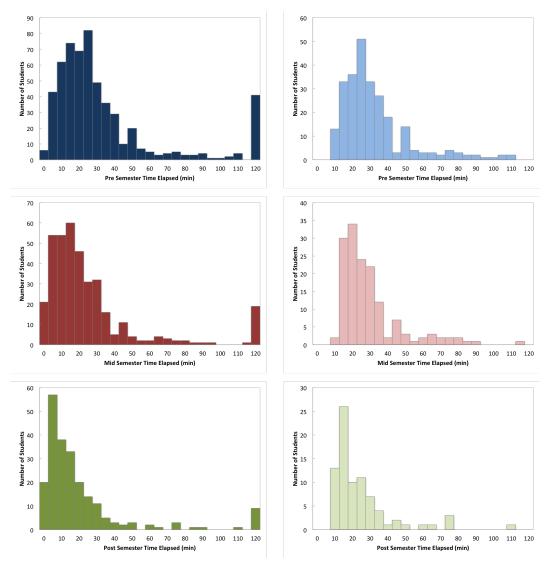


Figure 7.15: Histograms of time elapsed on the three Fall 2013 assessments - pre (top), mid (middle) and post (bottom) semester. The lefthand column represents the raw samples (N=518, 343, 202), and the righthand column represents the effort and time filtered sample (N=282, 175, 91). The median of the filtered samples (pre=30.0, mid=26.0, post=20.0) are systematically higher than those of the unfiltered sample (pre=26.5, mid=20.0, post=16.0). Furthermore, for all samples, the median time to completion decreases by \sim 5 minutes with each administration. We believe that this is primarily a result of increased apathy, as discussed in detail in the text.

Table 7.7: Effort Statistics

	Effort=1	Effort=2	Effort=3	Effort=4	Effort=5	
Question	Clicked	Didn't try	Quit	Tried pretty	Tried best on	Sig
	Randomly	very hard	midway	hard	most	
Overall, How difficult were the questions in this survey?	2.96	2.95	2.89	2.39	2.21	0.000
1= very easy, 5=very difficult	2.50	2.75	2.07	2.37	2.21	0.000
Which of the following best describes your general confi-						
dence in your answers? 1= conf. in very few 5=conf. in	1.84	2.64	3.02	4.11	4.28	0.000
all/almost all						
In your everyday life, how frequently do you encounter						0.000
situations similar to the problems in this survey? 1=almost	2.32	2.67	3.00	3.37	3.71	0.000
never, 5=daily I chose (or will choose) my major because:						
I feel it will help me get a job I will enjoy after graduation	47%	52%	53%	63%	70%	0.021
I feel it will help me get a well-paying after graduation	36%	36%	45%	57%	57%	0.021
How frequently do you do calculations in your everyday	3070	2070	1070	37,70	37,70	0.007
	3.04	2.60	2.93	3.11	3.24	0.000
life? 1=never 4=frequently						
How frequently do you encounter graphs in your daily	1.96	2.00	2.25	2.40	2.47	0.001
life? 1=never 4=frequently	1.50	2.00	2.20	2.40	2.17	0.001
Which of the following adjectives would you use to describe math?						
Interesting	24%	26%	34%	39%	45%	0.038
Easy	16%	26%	14%	21%	29%	0.009
Useful	36%	48%	55%	59%	69%	0.002
Rate the degree to which you agree with the following statement: 1=strongly agree, 4=strongly disagree						
I feel confident using numbers in my non-math courses	1.92	2.17	1.99	1.87	1.82	0.038
I feel confident using numbers in my everyday life	1.96	2.00	1.93	1.79	1.62	0.000
Numerical skills are important to the understanding of science	1.88	1.76	1.70	1.57	1.38	0.000
Numerical skills are important to my everyday life	2.00	2.05	1.95	1.81	1.77	0.031
	Effort=1	Effort=2	Effort=3	Effort=4	Effort=5	Total
Number of students	N=25	N=42	N=185	N=111	N=177	N=541

survey.

- (b) The length of the survey (25 quantitative questions, 25 non-quantitative demographic and attitudinal questions) was shown to be free of the effects of test fatigue among students who expend effort throughout. We note, however, that 25% of assessment takers report that their effort waned midway through the assessment.
- (c) Question wording was extensively vetted through multiple semesters of administration, and in particular through implementation of "free-responses" format questions. These were used to generate authentic distractors, vet and clarify question phrasing, and to ensure that students could not arrive at a correct answer through incorrect reasoning.
- (d) Both paper and online formats were explored. It was determined that the benefits of online assessment (out of class administration, ease of data collection, freedom from time constraints) outweighed the risks (lack of control for calculator and resource usage, idling).

Analysis of the data collected during development of the instrument further demonstrates the following:

- 1. Administration of demographic and attitudinal questions allows us to probe student motivations and to explore correlations between these variables and performance on the assessment.
- The administration of confidence rankings after each quantitative question allows us to probe student awareness of numerical deficits. Preliminary results suggest that students are often "unconsciously incompetent".

- 3. The addition of a question asking students to rank their effort on the assessment provided substantial improvements in our ability to interpret the results, including explaining why early QuaRCS administrations showed a decrease in student scores over the course of a semester.
- 4. Tracking of students among effort categories throughout the semester reveals that they are more likely to devote effort to the pre assessment than the post, regardless of whether the post is administered 75 or 95% of the way through the semester.
- 5. Collection of matched data allows us to correct for the effects of attrition throughout the semester.
- 6. Administration of the instrument 75% of the way through the semester rather than in the last two weeks suggests that this mitigates the effects of attrition and late-semester apathy, and results in a larger sample of students surviving effort filters.
- 7. Cronbach's α for the final version of the instrument is high (0.89) and item discriminations are above 0.3 for all but two items, as described in detail in the body of this paper. These statistics indicate that this instrument is valid for this population.
- 8. Scores of experts are significantly (0.000) different than those of students on the assessment, indicating that the instrument is capable of distinguishing the quantitatively literate for those who are not.

Having demonstrated the QuaRCS to be an appropriate and robust assessment for our purposes, we can now begin to use it to address some of the questions and concerns outlined in Section 7.2 of this paper.

It is important to note that the initial phase of the QuaRCS study reported herein (Phase I) was not an intervention study, and as such we did not expect students to demonstrate improvement in quantitative literacy, as indeed they do not in our analysis. This serves simply to verify our initial suspicions that the majority of introductory astronomy courses do not foster any improvement in students' skills or attitudes as they are currently taught, and argues for the need for research-validated intervention strategies.

One class in the Fall, 2013 administration and two classes in the Spring, 2012 administration engaged in significant remediation of quantitative skills, and analysis of these data provide some encouraging results regarding the ability of instructors in such courses to improve student attitudes toward the quantitative. However, we note that both of these administrations were of still problematic versions of the assessment, and we are therefore unable to draw any final conclusions about the effectiveness of the instrument at identifying courses in which improvement of student skills is being made. This is the goal of Phase II broader implementation of the QuaRCS. We hope that we will ultimately be able to answer definitively whether a semester of college science instruction is able to improve either (a) students' transferable quantitative skills, (b) their attitude toward mathematics, or (c) their ability to recognize their own numerical deficits. Ultimately, we hope to identify exemplary science educators who are making a difference in their students' quantitative abilities, and to inform the practice of all science educators in tackling the very important problem of innumeracy.

CHAPTER 8

EDUCATION CONCLUSIONS

With the validation and development stage (Phase I) of the QuaRCS study completed, I am now prepared to move on to the implementation phase (Phase II). As Phase II of the QuaRCS study relies heavily on broad implementation, I have already begun recruiting instructors and obtaining site authorizations (currently in place for Pima Community College, St. John's University and Queensborough Community College/CUNY) through networking at education conferences and through our faculty workshop series. We launched a newsletter in Fall, 2014 that will help to keep workshop participants informed of the results of the study and opportunities for involvement, and I have continued to post and advertise QL-focused astronomy curricular materials on my website.

Preliminary results that we are particularly eager to investigate during Phase II include the following observations from analysis of data collected during development of the QuaRCS:

- (a) students highly overrate their quantitative skills, often feeling confident that they have arrived at a correct answer when they have not
- (b) quantitative skills among students in introductory science courses span a broad range, making it difficult to show and to achieve statistically significant gains in ability, particularly given a documented waning of effort over the course of the semester
- (c) student attitudes, and in particular the likelihood that they will use positive adjectives to describe numerical skills, can and do rise over

the course of the semester, but only in classrooms where such skills are emphasized in very particular nonthreatening ways.

Phase II of this study will involve administering the fully-develped QuaRCS assessment in a variety of classrooms and analyzing the data. We hope that this phase will answer whether there are instructors who are improving their students' transferable quantitative skills over the course of a semester. If and when this is demonstrated, it will lead naturally into Phase III, when we will expand the study to include classroom observations and curricular analysis. This will be particularly important in the classrooms of educators revealed in Phase II to be improving their students' skills. What methods are instructors who are succeeding at improving their students attitudes and/or abilities using in the classroom?

Informed by these results, we will then begin the final phase (Phase IV) of the study - the development and testing of curricular materials utilizing the "best practices" identified in Phase III to assist educators in incorporating quantitative literacy into their classrooms. Specifically, I would like to develop and classroom test materials that emphasize the quantitative skills tested in our instrument, including:

- (a) A bank of numeracy-focused Think-Pair-Share (TPS) questions
- (b) A set of introductory astronomy labs that emphasize specific numerical skills
- (c) Activities that walk students through a review of each of the skills tested by our instrument

I have already begun developing these materials in my own classroom, and many are available in preliminary form on my website at http://www.katefollette.com/QL

though it remains to be seen whether these are the correct tools for improving QL skills, which Phase III results will inform. I also plan to expand our efforts to encompass other scientific disciplines, beginning with introductory geology for non-science majors. Together, Introductory Astronomy and Geology are taken by up to half of all college students (based on University of Arizona enrollment figures), thus involving our Geology colleagues in this endeavor is an important next step towards informing broader STEM education through our research.

In addition to continuing the QuaRCS study, I hope to raise awareness about the problem of quantitative illiteracy among my astronomy colleagues and to inform the debate on whether mathematical reasoning ought to be removed from introductory science curricula by presenting my work at science conferences and by continuing faculty development workshops. It is my sincere hope that my work will eventually demonstrate that it is indeed possible to make a difference in students' quantitative abilities through a semester of science instruction, and that I will be able to contribute to the broader dialogue regarding innumeracy among American students and adults.

APPENDIX A

SR21 DISK MODELING SUPPLEMENT

A.1 Optically Thick Scattered Light Models

In this appendix, we detail our attempts to derive a disk model in which the H-band scattered light emission originates from the surface of an optically thick small grain disk. Although we ultimately invoked an optically thin component in order to explain our data, this was not the most obvious solution, and more than 100 optically thick small grain scenarios were investigated before we arrived at the conclusion that the steepness of the observer H-band radial intensity profile is due to the effect of an optically thin small grain disk component.

Due to the relative expense of full scattered light models, we have not explored the entirety of the parameter space for the optically thick small grain disk solution. We therefore describe our efforts in this appendix, in the hope that they will help to inform future H-band scattered light modeling efforts.

A.1.1 Gapped and Differentially Depleted Models

If the r<36AU gap is optically thin at all wavelengths, as has been assumed by most previous models, then sufficient flux is incident on the dust wall at 36AU to create a sharp rise in the SED from $12-20\mu m$. We call the class of models with a depleted cavity that is optically thin in both large and small grains "Gapped" models.

Our initial working assumption, though perhaps naive, was that the observed SR21 cavity would also be seen in small grains. As discussed in

section 4.3, this hypothesis provides an excellent fit to the multiwavelength SED. However, if the cavity size and depletion factor suggested by the submm data is carried over to the small grains that scatter at H-band, the simulated H-band polarized intensity radial profiles produced by the Whitney code (Figure 17) suggest that we should have seen an observable discontinuity in the H-band scattered light PI profile at the r=36AU sub-mm wall. We see neither a slope nor a flux discontinuity at this radius, and can immediately exclude the simple scenario in which the large and small grains follow precisely the same distribution.

The simplest scenario that allows us to maintain a similar distribution for large and small grains is one in which the two populations are both depleted inside of the cavity, but each to a different degree (a different δ value). Physically, this would seem to make sense, as the large grains concentrated at the midplane should be subject to different physical processes (perhaps most notably clearing by forming protoplanets) than the grains at the surface of the disk where H-band scattered light originates. Second generation small grains are also expected to be generated by the planet formation process (e.g., Zuckerman 2001), and so even if small grain depletion also occurs at the midplane, it could in some sense be compensated for.

We began by exploring a variety of scenarios in which the small grains follow the same basic distribution as the large grains, are undepleted in the outer disk, but are depleted to a different degree than the large grains inside of the r<36AU cavity. Radial profiles for models following the same basic distribution as the large grain disk (α =2.15, β =1.15), but with depletion values ranging from 10^{-8} to zero (10^{0}) depletion, are shown overplotted on the observed H-band radial profile and on the observed SED in Figure

In moderate gap depletion cases (δ =10⁻³-10⁻⁵), the flux drops sharply inside of the gap and falls all the way into the saturation radius, which is inconsistent with our observation of a continually inwardly increasing profile. In cases where the depletion is high enough to make the disk optically thin (δ =10⁻⁶-10⁻⁸), the radial profile well inside the gap does increase sharply inward, however there is invariably a brightness discontinuity at the gap radius. This is due to the small amount of "wall" that is tipped into the line of sight by the disk inclination, which creates an excess at this radius that is evident in the radial profile. In low gap depletion cases (δ =10⁰-10⁻²), the radial profiles are relatively smooth, but too flat.

The "Gapped" models provide excellent fits to the SED and sub-mm imagery. However, the increased emitting area of the gap "wall" invariably produces a local excess at 0"3. In order for such models to fit the observed radial profile, two things would need to occur. First, scattered light emission from the wall (whose existence is required by the SED fits) needs to be somehow suppressed to avoid a sharp discontinuity in scattered brightness. If the disk were perfectly face-on, it may be possible to hide the excess emitting area of the wall, but the range of inclination estimates from a variety of multiwavelength work (A11, this work, P08) suggest that this is unlikely.

Second, they require that the optically thin grains inside of the gap and optically thick grains outside of the gap share a nearly identical radial PI profile despite very different physical conditions. Steep radial profiles can be produced by both optically thin and optically thick disks, but the apparent precise equality between the two makes this scenario seem somewhat contrived. We therefore exclude the "Gapped Disk" models as incompatible

with the scattered light imagery except under the most contrived scenarios.

A.1.2 Scale Height Discontinuity Models

We also consider scenarios that do not require the small grains to be depleted (and therefore optically thin) inside of the large grain gap. A step-like discontinuity in the small grain scale height could contribute to the steep SED rise, without requiring that the inner disk be depleted in small grains relative to the outer disk and allowing for identical (though discontinuous) distributions from inner to outer disk. We call this class of models "Scale Height Discontinuity" models, and the outputs for a representative model are shown in Figure 18. A related class of models utilize a localized "puff" in the large grain outer disk wall. Inflated disk rims are often invoked to explain variable emission in disks (e.g., Flaherty et al. 2012). Under such models, the high incident flux on a disk wall causes higher temperatures at that location and "puffs up" the rim. If this were to occur at the 36AU large grain disk wall, it is conceivable that it would also inflate the small grain scale height at this radius.

At the limit of a perfectly vertical step-like discontinuity and a face on disk, there is no break in the radial profile between inner and outer disk at H-band because the optical depth of these grains is large and the wall is obscured by the disk surface as viewed from face-on. However, as with the gapped disk models, it is difficult to mimic this situation in an inclined disk, even one with low inclination like SR21. These models thus solve the problem of requiring a different disk geometry inside and outside the gap however they also invariably produce a local excess at the gap radius that is not seen in the data, just as the "Gapped" models did.

In the case of puffed disk rims, the discontinuity that causes the puff is localized to the "wall", so the disk soon settles back into its normal scale height distribution. These "puffs" may have a small radial extent, but again it is difficult to construct a situation in which they wouldn't be visible as an excess in the radial profile. On the other hand, the puffed disk rim shadows a portion of the disk immediately exterior to it from incoming stellar flux, where the area of the shadowed region depends on the height of the "puff". This local excess ("puff") followed by a local minimum (shadowed region) could be smoothed out in the polarized intensity radial profile if the total radial extent is much smaller than the diffraction limit at H-band, mimicking our observed smooth radial profile. This precise balance again seems somewhat contrived, so we consider the puffed rim explanation to be relatively unlikely.

Owing to their continuous density distribution, brightness discontinuities in the "Scale Height Discontinuity" models are less drastic than those created by the "Gapped" disk models, so they are more likely to remain unresolved in our observational data. Non-diffraction limited performance, the peculiar shape of the HiCIAO PSF and the relative insensitivity of polarized intensity emission to sharp disk features (due mainly to low overall flux) could all conceivably smooth out such discontinuities in the radial profile to a greater extent than convolution with a 0.06" Gaussian PSF would suggest. However, we were unable to construct a disk model in which the radial profile was smooth enough to be consistent with our data and error bars, so we exclude these models as relatively unlikely candidates.

A.1.3 Steeply Curved Models

If we relax the assumption that the small grain disk needs to have precisely the same parameters inside and outside the gap, we find that there are cases where a turnover in the radial scale height exponent β at the 36AU wall can reproduce the steep SED rise. This situation does not require small grain depletion, nor does it require a discontinuity at the large grain dust wall. In such cases, the SED rise coincides with the region just inside of 36AU where the disk is most steeply curved, and in that sense mimics a "wall". From a physical standpoint, this transition from a large to a smaller radial scale height exponent at the large grain dust wall would need to be explained somehow by the dearth of large grains at the disk midplane in this region or the physical mechanism causing disk clearing, which we do not endeavor to do here. We call this class of models "Steeply Curved" models, and outputs for a representative model are shown in Figure 19. The scattered light imagery produced by each of these models is described in detail in Appendix A.

Under the third well-fitting SED scenario, a steep disk curvature (high β) in the inner disk mimics a "wall" at the transition to a shallower outer disk profile. Without a physical discontinuity in the small grain surface at 36AU, these models are also free of discontinuity in their radial polarized intensity profiles. However, all "Steeply Curved" models also produce flat (\mathbf{r}^0) radial profiles, which our observations argue against. This is because the geometric emitting area per pixel increases as the disk becomes increasingly flared, so pixels represent a larger amount of disk surface as you move outward. This mitigates the \mathbf{r}^{-2} effect of incident stellar radiation, and makes the flux in the disk extremely flat.

We find that the surface brightness of the disk is relatively insensitive to the radial density exponent α in our high- β models, even for very strongly inwardly peaked density profiles (e.g., r^{-10}). These models still show a flattening of the brightness profile in steeply curved regions that our data rule out. We therefore exclude these high- β undepleted scenarios as poor fits to the H-band data.

We find that the steepness of the radial profile significantly limits the disk geometry, particularly in the optically thick case. We tried more than 30 iterations of smooth optically thick disks with a wide range of α , β and α - β (the radial dependence of the surface density) values, and were unable to create a steep r^{-3} profile. The only optically thick disks we found that reproduce an r^{-3} radial polarized intensity profile are geometrically flat. As the geometrically flat case requires that the scale height of the disk be smaller than the radius of the star, it is impossible to create one in which the large grain disk required to reproduce the sub-mm flux does not stick up above the small grain disk. It is difficult to imagine a situation that would reproduce this, so we exclude these models on physical grounds.

A.1.4 Prospects for Envelope Component

Since SR21 may still be weakly accreting (the upper limit of $10^{-8.84}$ on the accretion rate is relatively high compared to other disks), we also explored the possibility that an envelope component is contributing to the flux in the interior of the gap. To provide observational constraints on the amount of circumstellar material, SR21 was observed on 2011 March 29 with SIR-POL (Kandori et al. 2006), a JHKs-simultaneous imaging polarimeter that is mounted on the Infrared Survey Facility (IRSF) 1.4m telescope in South

Africa. We conducted aperture polarimetry on this data and the results are summarized in Table 3. Our data reveal total polarization levels that are consistent with the range of total polarization observed for YSOs (Pereyra et al. 2009). We have not corrected the SR21 aperture polarization values reported in Table 3 for interstellar polarization, which could be significant due to the high foreground extinction (A_V =6.3). Moreover, because the observations were conducted in defocus mode to avoid saturation, the A and B components of SR21 (separated by 6") overlap in our aperture polarimetry, preventing us from using SR21B to determine how much of the measured polarization is due to foreground dust and how much is local to SR21A and/or B. There are no other nearby objects in the field suitable for use as field star interstellar polarization probes.

If we assume that the level of interstellar polarization and contamination from SR21B is minimal, we can use the measured polarization Position Angles (PAs) at each band to probe of the optical depth of the disk, as optically thick disks show polarization PAs oriented perpendicular to the PA of the disk major axis, while the polarization PA and major axis PAs of optically thin disks are oriented parallel to one another (Whitney & Hartmann 1992; Pereyra et al. 2009). The disk major axis PA ($86\pm11^{\circ}$, this work; 100° , A11) is roughly perpendicular to the measured polarization orientations at all three bands, suggesting that the disk may be optically thick in the NIR.

However, there are several caveats to be made. First, the 42° counterclockwise migration in the disk PA derived from isophotal fits (and described in section 3.1) suggest that an average PA is not a particularly good measure of the overall disk orientation. Secondly, these results depend on the robustness of our assumption that the interstellar polarization and contamination

from SR21B is small. Finally, and most importantly, we have managed to eliminate all of the optically thick small grain dust disk scenarios that we have investigated thus far. In fact, we find that optically thin disks provide a much better fit to the H-band radial profile, as described in Section 4.4, although we cannot rule out, and in fact require, an additional contribution from an underlying optically thick small grain disk component.

Table A.1: SIRPOL Aperture Polarimetry

Band	Polarization	PA
	(%)	0
J	2.32 ± 0.17	21.5 ± 2.1
Н	1.01 ± 0.05	18.0 ± 1.3
Ks	0.52 ± 0.10	179.2 ± 5.3

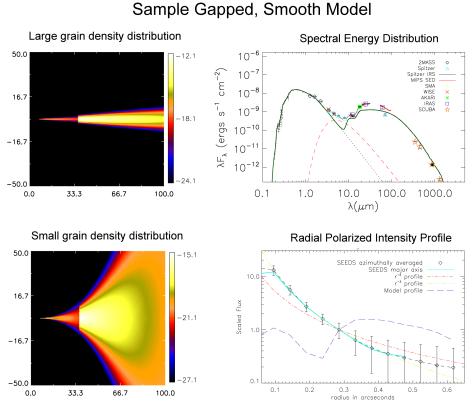


Figure A.1: Sample model outputs for a representative gapped disk model. The lefthand panels represent the density of large grains (upper) and small grains (lower). The right-hand panels provide model fits to the observed SED (upper) and H-band radial polarized intensity profile (lower).

Sample Scale Height Discontinuity Model Large grain density distribution Spectral Energy Distribution 10^{-6} 50.0 λF_{λ} (ergs s⁻¹ cm⁻²) 10 10⁻⁸ 16.7 10⁻⁹ 10^{-10} -16.7 10^{-11} 10 -50.0 0.1 1.0 10.0 100.0 1000.0 33.3 66.7 100.0 0.0 $\lambda(\mu m)$ Small grain density distribution Radial Polarized Intensity Profile 50.0 16.7 Scaled Flux -16.7 -50.0 100.0 0.3 0.4 radius in arcseconds

Figure A.2: Sample model outputs for a representative undepleted small grain disk with a scale height discontinuity at the 36AU large grain dust wall. The lefthand panels represent the density of large grains (upper) and small grains (lower). The righthand panels provide model fits to the observed SED (upper) and H-band radial polarized intensity profile (lower)

Large grain density distribution Spectral Energy Distribution 10^{-6} 50.0 $\lambda F_{\lambda}~({\rm ergs~s^{-1}~cm^{-2}})$ 10 10⁻⁸ 16.7 10^{-9} -18.1 10^{-10} -16.7 10^{-11} 10 -50.0 0.1 1.0 10.0 100.0 1000.0 0.0 33.3 66.7 100.0 $\lambda(\mu m)$ Small grain density distribution Radial Polarized Intensity Profile 50.0 16.7 17.8 33.3 66.7 100.0

Sample Steeply Curved Model

Figure A.3: Sample model outputs for a representative model in which the region of the sub-mm cavity is more steeply curved in the small grain disk. The lefthand panels represent the density of large grains (upper) and small grains (lower). The righthand panels provide model fits to the observed SED (upper) and H-band radial polarized intensity profile (lower).

Appendix B

$10\mu M$ Feature

As described in the main body of the text, we were able to create a model that fit both the global SED and the scattered light imagery well except in the region of the 10μ m feature, which is clearly being overproduced. This section describes our efforts to suppress the feature in the SED.

Model imagery in a $10\mu m$ narrowband filter reveals the gap wall as the dominant source of $10\mu m$ emission in our models. The location of this wall is well constrained by our images, and we find that other tunable parameters affect the strength of the $10\mu m$ feature only minimally (scale height, flaring, etc.). The simplest way to affect the strength of the feature seems to be modification of the grain prescription, as we were unable to suppress the $10\mu m$ feature to the observed level with a standard silicate-rich small grain dust prescription (e.g. Kim et al. 1994).

Among the small grain dust prescriptions that are standard in the Whitney code, we were best able to reproduce the small 10μ m feature with the prescription of Wood et al. (2002)'s Model 1, a mixture of ~55% amorphous carbon and ~45% astronomical silicates with a power law exponent of 3.0, a maximum grain size of 1mm and an exponential cutoff with a turnover at 50μ m. The best fitting SED generated with this small grain dust assumption, for which all parameters except for the small grain dust prescription and the cavity depletion factor (increased to 0.05) are the same as in Table 3.2, is shown in Figure B.1. This model provides a better fit in the 10μ m region, though whether it is a better overall fit is unclear, as the fit to the

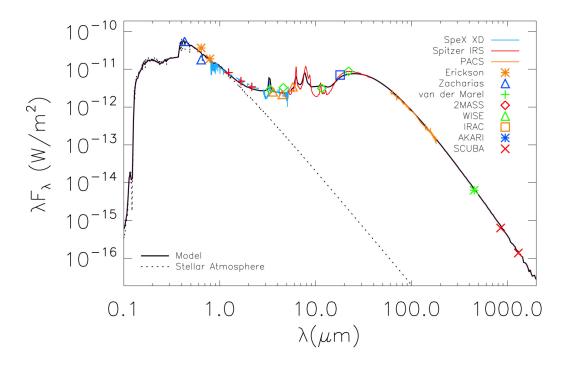


Figure B.1: Best-fitting Whitney model SED for Wood et al. (2002) Model 1 grains. This provides a better fit at 10μ m, but the IRS spectrum is poorly matched and the corresponding model imagery, shown in Figure B.2, is a much poorer fit to our observations.

Spitzer IRS spectrum is poorer.

The scattered light models generated under these assumptions, however, are a far poorer fit to our observed imagery. As shown in Figure B.2, the disk becomes significantly brighter to the south and fainter to the north in these models than is observed. This may be partially due to the necessity of increasing the depletion factor inside the gap, which increases the shadowing of the exposed Northern cavity rim, or it may be that these grains are more strongly forward scattering. As the model imagery was clearly a poor fit, we did not investigate in detail.

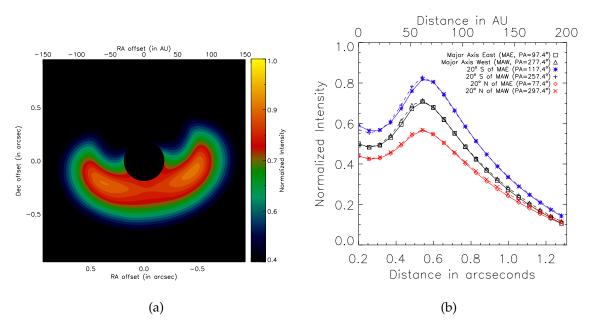


Figure B.2: K-band model image for Wood et al. (2002)'s Model 1 grains convolved with our PSF (left) and radial profiles taken through the same image (right). The H-band model and profiles are qualitatively similar. This model, though a better fit to the SED, is too bright in the South relative to the North, as revealed by the relative brightness of the model profiles along the major axis and to the south and north.

We thus reverted to the standard ISM-like dust prescription of Kim et al. (1994). We were able to mitigate the strength of 10μ m emission somewhat by modifying it to grains from only 0.1- 1.0μ m in size. Our best fit models still overproduce 10μ m emission, but the best fit model presented in the body of this paper is a much better fit to the observations as a whole.

Oph IRS 48 is neither the only disk with this problem (see, for example Follette et al. (2013b)), nor is the fit in this region of heavy PAH emission particularly good in any existing disk models. None of the other SED models in the literature (e.g. Bruderer et al. 2014; Maaskant et al. 2013) include PAH emission, and since PAHs also emit strongly in the 10μ m region, the 10μ m excess in other models, though present, is not as apparent.

We leave discussion of the mysterious lack of 10μ m silicate emission in this and other disks for future work, though we note that one way to reduce it may be to increase the minimum grain size to $>0.1\mu$ m.

APPENDIX C

RADIAL PROFILES OF OPH IRS 48

The radial profiles described in Section 3.3.3 were investigated relative to both the stellar residual and the center of the best fit to the cavity. Between these two potential reference points, profiles using the center of the cavity as the origin allow for clearer interprofile comparison, as they all peak at approximately the same radial location. Profiles taken relative to the cavity center also more closely match the circularly symmetric scattered light models discussed in Section 3.4.3. For completeness, we show profiles with respect to both reference points in Figure C.1. In the text, we have elected to use the lefthand (cavity referenced) profiles in our analysis.

In all cases, radial profiles are shown along the major axis to both the East and West of the origin (either the stellar residual or the center of the cavity), and 20 degrees offset from the major axis to both the North and South in each case.

Profiles taken through the final disk images at the two wavelengths also do not allow for a direct comparison between the structures at H and Ks bands and are potentially misleading. As the K band data are a factor of 2 smaller in FWHM, the rim appears much sharper in the profiles shown in Figures C.1a and C.1b than the H-band data in Figures C.1e and C.1f, so peak intensities cannot be directly compared.

In order to construct a direct comparison, we convolved our K band image with a Gaussian to simulate image quality degradation. The FWHM of the Gaussian (5.65pixels/0'05) was chosen in order to match the FWHM of the

median combined intensity image in both image sets, and these convolved profiles are shown in Figures C.1c and C.1d.

Convolving the Ks-band PI images with this Gaussian results in radial profiles that are remarkably similar between H and Ks bands for the Eastern half of the disk, and it is these two sets of radial profiles (Figure C.1c and C.1e that we have elected to use in the main body of the text.

A comparison between the convolved and unconvolved profiles also elucidates the reason why the Southern arc is not visible in the H-band data. Smoothing the K-band data in this way causes the arc to bleed into the stellar residual, erasing any trace of the cavity in this region. It is thus unsurprising that we did not resolve the Southern arc in the H-band data.

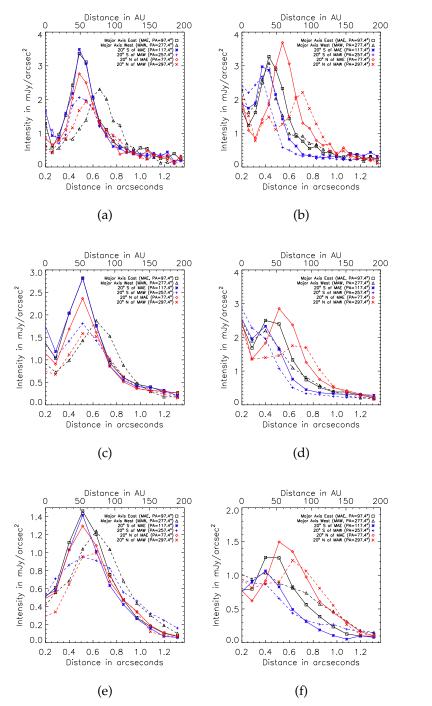


Figure C.1: Radial profiles taken through disk images deprojected to account for an inclination of 50° along PA=97.4°. Profiles were taken both with respect to the center of the circular K-band cavity (left column), and from the stellar location (right column). In all cases, solid lines represent profiles through the Eastern half of the disk and dashed lines represent profiles through the Western half of the disk. Black profiles were taken along the disk major axis and blue and red profiles were taken 20° North and South of the major axis, respectively. (a) and (b) Radial profiles through the Ks-band halo-subtracted and deprojected PI image. Profiles have been binned to 6 pixels (0.06″) radially, equivalent to half of the stellar FWHM.; (c) and (d) Radial profiles through a smoothed Ks-band halo-subtracted and deprojected PI image. The original image was smoothed with a 5.65 pixel (0.05″) Gaussian in order to match the stellar FWHM of the H-band image and allow for a direct comparison of the Ks and H-band profiles. Profiles have been binned to 12 pixels/0.11″ to match the H-band data.; (e) and (f) Radial profiles through the H-band halo-subtracted and deprojected PI image. Profiles have been binned to 12 pixels/0.11″ radially, equivalent to half of the stellar FWHM.

APPENDIX D

SKILL SURVEY RESPONSES

The figures in this Appendix are histograms of educator responses to our skills survey regarding the importance of numerical skills. A histogram for each skill is shown with responses from Astronomy educators in red and Math/Numeracy educators in blue. For each skill, responses regarding the importance of the skill for science literacy (solid lines) and for being a savvy consumer and citizen ("life", dashed lines) are shown.

- ☐ Science Literacy (Astronomy, N=19)
- ☐ Science Literacy (Math, N=29)
- Life (Astronomy, N=19)
- □ Life (Math, N=29)

Figure D.1: Legend applying to all subsequent figures in this appendix

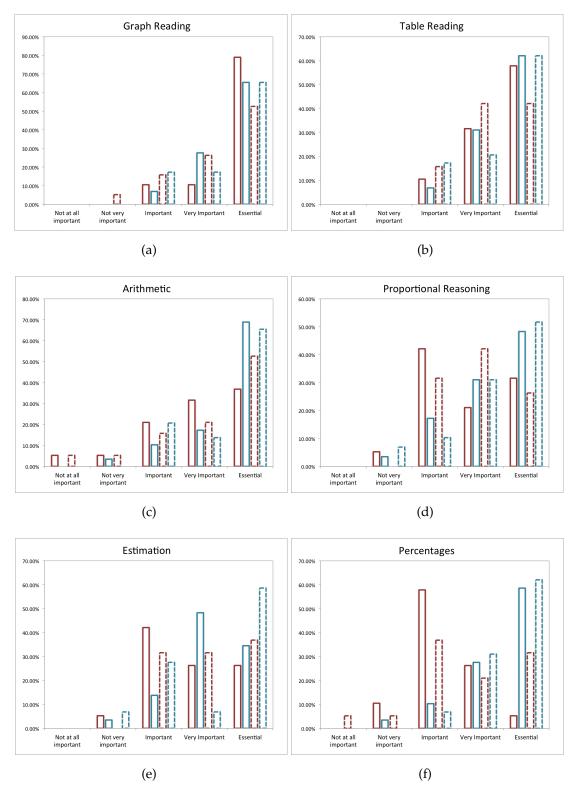


Figure D.2: Response distributions for astronomy and numeracy educators normalized to the number of respondents in each group for the six skills deemed most important by both groups.

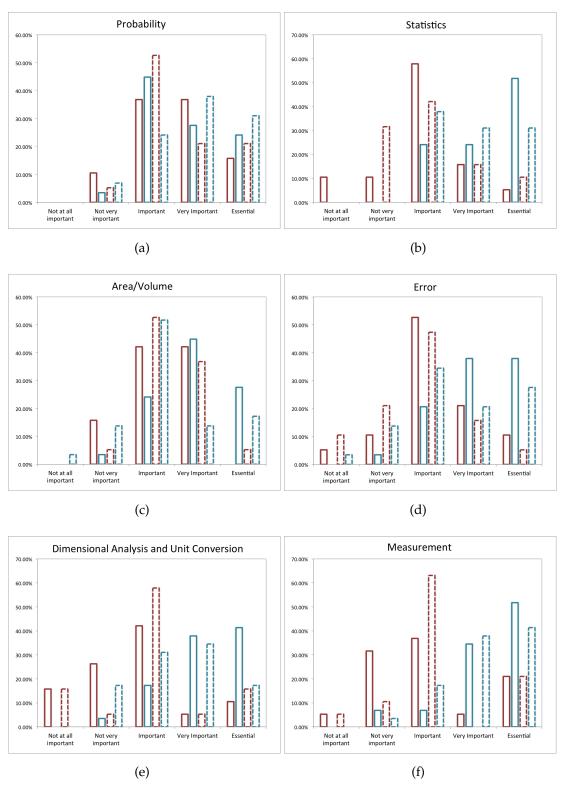


Figure D.3: Response distributions for astronomy and numeracy educators normalized to the number of respondents in each group for skills ranked 7-12 in importance.

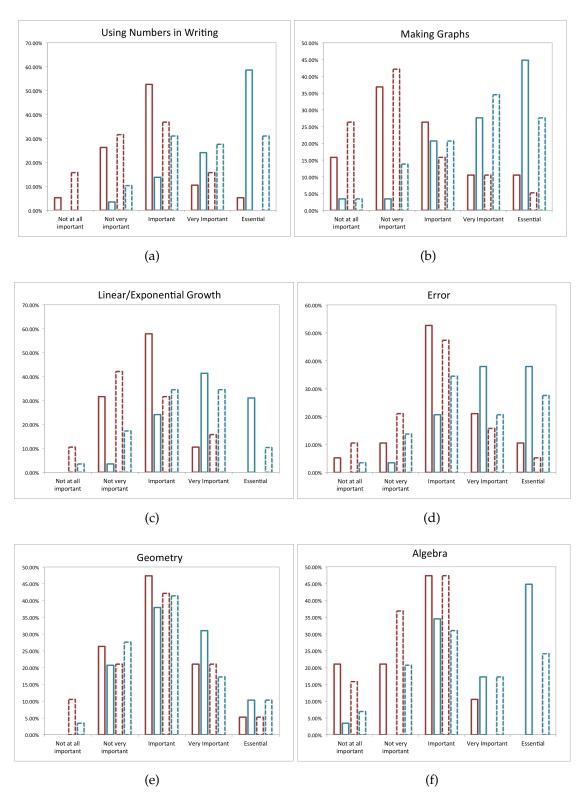


Figure D.4: Response distributions for astronomy and numeracy educators normalized to the number of respondents in each group for skills ranked 13-18 in importance.

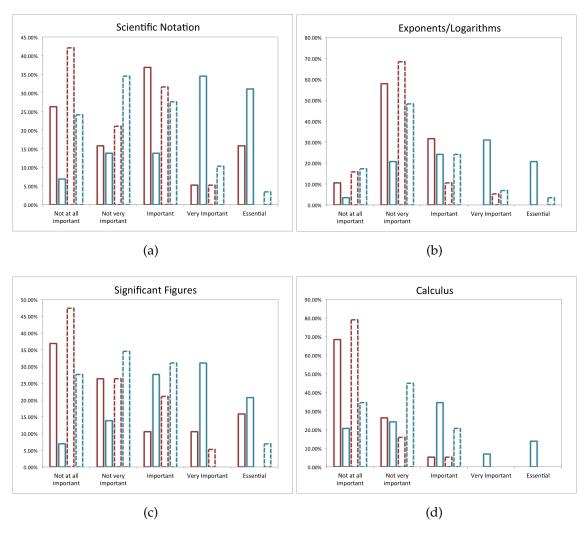


Figure D.5: Response distributions for astronomy and numeracy educators normalized to the number of respondents in each group for the four skills ranked as least important overall.

APPENDIX E

QUARCS DEMOGRAPHIC AND ATTITUDINAL QUESTIONS

This appendix lists the demographic and attitudinal questions in the QuaRCS as well as the statistics of student responses on the Fall 2013 and Fall 2014 pre-semester instruments combined (N=704).

Table E.1: QuaRCS Demographic Questions

Question	Response Options	Response Distribution
I am a college	a) freshman	62.1% Freshman
	b) sophomore	25.2% Sophomore
	c) junior	10.3% Junior
	d) senior	2.4% Senior
	e) other	0.4% Other
My age is	a) under 18	
	b) 18-25	4.0% Under 18
	c) 26-35	94.3% 18-25
	d) 36-45	1.7% 26-35
	e) 46-55	0.2% 36+
	f) 56+	
My gender is	a) Male	54.4% Male
	b) Female	45.2% Female
	c) Other	0.4% Other

Language) b) Arts (e.g. Art, Music, Dance, Film, Theater, Creative Writing) c) Social Sciences (e.g. Economics, History, Political Science, Psychology, Sociology, Anthropology, International Relations, Geography, Linguistics, Law) d) Education e) Science (e.g. Physics, Chemistry, Biology, Geology) Please select your major or majors Language) 6.4% Humanities 5.5% Arts 18.6% Social Sciences 4.4% Education 4.4% Science 6.1% Eng./Math/CS	(a) Humanities (e.g. English, Literature, Philosophy, Religion, Foreign	
c) Social Sciences (e.g. Economics, History, Political Science, Psychology, Sociology, Anthropology, International Relations, Geography, Linguistics, Law) d) Education e) Science (e.g. Physics, Chemistry, Biology, Geology) 6.4% Humanities 5.5% Arts 18.6% Social Sciences 4.4% Education 4.4% Science 6.1% Eng./Math/CS		Language)	
from the list below. f) Engineering, Mathematics or Computer Science g) Business-related (e.g. Business, Marketing, Management) h) Health-related (e.g. Nursing, Pharmacy, Nutritional Science, Public Health, Exercise Science) i) Trade-specific (e.g. Architecture, Agriculture, Justice, Library Science, Retail, Family and Consumer Sciences) j) Journalism k) Undecided l) Other	Please select your major or majors from the list below.	b) Arts (e.g. Art, Music, Dance, Film, Theater, Creative Writing) c) Social Sciences (e.g. Economics, History, Political Science, Psychology, Sociology, Anthropology, International Relations, Geography, Linguistics, Law) d) Education e) Science (e.g. Physics, Chemistry, Biology, Geology) f) Engineering, Mathematics or Computer Science g) Business-related (e.g. Business, Marketing, Management) h) Health-related (e.g. Nursing, Pharmacy, Nutritional Science, Public Health, Exercise Science) i) Trade-specific (e.g. Architecture, Agriculture, Justice, Library Science, Retail, Family and Consumer Sciences) j) Journalism k) Undecided	5.5% Arts 18.6% Social Sciences 4.4% Education 4.4% Science 6.1% Eng./Math/CS 36.0% Business 2.5% Health-related 2.3% Trade-Specific 1.8% Journalism 16.0% Undecided

I chose (or will choose) my major because: Check all that apply.	a) I like the subject b) I feel that it will help me get a job I will enjoy after graduation c) I feel that it will help me get a well-paying job after graduation d) I am good at it e) I chose a major that would avoid math as much as possible f) I chose a major that would avoid writing as much as possible g) I'm not sure yet h) Other	66.7% I like the subject 58.6% Job I'll enjoy 47.2% Well-paying job 13.8% To avoid math 5.1% To avoid writing 14.7% I'm not sure 6.0% Other
I completed the following classes before college. Check all that apply.	a) Algebra (including Algebra II and College Algebra) b) Geometry c) Trigonometry or Precalculus d) Calculus e) Statistics or Business Math f) Other g) None of the above	89.6% Algebra 86.7% Geometry 68.5% Trig./Precalc 25.4% Calculus 24.1% Statistics/Bus. Math 1.3% None of the Above 3.3% Other

I have completed or am currently enrolled in the following classes while in college. Check all that apply.	 a) Algebra b) Geometry c) Trigonometry or Precalculus d) Calculus e) Statistics or Business Math f) Other g) None of the above 	19.3% Algebra 4.3% Geometry 10.4% Trig./Precalc 15.6% Calculus 12.8% Statistics/Bus. Math 52.3% None of the Above 2.8% Other
How long ago was your last math course?	a) I am taking a math course nowb) Within the last yearc) Within the last two yearsd) Within the last three yearse) More than three years ago	29.3% Currently taking one 54.4% Within one year 11.5% Within two years 1.9% Within three years 3% 3+ years
I attended elementary, middle and high school entirely in the United States. (If no, please specify where and what grade(s))	a) yes b) no	87.9% Yes 12.1% No
How many more science courses do you plan to take after this one?	a) none b) one c) two d) three or more	24.5% None 37.4% One 23.3% Two 14.7% Three or more

	a) It is a prerequisite for courses in my major	16.0% Prerequisite
Why did you choose to take this course? Check all that apply.	b) To fulfill a university general education requirement	72.2% Gen. Ed. requirement
	c) It sounded interesting	58.6% Interesting
	d) It sounded easy	8.0% Easy
	e) I heard the class was good	13.6% Good Instructor
	f) I heard the instructor was good	11.1% Good course
	g) Other	2.5% Other

Table E.2: QuaRCS Attitudinal Questions

Question	Response Distribution	
How frequently do you do calculations in your everyday life?	a) Never b) Infrequently c) Sometimes d) Frequently	5.5% Never 19.5% Infrequently 40.4% Sometimes 34.6% Frequently
How frequently do you encounter graphs and tables in your daily life?	a) Never b) Infrequently c) Sometimes d) Frequently	17.3% Never 41.6% Infrequently 33.0% Sometimes 8.2% Frequently
Old: Which of the following adjectives would you use to describe math (doing calculations, reading graphs and tables, reasoning with numbers, etc.)? Check all that apply. New: Where would you put mathematics (including: doing calculations, reading graphs and tables, reasoning with numbers, etc.) on each of the following scales between two opposite adjectives.	Old: a) Interesting b) Useless c) Easy d) Boring e) Useful f) Hard g) Fun h) Scary New: Interesting 1 2 3 4 Boring Easy 1 2 3 4 Hard Useful 1 2 3 4 Useless Fun 1 2 3 4 Scary	Old: 37.5% Interesting 34.7% Boring 58.7% Useful 7.7% Useless 21.6% Easy 45.8% Hard 18.6% Fun 20.3% Scary New: See main body of paper

	a) Strongly Agree	28.1% Strongly Agree
Rate the degree to which you agree with the follow-	b) Agree	53.7% Agree
ing statement: I feel confident using numbers in my	c) Disagree	15.2% Disagree
non-math courses	d) Strongly Disagree	3.0% Strongly Disagree
	a) Strongly Agree	34.2% Strongly Agree
Rate the degree to which you agree with the following statement: I feel confident using numbers in my	b) Agree	53.7% Agree
	c) Disagree	10.4% Disagree
everyday life	d) Strongly Disagree	1.7% Strongly Disagree
	a) Strongly Agree	48.4% Strongly Agree
Rate the degree to which you agree with the follow-	b) Agree	47.1% Agree
ing statement: Numerical skills are important to the		3.7% Disagree
understanding of science	c) Disagree	0.9% Strongly Disagree
	d) Strongly Disagree	
Rate the degree to which you agree with the follow-	a) Strongly Agree	29.6% Strongly Agree
ing statement: Numerical skills are important to my	b) Agree	56.6% Agree
	c) Disagree	12.6% Disagree
everyday life	d) Strongly Disagree	1.3% Strongly Disagree
Rate the degree to which you agree with the follow-	a) Strongly Agree	18.3% Strongly Agree
, ,	b) Agree	45.4% Agree
ing statement: I am satisfied with my current level	c) Disagree	30.6% Disagree
of numerical/mathematical skill	d) Strongly Disagree	5.7% Strongly Disagree

REFERENCES

- Acke, B., & van den Ancker, M. E. 2006, A&A, 449, 267
- Acke, B., van den Ancker, M. E., & Dullemond, C. P. 2005, A&A, 436, 209
- Alexander, R., Pascucci, I., Andrews, S., Armitage, P., & Cieza, L. 2013, ArXiv e-prints
- Allen, M. J., & Yen, W. M. 1979, Introduction to measurement theory. Belmont, CA: Wadsworth
- Anders, E., & Grevesse, N. 1989, Geochim. Cosmochim. Acta, 53, 197
- Andrews, S. M., & Williams, J. P. 2005, ApJ, 631, 1134
- —. 2007, ApJ, 671, 1800
- Andrews, S. M., Wilner, D. J., Espaillat, C., Hughes, A. M., Dullemond, C. P., McClure, M. K., Qi, C., & Brown, J. M. 2011, ApJ, 732, 42
- Andrews, S. M., Wilner, D. J., Hughes, A. M., Qi, C., & Dullemond, C. P. 2009, ApJ, 700, 1502
- —. 2010, ApJ, 723, 1241
- Apai, D., Pascucci, I., Brandner, W., Henning, T., Lenzen, R., Potter, D. E., Lagrange, A.-M., & Rousset, G. 2004, A&A, 415, 671
- Apter, A. J., et al. 2009, Patient education and counseling, 75, 386
- Ashcraft, M. H. 2002, Current Directions in Psychological Science, 11, 181
- Bally, J., O'Dell, C. R., & McCaughrean, M. J. 2000, AJ, 119, 2919

Bardar, E. M., Prather, E. E., Brecher, K., & Slater, T. F. 2006, Astronomy Education Review, 5, 103

Barsony, M., Koresko, C., & Matthews, K. 2003, ApJ, 591, 1064

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. 2010, Proceedings of the National Academy of Sciences, 107, 1860

Bennett, J. 2014, Math for Life: Crucial Ideas You Didn't Learn in School (Big Kid Science)

Betz, N. E. 1978, Journal of counseling psychology, 25, 441

Betz, N. E., & Hackett, G. 1983, Journal of Vocational behavior, 23, 329

Biller, B., et al. 2012, ApJL, 753, L38

Birnstiel, T., Andrews, S. M., & Ercolano, B. 2012, A&A, 544, A79

Birnstiel, T., Dullemond, C. P., & Pinilla, P. 2013, A&A, 550, L8

Boehm, T., & Hirth, G. A. 1997, A&A, 324, 177

Bohlin, R. C., Savage, B. D., & Drake, J. F. 1978, ApJ, 224, 132

Brandt, T. D., et al. 2012, ArXiv e-prints

Brauer, F., Dullemond, C. P., & Henning, T. 2008, A&A, 480, 859

Brown, J. M., et al. 2007, ApJL, 664, L107

Brown, J. M., Blake, G. A., Qi, C., Dullemond, C. P., Wilner, D. J., & Williams, J. P. 2009, ApJ, 704, 496

Brown, J. M., Herczeg, G. J., Pontoppidan, K. M., & van Dishoeck, E. F. 2012a, ApJ, 744, 116

Brown, J. M., Rosenfeld, K. A., Andrews, S. M., Wilner, D. J., & van Dishoeck, E. F. 2012b, ApJL, 758, L30

Bruderer, S., van der Marel, N., van Dishoeck, E. F., & van Kempen, T. A. 2014, A&A, 562, A26

Calvet, N., et al. 2005, ApJL, 630, L185

Calvet, N., Muzerolle, J., Briceño, C., Hernández, J., Hartmann, L., Saucedo, J. L., & Gordon, K. D. 2004, AJ, 128, 1294

Cardelli, J. A., Clayton, G. C., & Mathis, J. S. 1989, ApJ, 345, 245

Casassus, S., Perez M., S., Jordán, A., Ménard, F., Cuadra, J., Schreiber, M. R., Hales, A. S., & Ercolano, B. 2012, ApJL, 754, L31

Casassus, S., et al. 2013, Nature, 493, 191

Chan, L. K., & Dally, K. 2001, Australian Journal of Learning Difficulties, 6, 12

Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, ApJL, 529, L45

Charette, M. F., & Meng, R. 1998, Canadian Journal of Economics, 495

Clarke, C. J., Gendrin, A., & Sotomayor, M. 2001, MNRAS, 328, 485

Close, L. M., et al. 2014, ApJL, 781, L30

Close, L. M., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8447, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

—. 2013, ApJ, 774, 94

—. 1997, ApJ, 489, 210

Corcoran, M., & Ray, T. P. 1997, A&A, 321, 189

Cotera, A. S., et al. 2001, The Astrophysical Journal, 556, 958

Couper, M. P., & Singer, E. 2009, Journal of empirical research on human research ethics: JERHRE, 4, 17

Cushing, M. C., Vacca, W. D., & Rayner, J. T. 2004, PASP, 116, 362

Cutri, R. M., et al. 2003, 2MASS All Sky Catalog of point sources.

Da Rio, N., Robberto, M., Soderblom, D. R., Panagia, N., Hillenbrand, L. A., Palla, F., & Stassun, K. 2009, ApJS, 183, 261

D'Antona, F., & Mazzitelli, I. 1994, ApJS, 90, 467

DeBoer, G. E. 2000, Journal of research in science teaching, 37, 582

Dodson-Robinson, S. E., & Salyk, C. 2011, ApJ, 738, 131

Dokter, E. 2008

Dong, R., et al. 2012, ApJ, 750, 161

Dotter, A., Chaboyer, B., Jevremović, D., Kostov, V., Baron, E., & Ferguson, J. W. 2008, ApJS, 178, 89

Dullemond, C. P., & Dominik, C. 2004, A&A, 421, 1075

—. 2005, A&A, 434, 971

Dullemond, C. P., & Monnier, J. D. 2010, ARA&A, 48, 205

Eisner, J. A., Monnier, J. D., Tuthill, P., & Lacour, S. 2009, ApJL, 698, L169

Ercolano, B., & Owen, J. E. 2010, MNRAS, 406, 1553

Erickson, K. L., Wilking, B. A., Meyer, M. R., Robinson, J. G., & Stephenson,L. N. 2011, AJ, 142, 140

Fedele, D., et al. 2013, A&A, 559, A77

Flaherty, K. M., Muzerolle, J., Rieke, G., Gutermuth, R., Balog, Z., Herbst, W., Megeath, S. T., & Kun, M. 2012, ApJ, 748, 71

Follette, K., & McCarthy, D. 2012, in Astronomical Society of the Pacific Conference Series, Vol. 457, Connecting People to Science: A National Conference on Science Education and Public Outreach, ed. J. B. Jensen, J. G. Manning, M. G. Gibbs, & D. Daou, 295

Follette, K., & McCarthy, D. 2012, Mercury Magazine

Follette, K., & McCarthy, D. 2014, in Astronomical Society of the Pacific Conference Series, Vol. 483, Ensuring Stem Literacy: A National Conference on STEM Education and Public Outreach, ed. J. G. Manning, M. K. Hemenway, J. B. Jensen, & M. G. Gibbs, 31

Follette, K. B., Close, L. M., Kopon, D., Males, J. R., Gasho, V., Brutlag, K. M., & Uomoto, A. 2010, in Society of Photo-Optical Instrumentation

Engineers (SPIE) Conference Series, Vol. 7735, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Follette, K. B., et al. 2013a, ApJL, 775, L13

—. 2014, ArXiv e-prints

—. 2013b, ApJ, 767, 10

Fraknoi, A. 2001, Astronomy Education Review, 1, 121

Fukagawa, M., et al. 2013, PASJ, 65, L14

Geers, V. C., Pontoppidan, K. M., van Dishoeck, E. F., Dullemond, C. P., Augereau, J.-C., Merín, B., Oliveira, I., & Pel, J. W. 2007, A&A, 469, L35

George, D., & Mallery, M. 2003, Boston, MA: Allyn y Bacon.[Links]

Girard, J. H. V., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8447, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Grady, C. A., et al. 2013, ApJ, 762, 48

Graham, J. R., Kalas, P. G., & Matthews, B. C. 2007, ApJ, 654, 595

Green, E. 2014, "Why Do Americans Stink at Math"

Grevesse, N., & Noels, A. 1993, in Origin and Evolution of the Elements, ed. N. Prantzos, E. Vangioni-Flam, & M. Casse, 15–25

Hales, A. S., Gledhill, T. M., Barlow, M. J., & Lowe, K. T. E. 2006, MNRAS, 365, 1348

Hashimoto, J., et al. 2012, ApJL, 758, L19

—. 2011, ApJL, 729, L17

Hayward, T. L., & McCaughrean, M. J. 1997, AJ, 113, 346

Heinemann, T., & Papaloizou, J. C. B. 2009, MNRAS, 397, 64

Henderson, C., & Dancy, M. H. 2008, American Journal of Physics, 76, 79

Henry, G. W., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2000, ApJL, 529, L41

Henyey, L. G., & Greenstein, J. L. 1941, ApJ, 93, 70

Hester, S., Buxner, S., Elfring, L., & Nagy, L. 2014, CBE-Life Sciences Education, 13, 54

Hillenbrand, L. A. 1997, AJ, 113, 1733

Hinkley, S., et al. 2009, ApJ, 701, 804

Hodapp, K. W., et al. 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6269, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Honda, M., et al. 2009, ApJL, 690, L110

Hughes, A. M., et al. 2009, ApJ, 698, 131

Isella, A., Pérez, L. M., & Carpenter, J. M. 2012, ApJ, 747, 136

Jackson, C. D., & Leffingwell, R. J. 1999, The Mathematics Teacher, 583

Jang-Condell, H. 2009, ApJ, 700, 820

Jang-Condell, H., & Boss, A. P. 2007, ApJL, 659, L169

Jang-Condell, H., & Turner, N. J. 2012, ApJ, 749, 153

Kandori, R., et al. 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6269, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Kastner, J. H., Franz, G., Grosso, N., Bally, J., McCaughrean, M. J., Getman, K., Feigelson, E. D., & Schulz, N. S. 2005, ApJS, 160, 511

Kenyon, S. J., & Hartmann, L. 1995, ApJS, 101, 117

Kim, S.-H., Martin, P. G., & Hendry, P. D. 1994, ApJ, 422, 164

Kirsch, I. S., et al. 1993, Adult Literacy in America: A First Look at the Results of the National Adult Literacy Survey. (ERIC)

Konacki, M., Torres, G., Jha, S., & Sasselov, D. D. 2003, Nature, 421, 507

Kopon, D., Close, L. M., Males, J., Gasho, V., Morzinski, K., & Follette, K. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8447, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Kopon, D., Males, J., Close, L. M., & Gasho, V. 2009, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7439, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Kraus, A. L., & Ireland, M. J. 2012, ApJ, 745, 5

Lafrenière, D., Marois, C., Doyon, R., Nadeau, D., & Artigau, É. 2007, ApJ, 660, 770

Lawrenz, F., Huffman, D., & Appeldoorn, K. 2005, Journal of College Science Teaching, 34, 40

Loinard, L., Torres, R. M., Mioduszewski, A. J., & Rodríguez, L. F. 2008, ApJL, 675, L29

Lombardi, M., Lada, C. J., & Alves, J. 2008, A&A, 480, 785

Lucy, L. B. 1999, A&A, 344, 282

Lynden-Bell, D., & Pringle, J. E. 1974, MNRAS, 168, 603

Lyra, W., Johansen, A., Klahr, H., & Piskunov, N. 2009, A&A, 493, 1125

Lyra, W., & Kuchner, M. 2013, Nature, 499, 184

Maaskant, K. M., et al. 2013, A&A, 555, A64

Maaskant, K. M., Min, M., Waters, L. B. F. M., & Tielens, A. G. G. M. 2014, A&A, 563, A78

Macintosh, B., et al. 2014, ArXiv e-prints

Madison, B. L., & Steen, L. A. 2003, Quantitative literacy: Why numeracy matters for schools and colleges (Woodrow Wilson Natl Foundation)

Males, J. R., et al. 2012, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 8447, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Mamajek, E. E. 2008, Astronomische Nachrichten, 329, 10

Mann, R. K., & Williams, J. P. 2010, ApJ, 725, 430

Marois, C., Lafrenière, D., Doyon, R., Macintosh, B., & Nadeau, D. 2006, ApJ, 641, 556

Mathis, J. S. 1990, ARA&A, 28, 37

Mayama, S., et al. 2012, ApJL, 760, L26

Mayor, M., & Queloz, D. 1995, Nature, 378, 355

McCarthy, D., & Follette, K. 2013, in Astronomical Society of the Pacific Conference Series, Vol. 473, Communicating Science: A National Conference on Science Education and Public Outreach, ed. J. Barnes, C. Shupla, J. G. Manning, & M. G. Gibbs, 79

McCaughrean, M. J., & O'dell, C. R. 1996, AJ, 111, 1977

McClure, M. 2009, ApJL, 693, L81

McClure, M. K., et al. 2010, ApJS, 188, 75

Menten, K. M., Reid, M. J., Forbrich, J., & Brunthaler, A. 2007, A&A, 474, 515

Menu, J., et al. 2014, A&A, 564, A93

Minowa, Y., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7736, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Muench, A. A., Lada, E. A., Lada, C. J., & Alves, J. 2002, ApJ, 573, 366

Muto, T., et al. 2012, ApJL, 748, L22

Muzerolle, J., Hartmann, L., & Calvet, N. 1998, AJ, 116, 2965

Natta, A., Testi, L., & Randich, S. 2006, A&A, 452, 245

Nicholson, S., & Mulvey, P. 2010, Statistical Research Center of the American Institute of Physics

O'dell, C. R., & Wen, Z. 1994, ApJ, 436, 194

O'dell, C. R., & Wong, K. 1996, AJ, 111, 846

Ogbu, J. U. 1990

Oppenheimer, B. R., et al. 2008, ApJ, 679, 1574

Owen, J. E., Clarke, C. J., & Ercolano, B. 2011, ArXiv e-prints

Palla, F., & Stahler, S. W. 1993, ApJ, 418, 414

—. 1999, ApJ, 525, 772

Panić, O., et al. 2013, MNRAS

Pascucci, I., & Sterzik, M. 2009, ApJ, 702, 724

Pascucci, I., et al. 2011, ApJ, 736, 13

Pereyra, A., Girart, J. M., Magalhães, A. M., Rodrigues, C. V., & de Araújo, F. X. 2009, A&A, 501, 595

Pérez, L. M., Isella, A., Carpenter, J. M., & Chandler, C. J. 2014, ApJL, 783, L13

Perez-Pena, R. 2013, US Adults Fare Poorly in a Study of Skills

Perrin, M. D., et al. 2014, ArXiv e-prints

Perrin, M. D., et al. 2010, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 7736, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, 5

Perrin, M. D., Graham, J. R., & Lloyd, J. P. 2008, PASP, 120, 555

- Perrin, M. D., Schneider, G., Duchene, G., Pinte, C., Grady, C. A., Wisniewski, J. P., & Hines, D. C. 2009, ApJL, 707, L132
- Peters, E., Hibbard, J., Slovic, P., & Dieckmann, N. 2007, Health Affairs, 26, 741
- Pollack, J. B., Hubickyj, O., Bodenheimer, P., Lissauer, J. J., Podolak, M., & Greenzweig, Y. 1996, Icarus, 124, 62
- Pontoppidan, K. M., Blake, G. A., van Dishoeck, E. F., Smette, A., Ireland, M. J., & Brown, J. 2008, ApJ, 684, 1323
- Poyneer, L., van Dam, M., & Véran, J.-P. 2009, Journal of the Optical Society of America A, 26, 833
- Prato, L., Greene, T. P., & Simon, M. 2003, ApJ, 584, 853
- Quanz, S. P., Schmid, H. M., Geissler, K., Meyer, M. R., Henning, T., Brandner, W., & Wolf, S. 2011, ApJ, 738, 23
- Ragazzoni, R. 1996, Journal of Modern Optics, 43, 289
- Rameau, J., Chauvin, G., Lagrange, A.-M., Thébault, P., Milli, J., Girard, J. H., & Bonnefoy, M. 2012, A&A, 546, A24
- Ratzka, T., Leinert, C., Henning, T., Bouwman, J., Dullemond, C. P., & Jaffe, W. 2007, A&A, 471, 173
- Raudenbush, S. W., & Kasim, R. M. 1998, Harvard Educational Review, 68, 33
- Rayner, J. T., Toomey, D. W., Onaka, P. M., Denault, A. J., Stahlberger, W. E., Vacca, W. D., Cushing, M. C., & Wang, S. 2003, PASP, 115, 362

Ricci, L., Robberto, M., & Soderblom, D. R. 2008, AJ, 136, 2136

Rigliaco, E., Pascucci, I., Gorti, U., Edwards, S., & Hollenbach, D. 2013, ApJ, 772, 60

Robitaille, T. P., Whitney, B. A., Indebetouw, R., & Wood, K. 2007, ApJS, 169, 328

Robitaille, T. P., Whitney, B. A., Indebetouw, R., Wood, K., & Denzmore, P. 2006, ApJS, 167, 156

Rodigas, T. J., et al. 2014, ApJ, 783, 21

—. 2012, ApJ, 752, 57

Roeser, S., Demleitner, M., & Schilbach, E. 2010, AJ, 139, 2440

Rosen, L. P., Weil, L., & von Zastrow, C. 2003, Quantitative literacy: Why numeracy matters for schools and colleges, 43

Rousselle, L., & Noël, M.-P. 2007, Cognition, 102, 361

Rutherford, F. J., & Ahlgren, A. 1991, Science for all Americans (Oxford university press)

Salyk, C., Blake, G. A., Boogert, A. C. A., & Brown, J. M. 2009, ApJ, 699, 330

Salyk, C., Herczeg, G. J., Brown, J. M., Blake, G. A., Pontoppidan, K. M., & van Dishoeck, E. F. 2013, ApJ, 769, 21

Schlingman, W. M., Prather, E. E., Wallace, C. S., Rudolph, A. L., & Brissenden, G. 2012, Astronomy Education Review, 11, 010107

Schneps, L., & Colmez, C. 2013, Math on trial: How numbers get used and abused in the courtroom (Wiley Online Library)

Schwartz, L. M., Woloshin, S., Black, W. C., & Welch, H. G. 1997, Annals of internal medicine, 127, 966

Seife, C. 2010, Proofiness: How You're Being Fooled by the Numbers (Penguin)

Shakura, N. I., & Sunyaev, R. A. 1973, A&A, 24, 337

Shang, H., Li, Z.-Y., & Hirano, N. 2007, Protostars and Planets V, 261

Siess, L., Dufour, E., & Forestini, M. 2000, A&A, 358, 593

Simpson, J. P., Burton, M. G., Colgan, S. W. J., Cotera, A. S., Erickson, E. F., Hines, D. C., & Whitney, B. A. 2009, ApJ, 700, 1488

Skrutskie, M. F., Dutkevitch, D., Strom, S. E., Edwards, S., Strom, K. M., & Shure, M. A. 1990, AJ, 99, 1187

Soummer, R., Pueyo, L., & Larkin, J. 2012, ApJL, 755, L28

Spencer, S. J., Steele, C. M., & Quinn, D. M. 1999, Journal of experimental social psychology, 35, 4

Steele, C. M., & Aronson, J. 1995, Journal of personality and social psychology, 69, 797

Steen, L. A. 1999, Educational Leadership, 57, 8

—. 2001, Princeton NJ

Strom, K. M., Strom, S. E., Edwards, S., Cabrit, S., & Skrutskie, M. F. 1989, AJ, 97, 1451

Takami, M., et al. 2013, ApJ, 772, 145

Tamura, M. 2009, in American Institute of Physics Conference Series, Vol.1158, American Institute of Physics Conference Series, ed. T. Usuda,M. Tamura, & Eamp; M. Ishii, 11–16

Tamura, M., et al. 2006, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 6269, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Terada, H., & Tokunaga, A. T. 2012, ApJ, 753, 19

Terada, H., et al. 2012, AJ, 144, 175

Thalmann, C., et al. 2010, ApJL, 718, L87

Vacca, W. D., Cushing, M. C., & Rayner, J. T. 2003, PASP, 115, 389

van der Marel, N., et al. 2013, Science, 340, 1199

van der Marel, N., van Dishoeck, E. F., Bruderer, S., & van Kempen, T. A. 2014, A&A, 563, A113

van der Plas, G., van den Ancker, M. E., Fedele, D., Acke, B., Dominik, C., Waters, L. B. F. M., & Bouwman, J. 2008, A&A, 485, 487

van Kempen, T. A., van Dishoeck, E. F., Salter, D. M., Hogerheijde, M. R., Jørgensen, J. K., & Boogert, A. C. A. 2009, A&A, 498, 167

Veran, J.-P., Rigaut, F., Maitre, H., & Rouan, D. 1997, Journal of the Optical Society of America A, 14, 3057

Weingartner, J. C., & Draine, B. T. 2001, ApJ, 548, 296

White, R. L. 1979, ApJ, 229, 954

Whitney, B. A., & Hartmann, L. 1992, ApJ, 395, 529

Whitney, B. A., Robitaille, T. P., Bjorkman, J. E., Dong, R., Wolff, M. J., Wood, K., & Honor, J. 2013, ApJS, 207, 30

Whitney, B. A., & Wolff, M. J. 2002, ApJ, 574, 205

Whitney, B. A., Wood, K., Bjorkman, J. E., & Cohen, M. 2003a, ApJ, 598, 1079

Whitney, B. A., Wood, K., Bjorkman, J. E., & Wolff, M. J. 2003b, ApJ, 591, 1049

Wiktorowicz, S. J., et al. 2014, in Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, Vol. 9147, Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series

Wilking, B. A., Gagné, M., & Allen, L. E. 2008, Star Formation in the ρ Ophiuchi Molecular Cloud, ed. B. Reipurth, 351

Wilking, B. A., Lada, C. J., & Young, E. T. 1989, ApJ, 340, 823

Wilking, B. A., Meyer, M. R., Robinson, J. G., & Greene, T. P. 2005, AJ, 130, 1733

Williams, J. P., & Cieza, L. A. 2011, ARA&A, 49, 67

Wood, K., Wolff, M. J., Bjorkman, J. E., & Whitney, B. 2002, ApJ, 564, 887

Wright, E. L., et al. 2010, AJ, 140, 1868

- Yamamura, I., Makiuti, S., Ikeda, N., Fukuda, Y., Oyabu, S., Koga, T., & White, G. J. 2010, VizieR Online Data Catalog, 2298, 0
- Zacharias, N., Monet, D. G., Levine, S. E., Urban, S. E., Gaume, R., & Wycoff, G. L. 2005, VizieR Online Data Catalog, 1297, 0
- Zhu, Z., Nelson, R. P., Dong, R., Espaillat, C., & Hartmann, L. 2012, ApJ, 755, 6
- Zhu, Z., Nelson, R. P., Hartmann, L., Espaillat, C., & Calvet, N. 2011, ApJ, 729, 47
- Zuckerman, B. 2001, ARA&A, 39, 549