TECHNIQUE DEVELOPMENT NEEDED TO INTERPRET RAMAN SPECTRA OF MINERAL MIXTURES. M. D. Dyar[1], L. B. Breitenfeld[1], P. Bartholomew[2], CJ Carey[3]. [1]Mount Holyoke College, Dept. of Astronomy, South Hadley, MA, 01075, mdyar@mtholyoke.edu, Department of Biology and Environmental Science, University of New Haven, West Haven, CT 06516, [3]College of Information and Computer Science, University of Massachusetts, Amherst, MA 01003.

Introduction: Planetary scientists who use reflectance and thermal emission spectroscopy are spoiled by the availability, breadth, and sophistication of unmixing software and spectral libraries available for commonly used techniques. Many techniques for extracting quantitative abundances of minerals from visible through mid-infrared (and thermal emission) spectra of mineral mixtures are available [1-3]. *However, there is no analogous methodology for mineral unmixing for the technique of Raman spectroscopy* because most development work has been focused on micro-Raman techniques [4] that probe pure minerals. Because the beam sizes of the Raman instruments on SuperCam (~1.3 mm) [5] and SHERLOC (50 μm) [6] on *Mars 2020* are larger than most expected grain sizes on Mars, it is likely that their Raman spectra will contain mixtures of mineral signatures. Other impending possible uses of Raman spectroscopy, such as on Europa and Venus landers, may also face this issue. An infusion of work is needed to support these instruments.

Three interrelated areas of research are reported here: 1) integration and review of new and existing mineral databases of well-characterized mineral samples, 2) acquisition of quantitative data to assess the relationship between Raman peak intensity and mineral abundance, and 3) development of mineral identification software with quantified accuracy.

Databases: The importance of robust Raman databases of appropriate well-characterized minerals cannot be overstated. In addition, spectra of single crystals may not be appropriate for identification (or quantification) of fine grained mineral mixtures and/or powders. Raman spectral data for powders and mixtures of minerals are nearly nonexistent and, while work has already begun to remedy this [7], the necessary investment of time and effort needed is significant.

While a number of public/internet Raman spectral databases exist, the largest, by far, is the RRUFF spectral database (http://rruff.info) created by Bob Downs and the University of Arizona mineralogy group through the support of both private donors and federal research grants [8,9]. While every spectrum in these databases has scientific value, not every spectrum meets the criteria to be included in a reference dataset for mineral identification [10]. These criteria include:

1. Use only spectra from verified samples – ideally with both structure and chemistry verified.
2. Strong, clear spectra are preferred, so a minimum S/N ratio must be set.
3. Use spectra that are free of artifacts – for Raman the concern is features due to photo-fluorescence.

Because many of the >4000 known mineral species are truly rare, an effective reference database for petrologic mineral identification also needs to contain flags to distinguish rock-forming minerals from rare ones, while still maintaining thorough coverage of rock-forming minerals. To this end, we are assembling and reviewing publicly-available Raman data from RRUFF and other on-line datasets to filter out weaker (low S/N) spectra, those contaminated by fluorescence, and from unverified samples. We are also evaluating gaps in the resulting dataset that are relevant to rock-forming and planetary minerals. We are obtaining Raman spectra on well-characterized samples in order to fill those gaps. Particularly important are pyroxenes and olivines. We plan to acquire data on samples synthesized by Don Lindsley at Stony Brook University that span the Ca-Fe-Mg pyroxene quadrilateral and the fayalite-forsterite solid solution series.

Finally, because fluorescence peaks are often not inherent characteristics of the individual minerals and have the potential to confuse matching algorithms, we plan to study mineral fluorescence under experimental conditions relevant to Raman spectrometry. Although the use of time-resolved Raman in some planetary missions will exclude fluorescence effects, nearly all existing spectral libraries (to which they will be compared) contain data with this problem.

To address the lack of data on fine-grained minerals, we plan to collect data on a petrologically relevant subset of >10,000 powdered minerals available in our laboratory as funding allows. We are also creating fine-grained mixtures of mineral and glass phases for Raman testing, including 256 binary mineral mixtures described in [11,12] and 140 glass-mineral mixtures designed to test detection limits.

Understanding Raman Peak Intensities: It is well-known that experimental factors affect Raman peak intensities (and thus mineral identification). Spectra of the same mineral species in existing datasets show variations in peak presence/absence and relative intensity. These can result from sample factors (grain size, transparency, crystallographic orientation, grain surface/boundary effects [4]), instrument factors (laser wavelength, power, and spot size, spectrometer apertures, gratings, and detectors), experimental factors

(angle of incidence and takeoff and the use/absence of polarizers) and data gathering factors (integration time, averaging, method/frequency of calibration). Some differences in peak intensities can be mitigated through pre-processing steps such as normalization, smoothing, and squashing [13,14]. We are also working on a calibration transfer methodology that will correct for many of these instrumental and experimental factors [15].

Raman cross-section of minerals is the primary *species-specific* intensity factor; it arises primarily from variations in bond polarizability in mineral structures. Mineral-specific proportionality factors (comparable to optical constants used to relate peak intensities to mineral abundances) of common rock-forming minerals are needed. These can be determined from calculations of remote Raman efficiency based on laboratory measurements [16] or by comparing peak intensities of minerals mixed with known amounts of a standard such as diamond [12]. Of these, the latter is more generally useful and easier to implement in practice. We plan to acquire those data in our laboratory for hundreds of mineral species before the *Mars 2020* landing.

Matching and Unmixing of Raman Data: Traditional Raman matching software packages, including the CrystalSleuth program [17], have been found to have limited accuracy, especially at the species level. However, machine learning techniques using automated whole spectrum matching (AWSM) [14] dramatically improve matching accuracy for Raman spectra of pure minerals. For example, a weighted-neighbor cosine similarity classifier [13] achieved 97.8% group-level and 89.2% species-level accuracy on average for a subset of RRUFF data, far outperforming existing methods for matching.

Moving forward, unmixing approaches that build on techniques currently used for FTIR should provide a starting point for Raman theoretical umixing models. Raman features arise from scattering of energy while FTIR spectra have absorption features. Raman provides information on covalency of molecular bonds, while FTIR indicates ionic character. Raman peaks reflect changes in bond polarization, while FTIR peaks record dipole changes. These commonalities suggest the potential of future unmixing algorithms to assist with interpretation of Raman data, but that promise is years away from becoming a reality.

As an alternative, our AWSM algorithms are now being adapted for the purpose of mineral unmixing. However, determining the components is one problem, but quantifying them is another. Consider the mixtures of minerals in known proportions shown in **Figure 1**. These are Raman data of three mixtures, all with 20 volume% gypsum and 80 volume % of another common phase (forsterite, siderite, or labradorite). In these

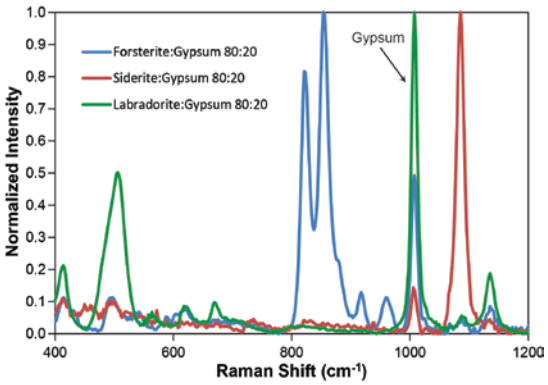

Figure 1. Mixtures of 20% gypsum with 80 volume% of other phases (pyroxene, carbonate, and feldspar). Note the large variations in the intensities of the gypsum peak ca. 1010 cm^{-1} despite its constant modal abundance.

mixtures, the amount of gypsum is identical but the magnitudes of the gypsum peaks are highly variable. This is emblematic of problems to be faced in quantifying mineral abundances using Raman spectra. We expect that Raman proportionality factors measured against our diamond internal standard will provide an empirical solution for this challenge.

Resources: Current models for Raman mineral matching and unmixing of up to three components [18] are available on our web site for beta-testing at http://nemo.cs.umass.edu:54321/. Ongoing development work and new data will be posted there for interested users. Please contact the first author for full site access.

Acknowledgments: This work was supported by NASA SSERVI grant NNA14AB04A (RIS^{4}E) and the Massachusetts Space Grant Consortium.

References: [1] Rull F. et al. (2014) *8th Intl. Mars Conf.*, Abstr. #1277. [2] Beegle L. W. et al. (2014) *11th Intl. GeoRaman Conf.*, Abstr. #5101. [3] Maurice S. et al. (2015) *LPS XLVI*, Abstr. #2818. [4] Haskin L. A. et al. (1997) *JGR*, 102, 19293-19306. [5] Clegg S. M. et al. (2015) *LPS XLVI*, Abstr. #2781. [6] Beegle L. et al (2014) *LPS XLV*, Abstr. #2835. [7] Dyar M.D. et al. (2016) *LPS XLVII*, Abstr. #2240 [8] LaFuente B. et al. (2015) *Highlights Mineral. Crystallography*, pp. 1-29. [9] Stone N. et al. (2015) *AbSciCon*, Chicago, IL. [10] Bartholomew P.R. et al. (2014) *GSA Abstr.* 46, 321. [11] Breitenfeld et al. (2016) *LPSXLVII*, Abstr. #2430. [12] Breitenfeld et al. (2016) *LPS XLVII*, Abstr. #2186. [13] Carey C. et al. (2015) *J. Raman Spectrosc.*, 46, 894-903. [14] Carey C. et al. (2015) *AI in Space, IJCAI 2015*. [15] Boucher, T. et al. (2016) *LPS XLVII*, Abstr. #2784. [16] Stopar J.D. et al. (2005) *Spectrochim. Acta A*, 61, 2315-2323. [17] Laetsch T. and Downs, R.T. (2006) *19th General IMA Meeting*, Kobe, Japan, 23–28. [18] C. Carey et al. (2016) *LPS XLVII*, Abstr. #2626.