Learning Benefits of Integrating Socioeconomic and Cultural Considerations into an Onsite Water Reclamation Course Project

Lt. Col. Andrew Ross Pfluger, Colorado School of Mines

Lt Colonel Andrew Pfluger, U.S. Army, is an active duty officer and an Assistant Professor at the United States Military Academy. He served in the Department of Geography and Environmental Engineering from 2010 to 2013 and is a faculty member in the Department of Chemistry and Life Science as of June 2018. He earned a B.S. in Civil Engineering from USMA, a M.S. and Engineer Degree in Environmental Engineering and Science from Stanford University, and is on track to earn a PhD from the Colorado School of Mines in the spring of 2018. He is a licensed PE in the state of Delaware.

Dr. Junko Munakata Marr, Colorado School of Mines

Dr. Munakata Marr is an Associate Professor of Civil and Environmental Engineering at the Colorado School of Mines in Golden, Colorado. She received her BS degree in Chemical Engineering from the California Institute of Technology and her MS and PhD degrees in Civil Engineering from Stanford University. Her research and teaching interests revolve primarily around microorganisms in engineered environmental systems, including biological wastewater treatment and methanogenesis from unconventional sources. She has nearly 20 years of experience in bioremediation. Other interests include sustainable water infrastructure, increasing diversity among STEM students and faculty, and sustainable community development.

Dr. Gary Vanzin
Dr. Robert L. Siegrist, Colorado School of Mines

Robert L. Siegrist, PhD, PE, BCEE is University Professor Emeritus of Environmental Science and Engineering at the Colorado School of Mines and also a Professor of Civil and Environmental Engineering. He is the former Director of the Environmental Science and Engineering Division at CSM and founding Director of the Small Flows Program. Before joining CSM in 1995, Prof. Siegrist held academic and research positions with the University of Wisconsin, Norwegian Institute for Georesources and Pollution Research, Ayres Associates Inc., and Oak Ridge National Laboratory. He earned his BS (High Honors) and MS in Civil Engineering and his PhD in Environmental Engineering at the University of Wisconsin where he conducted research within the Small Scale Waste Management Project. Prof. Siegrist is an internationally recognized expert in decentralized water reclamation and in situ remediation of contaminated land. During his 40-year career he has published over 300 technical papers and 3 books and was awarded 2 patents. His new textbook, Decentralized Water Reclamation Engineering, was just published by Springer (www.springer.com/us/book/9783319404714). He has given invited keynote lectures at more than 100 workshops and conferences in more than 30 countries worldwide. He has also served as a science and engineering advisor to private and government agencies and organizations in the United States and abroad and has been a Fellow with the NATO Committee for Challenges to Modern Society.

©American Society for Engineering Education, 2018

Learning benefits of integrating socio-economic and cultural considerations into an onsite water reclamation course project

Abstract

During the past decade, our university has offered a senior undergraduate/graduate-level course that focuses on onsite water reclamation covering the selection, design, and implementation of onsite and decentralized treatment systems. A major element used to assess student learning is a culminating project that asks students to critically review an onsite water reclamation or reuse technology, identify lessons learned from an onsite case study, or design an onsite treatment system for a specific application. During course deliveries in 2014 and before, non-technical considerations focused on regulatory requirements and project owner needs. In 2016 and 2017, a different instructor integrated socio-economic and cultural considerations, through course content focused on onsite water, sanitation, and hygiene (WaSH) efforts in developing countries, as a major course theme. To assess whether students valued the integration of non-technical considerations, 22 student projects spanning a period of four academic semesters between 2014 and 2017 were analyzed using two approaches. Projects were analyzed (1) for the degree of integration of non-technical considerations and (2) by term frequency mining and term frequency-inverse document frequency (tf-idf). The integration of socio-economic and cultural considerations into the course project increased in 2016 and 2017, with five of twelve student teams in 2014 and eight of ten student teams in 2016 and 2017 integrating non-technical considerations in their analysis. Gender demographics and graduate standing were not correlated with the degree of integration of non-technical considerations. Term frequency analysis and tf-idf showed that key terms in the “social” and “energy” categories were used significantly more in the 2016 and 2017 course projects, while use of technical terms did not change. Increasing emphasis and content of nontechnical concepts through integration of WaSH principles appears to have enhanced student consideration of these concepts while maintaining technical content.

1. Introduction

Appropriate and sustainable solutions to water and sanitation need to be effective while being affordable, socially acceptable, and sustainable. In the U.S. and most industrialized nations, centralized treatment and distribution/collection approaches are common for urban areas and population centers. However, in many situations within the U.S. and in developing countries, such practices are neither cost-effective nor sustainable for a number of reasons, including low-density development, rugged topography, and cultural constraints, among others. In these instances, onsite or satellite water treatment systems can be used to protect public health and environmental quality while having low energy and chemical consumption and enabling beneficial reuse of water and nutrients (Siegrist, 2017). In addition to technical considerations, non-technical factors such as socio-economic, political, and regulatory issues need to be considered in developing holistic solutions that are acceptable to the users. This can be especially true in developing countries, and in contexts where WaSH (Water, Sanitation, and Hygiene) interventions are required (UNICEF, 2015).

Our university offers a semester-long course to educate students on onsite water treatment topics, including the selection, design, and implementation of onsite and distributed (hereafter referred to as decentralized) treatment systems. The course was developed over a period of several years and first offered to senior undergraduates and graduates in 2006 (Siegrist, 2014). During course deliveries in 2014 and before, non-technical considerations incorporated into the course principally focused on regulatory requirements and project owner requirements. However, since its development, the course has been included as an elective in our university's Humanitarian Engineering minor program. In 2016 and 2017, a different course instructor integrated other non-technical themes, specifically socio-economic, political, and cultural considerations, throughout the course content with an emphasis on WaSH efforts in developing countries.

To evaluate the impact of integrating the aforementioned non-technical considerations into the course, term frequency analyses were employed. Term extraction, or term mining, is a technique for identifying relevant terms and the frequency they occur within a given corpus, or collection, of documents for further analysis (Jacquemin & Bourigault, 2005). Term frequency analyses used in engineering education to identify and examine significant words to student education include simple term counts or term frequency-inverse document frequency (tf-idf), amongst others (Chen et al., 2011). Tf-idf is one of the most commonly used approaches for term weighting (Aizawa, 2003) and is described further in Section 4. Such term frequency analyses have been used in numerous educational contexts, from examination of Master's thesis and PhD dissertations to identify common topics (Rivera & Larrondo-Petrie, 2017), to engineering term language gaps between professors and students in freshman-level engineering courses as a barrier to learning (Variawa et al., 2013; Variawa & McCahan, 2012).

2. Course Description and Student Demographics

Our university's onsite course focuses on the selection, design, and implementation of decentralized systems for water reclamation and reuse. Topics include process analysis and system planning, engineered and natural system treatment units, alternative collection systems, and effluent dispersal and reuse options. During the course, students are expected to demonstrate that they can: (1) describe multiple treatment units, collection approaches, and effluent dispersal and reuse options for onsite/decentralized sanitation; (2) evaluate advantages and disadvantages of options; (3) identify and use relevant design equations; (4) consider socio-cultural contexts; and (5) work in teams to write a project report and present their findings.

In 2016, to help students think holistically about the integration of decentralized technologies in a variety of contexts, a new instructor placed increased emphasis on integrating non-technical concepts into the course curriculum, to include socio-economic considerations, cultural considerations, and political considerations, beyond regulatory and project owner considerations that already were part of the course curriculum. The new instructor also emphasized concepts related to WaSH in developing countries, including an overview of low-complexity sanitation systems. The added topics, reading assignments, discussion and session objectives are summarized in Table 1.

Classes included a mix of graduate students and undergraduates. All student demographics during the study period are described in Table 2.

3. Description of Course Project

The course project was designed for students to demonstrate their synthesis of material learned throughout the course. Students were given three example project types, although teams were told they could propose a different type of project if they desired. Project types included: (1) technical analysis of a specific device or technology; (2) a review of lessons learned from case studies of decentralized system applications; (3) design of a decentralized system for a specific project. The project requirements were the same for all semesters examined in this study; however, for courses taught in 2016 and 2017, students were specifically encouraged to consider non-technical aspects of their selected project (type 1, 2 or 3), as appropriate.

Students self-organized into teams of one to five students to prepare a project report less than 20 pages in length. In the report, students were instructed to describe the scope of their project, clearly and concisely present their evaluation and analysis, and to provide recommendations for implementation and future work. Student teams were required to meet several mid-semester milestones, to include a mid-course project proposal (due six weeks into the course), a description of a chosen project (due nine weeks into the course), a report outline (due 11 weeks into the course), and a draft bibliography (due 14 weeks into the course).

The projects were graded on the following criteria in 2016 and 2017: (1) clarity of the context and stated purpose (13.3%); (2) technical evaluation with supporting tables and figures (26.6%); (3) consideration of non-technical (social, cultural, economic) factors (13.3%); (4) discussion of implications of the work, as well as issues and constraints (13.3%); (5) clarity of conclusions and recommendations (13.3%); (6) quality of references (6.7%); (7) project formatting, spelling, grammar, and editorial quality (13.3%). Grading criteria were similar in 2014; however, non-technical factors were not graded.

4. Methods of Assessment

To access the effect of integrating more emphasis on current and additional non-technical concepts into the course, the project reports were assessed using two methods. The first assessment method involved a survey of course projects and project assessment sheets for non-technical content. After the level of non-technical content was assessed, reports were binned into four categories: (1) no non-technical integration; (2) brief discussion of non-technical content; (3) substantial consideration of non-technical content; and (4) primary focus on non-technical concepts. Results from the assessment were aggregated and assessed by term the course was taught (i.e., pre-intervention (terms 14-1 and 14-2), and post-intervention (terms 16-2 and 17-2)). Terms 14-1 and 14-2 were representative of earlier course deliveries and were used as a reference for assessing the integration of non-technical concepts in 2016 and 2017. The impact of team demographics on the non-technical content was also evaluated. Specifically, the level of non-technical content for each project was compared to the ratio of male-to-female students and the ratio of graduate-to-undergraduate students per project group and subjected to regression statistics to determine statistical significance.

The second method of assessment involved term frequency analyses. Term document matrices were generated in R version 3.4.2. All course project reports were first merged into a single corpus using R library pdftools (Ooms, 2017). Two analyses were then conducted using the TermDocumentMatrix function within the R library tm version 0.7-3: (1) term frequency and (2) term frequency-inverse document frequency (tf-idf) (Feinerer & Hornik, 2017). Parameters included converting all letters to lower-case, as well as removing stop words, numbers, and punctuation. Common words, such as “the” and “and”, were eliminated during the term mining process. While term frequency counts each individual word occurrence and bins them by document, tf-idf is a statistical method that reflects how important a word is to a particular document within a corpus of documents (e.g., term project reports). More specifically, tf-idf determines the relative frequency of terms in an individual document compared to the inverse proportion of that word over the entire document corpus (Blei et al., 2003; Robertson, 2004). For both analyses, a minimum threshold of five occurrences within the 22 project reports was established for term examination. Relevant terms were binned into six categories for analysis: (1) social, (2) sustainability, (3) WaSH, (4) economic, (5) energy, and (6) technical (not energy-

Table 1. Summary of course content added in 2016 to increase emphasis on WaSH.

Class session	Reading	Discussion	Objectives
Water, sanitation and hygiene overview	WHO/UNICEF Joint Monitoring Programme update, focus on pp. 2-3, 5, 15-25	• What is WaSH? • In groups of 4-5, script race to define each of these terms • Why are these lumped together?	Define WaSH, goals
WaSH, continued	UN Water Investing in Water and Sanitation: Increasing Access, Reducing Inequalities	Student-centered discussion on specific WaSH efforts, stakeholders 1. What is JMP? What is WHO? UNICEF? 2. Who works on WaSH? 3. What is your role, as engineers?	Identify who works on sanitation. Consider perspectives of different stakeholders
Overview of sanitation systems	Eawag Sanitation Systems and Technologies user interfaces, collection and storage/treatment (3.1-3.4) conveyance and treatment (3.5-3.11)	1. Who is Eawag? 2. What is sanitation? 3. What is environmental sanitation? 4. What kinds of interfaces, collection, etc. do we use in US? Interfaces—options? 5. Sewers: learned about in other classes? 6. What design guidelines are specified? 7. Planning and design—intro to next week's topic	Provide overview of sanitation systems (Eawag sanitation system drawing tool) (Eawag/Sandec, 2008)
Overview of sanitation systems, continued	non-technical aspects (4)	This reading is focused on developing country scenarios. What are the broad categories of considerations? Which of the **technical and physical** criteria are also relevant in developed countries? **Socio-cultural** aspects? **Political and institutional**? **Financial and economic**? **Financing schemes**?	Appreciate that considerations are common across scenarios
System planning, analysis and selection	Mihelcic, et al. (2009) Field Guide Ch 3	Planning exercise to evaluate sanitation options for a local community: Formulate an approach to conducting this evaluation. What information would you gather, from where/whom? Why should you be hired? (ignoring schedule and cost)	Describe methodical approach to project design and management
Latrines	Mihelcic et al. Field Guide Ch. 20 through	Overview of latrines, pros & cons Design challenge for latrines in different settings	

	Ch. 20.2.2 (up to p. 394)		
Composting toilets	Mihelcic et al. Field Guide Ch. 20.2.3 on (p. 394+)	Education: community-dependent; community/school classes; latrine info (materials, maintenance, different types, construction); cost; accountability plan; games/competition; partner with community members Feasibility: stakeholder mapping; cost/benefit analysis; materials/construction; economic/business development opportunities; behavior change; technical considerations Maintenance: incentives to attend classes; job creation or rotation plan; set definite dates; materials easily accessible; periodic check-in, accountability	

Table 2. Report titles by semester, student demographics, and binned category. Graduate-to-undergraduate ratios and gender ratios (male-to-female) are displayed, Categories were weighted: “technical only” = 1; “non-technical discussed” = 2; “non-technical considered” = 3; “non-technical only” = 4.

Semester of Study	Report #	Title	Graduate-to-Undergraduate	Male-to-Female	Binned Category
Term 14-1 Pre-intervention	1	Anaerobic Treatment and Nitrogen Management Livestock Waste from Small Farming Operations	2:0	1:1	Technical Only
	2	Passive onsite and natural water treatment train design for kitchen graywater effluent from Butare Central Prison, Rwanda	3:0	2:1	Non-technical Discussed
	3	Chlorine, UV and Ozone Disinfection Technologies for Onsite and Decentralized Wastewater Treatment	1:0	0:1	Technical Only
	4	Developing Nations: A look into onsite and decentralized systems	0:3	2:1	Non-technical Discussed
	5	Living Machine for Wastewater Treatment for New Apartment Complex in Tucson, AZ	2:0	1:1	Technical Only
	6	Membrane Bioreactors and Domestic Wastewater	2:0	1:1	Non-technical Discussed
	7	Decentralized System Design for Pine Ridge, South Dakota	2:2	2:2	Non-technical Considered

Term 14-2 Pre-intervention	8	Lessons Learned from Case Studies of Decentralized System Applications of Constructed Treatment Wetlands	3:0	0:3	Technical Only
	9	Onsite Treatment Evaluation: Elm Hall (Colorado School of Mines)	1:0	0:1	Technical Only
	10	Summit View Village: Onsite Wastewater Treatment Considerations	0:4	1:3	Technical Only
	11	Treatment Solutions for Red Rocks Amphitheater	3:0	1:2	Technical Only
	12	Design of a Constructed Wetland in Leadville, CO	1:2	2:1	Non-technical Discussed
Term 16-2 Post-intervention	1	Biochar BEST System for Onsite Remediation of Abandoned Mines Sites	0:5	1:4	Non-technical Discussed
	2	Design, Improvement, and Implementation of Compositing Sanitation Systems in Rocky Mountain National Park	1:4	2:3	Non-technical Discussed
	3	An Analysis of a Proposed Onsite Wastewater Treatment System for the Madha Community	2:3	0:5	Non-technical Focus
	4	Microbial Fuel Cells in Onsite Applications	0:5	2:3	Technical Only
	5	Technical Analysis of Sequencing Batch Membrane Bioreactor Report	0:5	2:3	Technical Only
Term 17-2 Post-intervention	6	Preliminary Design Proposal to Address 2035 NPDES Permit Compliance for Page Wastewater Treatment Plant	5:0	3:2	Non-technical Considered
	7	Onsite Graywater Reuse at Colorado State University	1:2	1:2	Non-technical Considered
	8	Constructed Wetlands for Wastewater Treatment	0:5	2:3	Non-technical Discussed
	9	Sewage-Fed Aquaculture in Vietnam: A Comparative Study of Rural and Urban Systems	0:4	0:4	Non-technical Focus
	10	Onsite Wastewater Treatment Feasibility Study for The Fort	0:4	2:2	Non-technical Discussed

related). Terms that did not fall into one of the 6 categories were not considered. Similarly, terms that have several meanings and were used in different contexts within course projects were not considered. For example, the term “diversity” was not considered because it can be applied in several contexts, such as microbial diversity or social diversity. The paired t-test was used to determine statistically significant differences in the use of key terms within each of the aforementioned six categories. First, the mean use of each key term was determined for pre-intervention course projects (n = 12) and post-intervention course projects (n = 10) for comparison. Second, all key terms within the category (e.g., for the “social” category, all key terms including “stakeholder”, “education”, etc.) were subjected to the paired t-test. If the resulting p value was < 0.05, then the use of key terms within a category was determined to be statistically different post-intervention.

5. Results & Discussion

5.1. Survey of Course Projects for Non-Technical Content

As shown in Table 2, the aggregate level of non-technical (i.e., socio-economic, political, and cultural) content integration into course projects increased from 2014 to 2016 / 2017. Specifically, in 2014, 5 of 12 course projects showed some level of non-technical content integration, while 8 of 10 showed some level of non-technical integration in 2016 / 2017. Post-intervention, two projects examined non-technical concepts centered on WaSH interventions in Vietnam and in Maharashtra, India, while only two project groups failed to integrate any socio-economic, political, and cultural considerations into their analysis. Comparatively, pre-intervention, the majority of course projects (7 of 12) considered technical aspects only and did not consider the aforementioned non-technical themes.

Assigning a weighting factor, i.e. technical only = 1, non-technical discussed = 2, non-technical considered = 3, non-technical focus = 4, to each project and averaging the values for projects pre- and post-intervention, an increase from 1.7 to 2.4 is observed. The increase in integration of socio-economic, political, and cultural considerations into course projects post-intervention suggests that the increased emphasis on these non-technical themes, in the context of WaSH in developing countries, within the course helped enhance student understanding and aided in the synthesis of these concepts as part of a holistic solution for decentralized users.

5.2. Impact of Student Demographics on Non-Technical Integration

Most project groups were gender integrated (17 of 22, 77%). The five groups that were comprised of one gender were each female only; however, two of the five single-gender groups had only one student. Of the three female-only groups with more than one student, two produced the only “non-technical only” course projects (Table 2). The third female-only group with more than one student selected a “technical only” project in 2014 (pre-intervention). An analysis of gender-integrated groups showed no statistically significant trends in the integration of non-technical content. A similar analysis of non-technical content integration and graduate student composition within each group also showed no statistically significant trend in the degree of non-technical integration.

5.3. Term Frequency in Course Reports Pre- and Post-Intervention

Mining of terms yielded 2866 unique terms. The terms with the greatest frequency in the 22 course projects were: “treatment” (1042 occurrences), “water” (893 occurrences), “system” (833 occurrences), “wastewater” (668 occurrences), and “design” (435 occurrences). After terms with less than 5 occurrences were removed, 2011 unique terms remained and were binned into the aforementioned 6 categories (see Section 4). Tables 3 to 8 show results for each category. A statistically significant increase in the use of key terms in the “social” category and the “energy” category was observed (Tables 3 and 7). A statistically significant increase in several key terms within the social category, to include the term “stakeholder”, which increased from 0 occurrences pre-intervention to 14 occurrences post-intervention, was also observed. Further, while the differences were not statistically significant, the mean use of key terms per project increased in the sustainability (81% increase; Table 4) and economic (33% increase; Table 6) categories. A decrease in the mean use of terms per project in the technical category (6%; Table 8) from pre- to post-intervention was also observed. Interestingly, terms in the “WaSH” category showed very little aggregate difference pre- to post-intervention (Table 5). These results suggest that the integration of non-technical concepts increased student awareness and integration of such concepts in the course project.

Table 3. Change in frequency for “social” terms pre-intervention to post-intervention. Paired t-test of the two populations indicates a statistically significant change in the mean use of key terms per project pre- and post-intervention (two-tailed p = 0.045). Percent change is defined as: 100*(post-intervention – pre-intervention)/pre-intervention.

Term	Pre-Intervention Frequency (n=12 projects)	Post-Intervention Frequency (n=10 projects)	Change (%)
Stakeholders	0	14	N/A
Education	1	17	1600% increase
Social	2	24	700% increase
Community	24	125	420% increase
Society	1	5	400% increase
Human / Humans	10	38	280% increase
Culture / Cultural	5	10	100% increase

Population	13	26	100% increase
Acceptance	8	12	50% increase
People	22	26	9% increase
Category Total	**87**	**295**	**239% increase**
Mean per Project	**7.25**	**29.5**	**307% increase**

Table 4. Change in frequency for "sustainability" terms pre-intervention to post-intervention. Paired t-test of the two populations indicates no statistically significant change in the mean use of key terms per project pre- and post-intervention (two-tailed p = 0.093). Percent change is defined as: 100*(post-intervention – pre-intervention)/pre-intervention.

Term	Pre-Intervention Frequency (n=12 projects)	Post-Intervention Frequency (n=10 projects)	Change (%)
Scarcity	1	10	900% increase
Sustainability	2	14	600% increase
Development	10	25	150% increase
Sustainable	19	19	No change
Goals	20	17	15% decrease
Policy	5	1	80% decrease
Category Total	**57**	**86**	**51% increase**
Mean per Project	**4.75**	**8.6**	**81% increase**

Table 5. Change in frequency for "WaSH" terms pre-intervention to post-intervention. Paired t-test of the two populations indicates no statistically significant change in the mean use of key terms per project pre- and post-intervention (two-tailed p = 0.921). Percent change is defined as: 100*(post-intervention – pre-intervention)/pre-intervention.

Term	Pre-Intervention Frequency (n=12 projects)	Post-Intervention Frequency (n=10 projects)	Change (%)
Health	20	37	85% increase
Toilet(s)	61	83	36% increase
Sewage	21	27	29% increase
Sanitary	4	5	25% increase
Sanitation	35	40	14% increase
Urine	29	23	21% decrease
Latrines	7	4	43% decrease
Fecal	24	5	79% decrease
Washing	6	1	83% decrease
Disease	8	0	N/A
Category Total	**286**	**244**	**15% decrease**
Mean per Project	**23.8**	**24.4**	**2% increase**

Table 6. Change in frequency for "economic" terms pre-intervention to post-intervention. Paired t-test of the two populations indicates no statistically significant change in the mean use of key terms per project pre- and post-intervention (two-tailed p = 0.071).

Term	Pre-Intervention Frequency (n=12 projects)	Post-Intervention Frequency (n=10 projects)	Change (%)
Examined Terms (in decreasing order of term frequency): cost(s), benefit(s), economic, capital, million, expense, benefit, savings, inexpensive, owner, labor, money, funding, commission, investment, revenue, economically, valuable, fund, income, costly, purchase, financial			
Category Total	**305**	**338**	**11% increase**
Mean per Project	**25.4**	**33.8**	**33% increase**

Table 7. Change in frequency for "energy" terms pre-intervention to post-intervention. Paired t-test of the two populations indicates a statistically significant change in the mean use of key terms per project pre- and post-intervention (two-tailed p =0.022).

Term	Pre-Intervention Frequency (n=12 projects)	Post-Intervention Frequency (n=10 projects)	Change (%)
Examined Terms (in decreasing order of term frequency): energy, power, electricity, fuel, electrical, gas, heat, kWh, oil			
Category Total	**100**	**229**	**129% increase**
Mean per Project	**8.3**	**22.9**	**175% increase**

Table 8. Change in frequency for "technical" terms pre-intervention to post-intervention. Paired t-test of the two populations indicates no statistically significant change in the mean use of key terms per project pre- and post-intervention (two-tailed p = 0.279).

Term	Pre-Intervention Frequency (n=12 projects)	Post-Intervention Frequency (n=10 projects)	Change (%)
Examined Terms: n = 348			
Category Total	**10,387**	**8,074**	**22% decrease**
Mean per Project	**856.6**	**807.4**	**6% decrease**

Results from the tf-idf analysis were binned into 6 categories in the same manner. Heat maps were constructed for the two term categories showing a statistically significant increase in terms post-intervention: social (Table 9) and energy (Table 10). Heat maps provide a clear visual depiction of term prevalence within each report, as well as which terms had the greatest impact. As shown in Table 9, several social terms had increased values relative to others, principally in course projects post-intervention. More specifically, post-intervention course project report 3 had an increased prevalence of the social terms stakeholders, community, social, and education. Similarly, post-intervention course

Table 9. Heat maps of tf-idf weighted values for terms in the "social" category. Cells colored red indicate that the term did not appear in that particular course project. As the tf-idf value increases, indicating increasing importance of the term within the document and corpus of course projects, the cell color shifts to yellow and then to green. As depicted, most social terms had increased tf-idf values in post-intervention course projects.

	Pre-Intervention (2014)											
Term	Report 1	Report 2	Report 3	Report 4	Report 5	Report 6	Report 7	Report 8	Report 9	Report 10	Report 11	Report 12
stakeholders	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
education	0.0E+00	0.0E+00	0.0E+00	5.0E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
social	0.0E+00	4.8E-04	0.0E+00	0.0E+00	6.3E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
community	0.0E+00	1.5E-04	2.0E-04	1.2E-03	5.9E-04	7.5E-04	2.1E-04	8.1E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00
society	0.0E+00	3.5E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
human	0.0E+00	2.4E-04	0.0E+00	0.0E+00	1.3E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	5.3E-04	1.1E-03
humans	0.0E+00	3.1E-04	0.0E+00	0.0E+00	8.1E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	4.6E-04
culture	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.1E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
cultural	0.0E+00	8.2E-04	0.0E+00	0.0E+00	1.1E-03	0.0E+00	1.1E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
population	0.0E+00	4.3E-04	0.0E+00	2.0E-04	1.1E-03	0.0E+00	3.0E-04	2.9E-04	0.0E+00	0.0E+00	0.0E+00	3.2E-04
acceptance	0.0E+00	0.0E+00	0.0E+00	0.0E+00	9.8E-04	4.6E-04	5.2E-04	5.0E-04	0.0E+00	6.0E-04	0.0E+00	5.5E-04
people	1.8E-04	3.7E-04	0.0E+00	4.5E-04	6.4E-04	0.0E+00	0.0E+00	1.6E-04	1.7E-04	3.9E-04	1.3E-04	1.8E-04

	Post-Intervention (2016, 2017)									
Term	Report 1	Report 2	Report 3	Report 4	Report 5	Report 6	Report 7	Report 8	Report 9	Report 10
stakeholders	0.0E+00	0.0E+00	6.0E-03	0.0E+00	0.0E+00	2.7E-03	2.0E-03	0.0E+00	0.0E+00	0.0E+00
education	0.0E+00	8.1E-04	3.5E-03	0.0E+00	0.0E+00	3.1E-04	5.8E-04	2.2E-03	2.3E-03	0.0E+00
social	8.6E-04	0.0E+00	4.4E-03	0.0E+00	0.0E+00	5.4E-04	2.0E-03	1.4E-03	2.8E-03	0.0E+00
community	5.4E-04	2.2E-04	9.9E-03	2.0E-04	1.7E-04	1.4E-03	0.0E+00	2.1E-03	4.3E-04	2.4E-04
society	0.0E+00	0.0E+00	0.0E+00	0.0E+00	7.7E-04	7.9E-04	0.0E+00	0.0E+00	1.0E-03	1.1E-03
human	0.0E+00	7.1E-03	4.4E-04	0.0E+00	0.0E+00	0.0E+00	1.0E-03	0.0E+00	1.2E-02	0.0E+00
humans	0.0E+00	9.2E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00	6.6E-04	0.0E+00	4.4E-03	0.0E+00
culture	1.4E-03	0.0E+00	0.0E+00	1.1E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00	7.0E-03	0.0E+00
cultural	0.0E+00	0.0E+00	1.5E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
population	0.0E+00	0.0E+00	7.8E-04	5.7E-04	2.4E-04	7.2E-04	1.1E-03	6.5E-04	1.2E-03	3.3E-04
acceptance	0.0E+00	5.6E-04	3.4E-04	0.0E+00	0.0E+00	0.0E+00	8.0E-04	3.0E-03	0.0E+00	0.0E+00
people	0.0E+00	1.1E-03	4.5E-04	0.0E+00	1.3E-04	2.1E-04	2.6E-04	1.2E-04	1.1E-03	1.9E-04

Table 10. Heat maps of tf-idf weighted values for terms in the "energy" category. Cells colored red indicate that the term did not appear in that particular course project. As the tf-idf value increases, indicating increasing importance of the term within the document and corpus of course projects, the cell color shifts to yellow and then to green. As depicted, tf-idf values changed little pre- to post-intervention for energy terms with the exception of post-intervention report 4, which had an increased prevalence of several energy-related terms.

	Pre-Intervention (2014)											
Term	Report 1	Report 2	Report 3	Report 4	Report 5	Report 6	Report 7	Report 8	Report 9	Report 10	Report 11	Report 12
energy	6.5E-05	2.0E-04	5.7E-05	8.1E-05	8.7E-05	1.4E-04	9.2E-05	1.2E-04	1.6E-04	3.6E-05	4.8E-05	3.3E-05
power	0.0E+00	1.3E-03	9.7E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	1.2E-03	0.0E+00
electricity	0.0E+00	2.7E-04	3.5E-03	0.0E+00	7.1E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	5.9E-04	0.0E+00
fuel	0.0E+00	1.9E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
electrical	0.0E+00	7.1E-04	2.7E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
gas	0.0E+00	0.0E+00	9.4E-03	1.5E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
heat	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	8.3E-04	0.0E+00	0.0E+00	2.0E-03	0.0E+00
kwh	0.0E+00	4.1E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
oil	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	7.0E-04

	Post-Intervention (2016, 2017)									
Term	Report 1	Report 2	Report 3	Report 4	Report 5	Report 6	Report 7	Report 8	Report 9	Report 10
energy	0.0E+00	6.6E-05	2.0E-05	1.0E-03	9.7E-05	8.7E-05	4.7E-05	2.2E-05	3.2E-05	3.4E-05
power	0.0E+00	3.3E-03	6.8E-04	1.4E-02	4.1E-03	2.1E-04	3.6E-03	0.0E+00	0.0E+00	5.8E-04
electricity	0.0E+00	0.0E+00	4.0E-03	1.8E-02	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	8.5E-04
fuel	0.0E+00	0.0E+00	2.6E-03	3.9E-02	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
electrical	0.0E+00	1.0E-03	0.0E+00	2.3E-02	0.0E+00	1.2E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00
gas	0.0E+00	0.0E+00	7.4E-04	1.1E-03	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00	0.0E+00
heat	0.0E+00	1.8E-03	0.0E+00	8.2E-04	6.8E-04	6.9E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00
kwh	0.0E+00	1.2E-03	0.0E+00	4.3E-03	0.0E+00	9.1E-04	0.0E+00	0.0E+00	0.0E+00	0.0E+00
oil	0.0E+00	0.0E+00	0.0E+00	3.8E-03	0.0E+00	0.0E+00	2.0E-03	0.0E+00	0.0E+00	0.0E+00

project report 9 had increased prevalence of the social terms human, humans, and culture. The heat map depicting tf-idf values for energy terms (Table 10) shows less distinction between course projects pre- and post-intervention with the exception of post-intervention course report 4, which had increased values for many energy-related terms, including “fuel”, “electricity”, “electrical”, and “power”. The tf-idf analysis shows in granular detail how key terms are distributed amongst course projects, and how specific project reports can influence the overall term frequency analysis. In this analysis, the increased prevalence of social terms was most distributed amongst post-intervention projects; however, the increased prevalence of energy terms was concentrated in one post-intervention project, which heavily influenced the overall term frequency analysis.

6. Conclusions & Future Work

In an onsite water reclamation course, the intentional addition in focused class sessions on socio-economic, cultural, and political considerations appeared to increase student integration of such concepts, evidenced by the increased proportion of student project reports that included such consideration, as well as the quadrupled and tripled per-project relative frequency of terms specific to social and energy considerations, respectively. Such added consideration did not detract from the technical content of the reports. Team demographics based on student standing (senior vs. graduate student) and gender did not significantly impact consideration of non-technical factors. These results suggest that inclusion of WaSH concepts in an onsite water reclamation course may be an effective means to providing better context and thus designing for more sustainable onsite sanitation solutions.

Future efforts may compare corresponding student presentations for these projects to look for similar trends. We also may control for instructor by evaluating projects under pre-intervention conditions with the second instructor. We also recommend testing the effect of including vs. excluding the explicit grading criterion for consideration of non-technical factors.

References

Aizawa, A. (2003). An information-theoretic perspective of tf-idf measures. *Information Processing and Management*, *39*(1), 45–65. https://doi.org/10.1016/S0306-4573(02)00021-3

Blei, D. M., Edu, B. B., Ng, A. Y., Edu, A. S., Jordan, M. I., & Edu, J. B. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, *3*, 993–1022. https://doi.org/10.1162/jmlr.2003.3.4-5.993

Chen, X, Sambamurthy, N., Schimpf, C., Xian, H., Madhavan, K. (2011). Weighted

Social Tagging as a Research Methodology for Determining Systemic Trends in Engineering Education Research. In *ASEE Annual Conference and Exposition, Conference Proceedings*.
Eawag/Sandec (2008). Sandec Training Tool 1.0 - Module 4: Sanitation Systems & Technologies. Dubendorf, Switzerland: Eawag/Sandec.
Feinerer, I., Hornik, K. (n.d.). tm: Text Mining Package. R package version 0.7-3.
Jacquemin, C., Bourigault, D. (2005). Term extraction and automatic indexing. In R. Mitkov (Ed.), *The Oxford Handbook of Computational Linguistics* (1st Editio). https://doi.org/10.1093/oxfordhb/9780199276349.013.0033
Mihelcic, J., Fry, Lauren M., Myre, Elizabeth A., Phillips, Linda D., Barkdoll, B. D. (2009). *Field Guide to Environmental Engineering for Development Workers: Water, Sanitation, and Indoor Air*. Resten, VA: ASCE Press. Retrieved from https://doi.org/10.1061/9780784409855
Ooms, J. (2017). Text Extraction, Rendering and Converting of PDF Documents. R package version 1.5. Retrieved from https://cran.r-project.org/package=pdftools
Rivera, L., Larrondo-Petrie, M. (2017). An Initial Study Applying Data Analysis and Machine Learning Techniques to Analyze Dissertations and Theses in the Engineering Education Field. In *ASEE Annual Conference and Exposition, Conference Proceedings*.
Robertson, S. (2004). Understanding inverse document frequency: on theoretical arguments for IDF. *Journal of Documentation*, *60*(5), 503–520. https://doi.org/10.1108/00220410410560582
Siegrist, R. L. (2014). Onsite and decentralized systems water and wastewater engineering: course development and delivery experiences in higher education. In *Innovations in Soil-based Onsite Wastewater Treatment, Proceedings Soil Science Society of America Conference* (p. 13). Albuquerque, NM.
Siegrist, R. L. (2017). *Decentralized Water Reclamation Engineering: A Curriculum Workbook*. Cham, Switzerland: Springer Interlational Publishing AG. Retrieved from www.springer.com/us/book/9783319404714
UNICEF (2015). *Annual Results Report: Water, Sanitation, and Hygiene*. Retrieved from https://www.unicef.org/wash/files/WASH_2015_Report.pdf
Variawa, C., McCahan, S., Chignell, M. (2013). An Automated Approach for Finding Course-specific Vocabulary. In *ASEE Annual Conference and Exposition, Conference Proceedings*.
Variawa, C., McCahan, S. (2012). Computational Method for Identifying Inaccessible Vocabulary in Engineering Educational Materials. In *ASEE Annual Conference and Exposition, Conference Proceedings*.
WHO (2014). *UN-water global analysis and assessment of sanitation and drinking-water (GLASS) 2014 report: investing in water and sanitation: increasing access, reducing inequalities*. Geneva, Switzerland.
WHO and UNICEF (2015). *Progress on Sanitation and Drinking Water - 2015 update and MDG assessment*. Geneva, Switzerland.