Investigating Students' Habits of Mind in a Course on Digital Signal Processing

Tarun Yellamraju, Alejandra J. Magana and Mireille Boutin

***Abstract*—**
***Contribution*– Knowledge of students' Habits of Mind in a signal processing course, and a method for education research. The method identifies factors that may influence students' performance, but are not evident when analyzing agglomerated data; it is an alternative to the traditional case study method as it derives the cases after applying a clustering approach.**

***Background*– Habits of Mind refer to mathematical, logical and attitudinal modes of thinking required for students of science, mathematics, technology and engineering to become effective problem solvers, capable of transferring such modes to new contexts. These are particularly relevant in a signal processing course in which students must learn to address engineering problems using tools and techniques previously acquired in an abstract context (mathematics).**

***Research Questions*– 1) What are the different Habits of Mind patterns exhibited by the students? 2) Are some of these patterns associated with differences in course grades?**

***Methodology*– A qualitative method is combined with random signal modeling and machine learning. Students' work is first annotated manually based on a custom-built rubric of Habits of Mind. Data is modeled and clustered to obtain statistically significant patterns of Habits of Mind corresponding to divisions of the students into groups.**

***Findings*– The data obtained suggests that the student group is inhomogeneous in terms of their Habits of Mind, and this inhomogeneity is associated with grade differences. In particular, the course grade is found to be dependent on inhomogeneity based on at least two Habits of Mind: "Computation and Estimation" and "Values and Attitudes."**

***Index Terms*—Clustering, communication skills, computing skills, critical thinking, peer instruction, peer review, rubric.**

I. INTRODUCTION

Educators, policymakers and engineering education researchers have attempted to produce a clear understanding of the qualities and knowledge engineering graduates should possess [1]. For example, strong foundations in mathematics, engineering, and technology are highly emphasized in engineering programs [2]. Bodies of accreditation have identified not only the required knowledge and skills that engineering graduates should exhibit, but also the attitudes and behaviors needed to confront complex problems. For instance, ABET criteria [3] stipulate student outcomes ranging from abilities to apply knowledge of mathematics, science and engineering, conduct experiments, analyze and interpret data, and design a system, component, or process to meet desired needs, to abilities to function on multidisciplinary teams, demonstrate professional and ethical responsibility, and communicate effectively.

Furthermore, the Engineer of 2020 proposes a set of aspirations for engineering students needed to operate in societal, geopolitical, and professional contexts within which engineering and its technologies will occur [4]. These aspirations include traits such as strong analytical skills, creativity, ingenuity, professionalism, and leadership [4]. Such aspirations and traits relate to students' "Habits of Mind," which are defined in [5] as modes of thinking required for STEM students to become effective problem solvers capable of transferring such skills to new contexts. An example of Habit of Mind is a willingness to make mistakes while trying to solve a problem, an attitude that allows engineers to successfully attack complex problems.

This work focuses on the Habits of Mind of students learning the theory and application of signals and systems. More specifically, the work is focused on foundational concepts of digital signal processing taught to undergraduate engineering students. Two questions are investigated: 1) What are the different Habits of Mind patterns exhibited by the students?, and 2) Are some of these patterns associated with differences in course grades?

The rationale for centering the investigation around signals and systems is that these concepts are fundamental for electrical engineers and require a strong mathematical background [6], [7]. Furthermore, research has shown that the content of such courses is difficult to master [6], [8]. Previous studies of students learning signals and systems concepts have used quantitative approaches such as concept inventories [6], as well as qualitative approaches using textual analysis of students' responses [8]. Both approaches have advantages and limitations. The approach taken herein to characterize students' 'Habits of Mind' combines a qualitative method with random signal modeling and machine learning techniques. This method combines the advantages of qualitative approaches by first uncovering details of student performance from qualitative data and subsequently dividing students into groups (clusters) based on distinguishing characteristics.

There are many ways to analyze qualitative data [9]. For example, word clouds can be used to visualize text data. After the data has been coded, it can also be analyzed as a whole by visualizing histograms of the frequencies of the rubric elements; The entire dataset can also be summarized using global descriptors for the performance levels of the rubric elements, such as averages, percentages and other descriptive statistics. Correlations can be sought between the performance levels of the rubric elements, or between these performance levels and other course outcomes. Such methods can reveal general trends in the data. However, when the population studied is not homogeneous, that is to say when it is composed of distinguished sub-groups, the results can be misleading. For example, different subgroups may follow different trends in ways that cancel each other in the aggregated data. Thus, the trends of the sub-groups may not manifest themselves when analyzing the dataset as a whole. Manually looking for subgroup trends can be impractical, especially when there are several different ones. This is one reason why automatic clustering is used for analyzing non-homogeneous data.

The method proposed in this paper does not assume that the data is homogeneous. Rather, it seeks to find distint subgroups based on students' Habits of Mind, and statistically validates the existence of these subgroups to confirm the non-homogeneity of the data. It does so automatically using a machine learning framework.

The groups are found by first transforming the qualitative data into quantitative data. Specifically, the qualitative data is transformed into real-valued feature vectors by random signal modeling so it can be automatically clustered using machine learning approaches. The machine learning method selected is well-suited to analyze small datasets in high-dimensions. It also easily lends itself to statistical validation.

More specifically, the students' work is first annotated manually based on a custom-built rubric of Habits of Mind and skill levels. The annotations are vectors, which are stored in sequence for each student. The sequence of vectors of each student is modeled as a random process whose parameters are estimated from the observed data. The parameters of the random process associated with all the students are later clustered using a non-deterministic approach that yields several statistically significant patterns of Habits of Mind. These patterns correspond to binary groupings of the students, that is, dividing them into two groups. The corresponding groups of students are then described using their Habits of Mind histogram as well as course grades. A more detailed statistical analysis is then given using the cumulative distribution function of the difference in average course grade of all the binary groupings. A statistical test is used to determine if the grade differences are significant. Repeating the analysis after removing certain individual Habits of Mind provides a visualization of the contribution of each Habit of Mind to the course grade.

The data analysis approach proposed is a new method that can generally be used to characterize and measure different aspects of professional formation processes in engineering education. The study itself provides a baseline for future efforts in engineering education research methods and assessment.

The rest of the paper is organized as follows: The underlying conceptual framework is presented in Section II, followed by a description of the methods used in Section III. The experimental results are presented in Section IV followed by discussions of the results and the conclusion of the work in Sections V and VI respectively.

II. Conceptual Framework

This investigation is guided by the Scientific Habits of Mind conceptual framework. Habits of Mind are individuals' responses to situations and problems where the answers are not immediately known [10]. Specifically, scientific Habits of Mind refer to modes of mathematical, logical and attitudinal thinking needed for students in the fields of science, mathematics, technology and engineering to become effective problem solvers that can use these modes in new contexts. Effective use of Habits of Mind can allow students to search for solutions moving from highly theoretical to the entirely concrete.

The implications of the conceptual framework for the design of this study relate to the operationalization and characterization of different Habits of Mind in an engineering context. The Habits of Mind explored and operationalized are described in Table I.

III. Methods

The methodological framework for this investigation is a comparative case study method [11]. According to [12], a case study is a research strategy focused on understanding the dynamics within single settings. In [13], it is described as an empirical inquiry that investigates a contemporary issue in depth within its real-life context, especially when the boundaries between phenomenon and context are not clearly evident. A comparative case study was chosen because it facilitates in-depth investigations of two or more instances of groups of students exhibiting similar Habits of Mind.

According to [13], a case study should include data collected from multiple data sources so as to allow the identification of individuals' behaviors, perceptions and attitudes. The use of multiple cases is a common strategy for identifying contextual variations [14]. By comparing cases, one can establish the range of generality of a finding or explanation, and at the same time, pin down the conditions under which that finding will occur [15]. The cases for this study were groups of students exhibiting Habits of Mind in similar ways. The groups were found using a clustering method called n-TARP [16], [17], [18] where TARP stands for "Thresholding After Random Projection". This method was used in [19] to identify clusters in MOOC data. Here, it is applied to feature vectors containing the parameters of a random process modeling a student's Habits of Mind expressed in an active learning activity. As described in the next section, the sources of data considered included student-produced material, peer review material and course outcome data.

The clustering method looks for a good separation of the students into groups after a random projection of their representation (i.e., the feature vector containing the parameters of the student's own random process) down to one dimension [17], [18], [20], [21], [22]. Since the structures of concern (i.e., separations) are found in a one-dimensional space, it is possible to find such groupings even if the number of points projected is fairly small. The one-dimensionality of the data also greatly facilitates the statistical validation of these small groups [17]. These two groups were then analyzed as separate cases and their characteristics, including similarities and differences, were further explored.

A. Participants, Procedures and Dataset

The study context is a course on signal processing in which students were asked to produce learning material and share it on a public website [23]. Specifically, the instructor

TABLE I
DEFINITION OF HABITS OF MIND AS PROPOSED BY [5] AND OPERATIONALIZATION HEREIN

Habits of Mind	Definition as per [5]	Operationalization
Computation and Estimation	Ability to judge an appropriate computation method to be used based on specific circumstances.	Ability to choose an appropriate computation method and carry out the mathematical procedure accurately
Mathematical Rigor	Ability to make careful observations and handle information.	Ability to handle mathematical rigor and remember details of a definition.
Communication Skills	Ability to communicate ideas and share information with fidelity and clarity.	Ability to communicate effectively, explain background and present a good and meaningful flow of ideas.
Critical Response Skills	Ability to detect the symptoms of doubtful solutions, assertions and arguments.	Ability to detect the symptoms of doubtful solutions, assertions and arguments in one's own work and in peers' work.
Values and Attitudes	General social values and people's attitudes toward their own or others ability to understand science, engineering and mathematics.	Student's attitudes towards their peers' work and their own ability to make assessments on others' work.

pre-defined nine topics covered in the course, and students prepared a slecture [24] explaining the course material for a topic of their choice in their own words. The term "slecture" is a concatenation of the words "student" and "lecture." Invented by Boutin in 2010, the idea is to have students create online learning material based on the teaching of a professor.

In addition to creating a slecture, the students were also instructed to review and comment on the slectures prepared by their peers (one slecture per topic for each student). Note that online discussion comments have been previously used to uncover students' Habits of Mind [25].

Specifically, the unit of analysis, the major entity that is being analyzed in a study, was each student's individual contribution to a public website. Two additional data sources were the feedback provided to their peers in the form of a review, and the final grade as a measure of performance. The cases for this study were groups of students exhibiting Habits of Mind in similar ways. Such cases were uncovered by the clustering method and were further compared and analyzed regarding their performance and observed Habits of Mind.

A total of 28 students participated: 27 of these presented the slecture in written form, while one presented it as a video. The 27 written slectures were used in this study. There were 3.0 slectures per topic and 6.89 reviews per slecture, on average. This is because some students did not complete the review assignment while others provided more/less than nine reviews. All students who completed the tasks received full credit on the assignment, so the exercise in itself did not produce any difference in grades among the participants.

B. Data Scoring

The data scoring was performed using a rubric, created and validated iteratively, starting with an inductive approach, followed by a deductive approach. For the inductive approach, one of the authors with expertise in education research built a Habits of Mind focused criteria, Table I, derived from the literature [5]. This author is a subject expert who is well acquainted with the student population and the skill level expectations within the field of studies. For the deductive approach, the initial definitions were then further operationalized for the context of the study. Based on the initial operationalization of each construct or criterion, levels of performance were identified by this subject expert, Table II. A second author, also with expertise in signals and systems and well acquainted with the student population and the skill level expectations within the field of studies, then used the rubric to annotate the slectures and the reviews. The first pass of the data scoring was then validated and reviewed by a third author. In the process, the rubric was modified to better capture students' patterns, and when modified, it was tested against the data following a deductive approach. The process and findings were discussed among the three authors. This iterative approach was performed three times resulting in the rubric presented in Table II (reported in [26]).

The element "Values and Attitude" was initially focused on perceived importance or confidence in the subject domain. Traditionally, this habit of mind is assessed via surveys asking students to report their perceived confidence on the subject matter or their self-perceived abilities to understand the concepts. Because an opportunity to survey students was not available, "Values and Attitudes" focusing on students' abilities to evaluate their own and their peers' work was indirectly characterized. Specifically, this study found that their critical views of their own work and that of their peers' was a good indicator of students' confidence and abilities in their own knowledge and skills. So the focus was shifted to analyzing if students were able to provide a meaningful critique of their peers' work and how their attitude appeared in their feedback. Below are two examples of Values and Attitudes ratings.

- "*I think specific outline is very helpful and make easy to follow the formula and graphs. Formulas and graphs are very clear to understand.*" − Basic Level rating
- "*I think an important aspect that you did not include in your final answer is that the DTFT of a DT signal must be periodic. Your answer must be"rep-ed" to denote it's periodicity. Otherwise your answer is only correct for* $0 \leq \omega \leq 2\pi$*. The DTFT of* $x[n]$ *is* $rep_{2\pi}(2\pi\delta(\omega-\omega_0))$ *Overall color coating was very helpful, and the slecture was concise and clear.*" − Advanced Level rating

Another example is the element "Computation and Estimation", which initially focused on the ability to choose an appropriate computation method and recognize when approximations can be made. The topics covered by the students did not involve any approximations, but rather mathematical computations. Hence, the focus for this element was shifted

TABLE II
RUBRIC GENERATED FROM STUDENT EXHIBITED HABITS OF MIND.

	Description		Performance Level			
Tag	Element	Definition	Below Basic 1	Basic 2	Proficient 3	Advanced 4
A	Computation and Estimation	Ability to choose an appropriate computation method and carry out the mathematical procedure accurately.	Student selected an incorrect method and the solution was completely off.	Student selected a correct method but the solution was incorrect.	Student selected an appropriate method and the solution was correct. However, the student did not provide a justification for the method based on the circumstances, or the justification was inadequate.	Student selected an appropriate method, provided correct justification for the method selection based on the circumstances and the solution was correct.
B	Mathematical Rigor	Ability to handle mathematical rigor and remember details of a definition	Student was not at all rigorous in the involved mathematics.	Student displayed some rigor but there were major errors.	Student was very rigorous but made small errors.	Student was very rigorous and made no errors.
C	Communication Skills	Ability to communicate effectively, explain background and present a good and meaningful flow of ideas	Student presented an unclear and unjustified procedure.	Student presented a somewhat clear procedure but it was unjustified.	Student presented a clear procedure with a reasonable justification.	Student presented a clear procedure with a detailed justification based on the theory or principles.
D	Critical Response Skills	Ability to detect the symptoms of doubtful solutions, assertions and arguments in one's own work and in peers' work	Student was unable to identify incorrect procedures and provided no evidence of procedures for validation of their solution.	Student was able to identify incorrect procedures but was unable to correct them. Student provided no evidence of procedures for validation of their solution.	Student was able to identify incorrect procedures and corrected them properly. However, student provided no evidence of procedures for validation of their solution.	Student was able to identify incorrect procedures and correct them properly or did not demonstrate any incorrect procedures. In addition, student demonstrated evidence of applying procedures for validation of their solution.
E	Values and Attitudes	Students attitude towards their peers' work and their own ability to make assessments on others work	Student made negative comments about others' work or was indifferent to it.	Student made generic comments that do not provide any insight or critique.	Student made good comments providing insight and a somewhat reasonable critique.	Student made excellent comments, correcting mistakes and providing insightful critique.

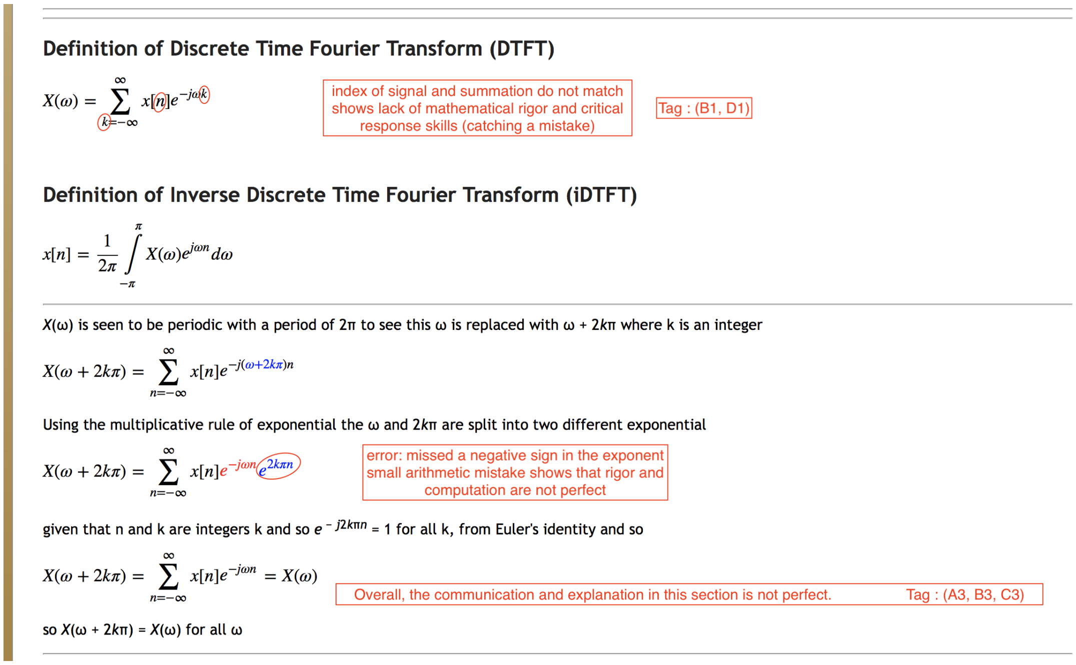
to appropriateness of the computational method used and mathematical accuracy of the computation. A final example is the element "Mathematical Rigor", which in an earlier version of the rubric [26] was called "Manipulation and Observation". Mathematical rigor in this case refers to the correctness of the notation and attention to detail in the writing of mathematical expressions. This element was mostly found in definitions and mathematical statements either within a computation or standing on its own within the text of a slecture. (See the top of Fig. 1, tagged as (B1,D1), for an example of poor rigor). The change in the rubric element "Manipulation and Observation" to "Mathematical Rigor" was motivated by the fact that mathematical rigor within arguments and explanations was found to play a more critical role than in handling basic mathematical manipulation and observation. In fact, manipulation and observation can be bundled in with computation and estimation. The final rubric is presented in Table II. Please note the alphabetical labeling of the items, used in the annotation process described below.

C. Annotating Slectures and Comments with Rubric Tags

The annotated material of each student was recorded as a sequence of vectors representing the sequence of Habits of Mind elements and levels of performance. For example, one part of a slecture might have been tagged with the vector $(A4, B4, C2, D4)$ to denote that the student carried out the computations effectively with the necessary rigor and validation but the explanation was lacking in terms of communication. The length of the text used for each labelled block varied with the context and student communication style, so as to include separate ideas and concepts. The lengths varied from one sentence or equation for concise ideas to several paragraphs for lengthy or redundant explanations and were decided subjectively by the rater on a case to case basis.

Two examples are provided to illustrate how the slecture material is tagged/annotated, with the annotations and tags shown in red bounding boxes. The first, Fig. 1, shows a low value tag because of errors made in the slecture. This figure corresponds to the first half of a slecture and serves as an example of lower score annotations. The second, Fig. 2, shows a high value tag since the material was almost flawless.

Table III compares of two examples of identified Habits of Mind with different levels of performance. Example 1 has a computation error under Computation and Estimation, and incorrect mathematical statements for Mathematical Rigor and Critical Response Skills. For Communication Skills, it provides a basic explanation of an idea through a set of mathematical equations without much context or explanation, and for Values and Attitudes, has a very vague comment



Definition of Discrete Time Fourier Transform (DTFT)

$$X(\omega) = \sum_{k=-\infty}^{\infty} x[n]e^{-j\omega k}$$

index of signal and summation do not match shows lack of mathematical rigor and critical response skills (catching a mistake)

Tag : (B1, D1)

Definition of Inverse Discrete Time Fourier Transform (iDTFT)

$$x[n] = \frac{1}{2\pi}\int_{-\pi}^{\pi} X(\omega)e^{j\omega n}d\omega$$

$X(\omega)$ is seen to be periodic with a period of 2π to see this ω is replaced with $\omega + 2k\pi$ where k is an integer

$$X(\omega + 2k\pi) = \sum_{n=-\infty}^{\infty} x[n]e^{-j(\omega+2k\pi)n}$$

Using the multiplicative rule of exponential the ω and $2k\pi$ are split into two different exponential

$$X(\omega + 2k\pi) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n}e^{2k\pi n}$$

error: missed a negative sign in the exponent small arithmetic mistake shows that rigor and computation are not perfect

given that n and k are integers k and so $e^{-j2k\pi n} = 1$ for all k, from Euler's identity and so

$$X(\omega + 2k\pi) = \sum_{n=-\infty}^{\infty} x[n]e^{-j\omega n} = X(\omega)$$

Overall, the communication and explanation in this section is not perfect. Tag : (A3, B3, C3)

so $X(\omega + 2k\pi) = X(\omega)$ for all ω

Fig. 1. Low value tag in a slecture

that does not really provide any insight or critique. These justify a lower score than for Example 2, for which the corresponding items are mostly correct, mathematically error free, and provide greater insight, critique and context.

D. Inter-Rater Reliability

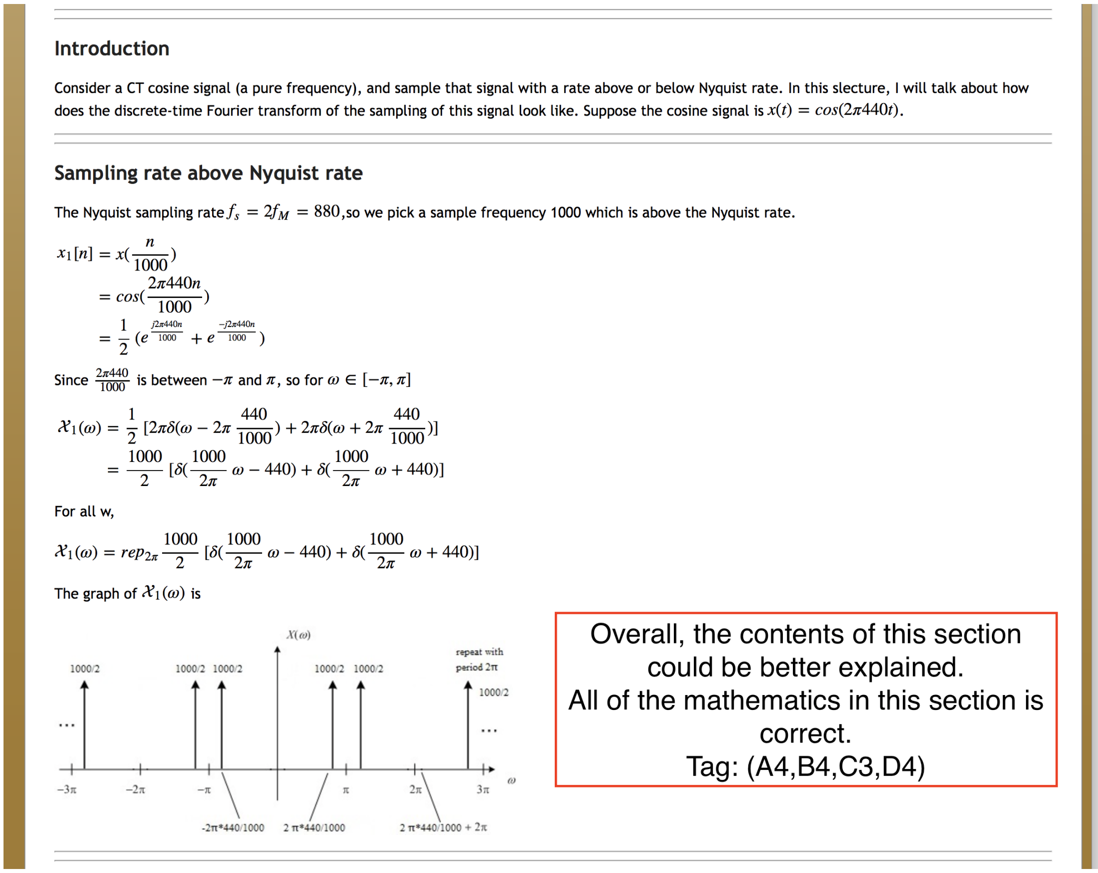
The reliability of the data scoring was estimated by having another author annotate the slectures and reviews of 11 students in two phases. Data from 11 students was chosen at random to avoid bias and to span at least a third of the data. First, the rubric was consulted and discussed with the first rater. Then the slectures and reviews of five students were rated by the second rater. A discussion followed in which the two raters compared their scoring and discussed the reasons behind the differences. The second rater then rated the slectures and reviews of six other students. The reliability of the first (five students) and second phase (six students) were measured using correlation coefficients (Pearson product-moment coefficients [27], [28]) for each phase. The correlation coefficient measures the reliability [29] of ratings by estimating how much of the variance in the ratings is due to the true ratings (numerator of the coefficient) versus how much is due to noise (denominator of the coefficient.) Two correlation coefficients were computed for each phase: one represents the reliability of the detection of the different rubric elements present in the work; the other represents the reliability of the accuracy of the scores for all the rubric elements.

More specifically, the reliability of the detection of the initial rater was estimated using the letter tags (without scale value) for both raters: when a part of text was coded with a given letter by a rater, a "detection", denoted by a "1", was recorded for that rater; if the other rater also coded the same text with that letter, then the other rater was considered to have also detected that event, and a "1" was recorded for that rater as well. Conversely, if the other rater did not code that text with that letter, then this rater was considered to have missed that event and a "0" was recorded. The reliability of this detection process was measured using the correlation coefficient [30] of the sequences of 0's and 1's for the two raters. The reliability of the accuracy of the labeling (letter and score) was only considered for those events (rubric elements) detected by both raters. A sequence of scores for each rater was built by collecting all the numerical scores for all the commonly detected events of a given type (tag) into a vector; the correlation coefficient of the two vectors for that tag was then computed.

The reliability of the detection of the rubric elements in the first phase of the inter-rater reliability testing was found to be 0.62 (correlation coefficient). In the second phase, that number increased to a much higher value of 0.82. The reliability of the accuracy of the scoring was found to be 0.94, already a very high value, which increased modestly to 0.96 in the second phase. Thus both the detection of rubric elements and accuracy of the data scoring were considered to be very reliable. These reliability estimates were computed after the two phases of rating were completed.

E. Data Analysis

1) Class Statistics: The Habits of Mind of the class were first summarized using a 2D histogram of tag values (in a 5×4 grid) to analyze the distribution of the annotation tags for the entire set of students in the study, and to look for some global trends in the class with regard to their Habits of Mind. The final grade distribution for the entire course is also examined to characterize their academic performance in the course as a whole.



Introduction

Consider a CT cosine signal (a pure frequency), and sample that signal with a rate above or below Nyquist rate. In this slecture, I will talk about how does the discrete-time Fourier transform of the sampling of this signal look like. Suppose the cosine signal is $x(t) = cos(2\pi 440t)$.

Sampling rate above Nyquist rate

The Nyquist sampling rate $f_s = 2f_M = 880$,so we pick a sample frequency 1000 which is above the Nyquist rate.

$$\begin{aligned} x_1[n] &= x(\frac{n}{1000}) \\ &= cos(\frac{2\pi 440n}{1000}) \\ &= \frac{1}{2}(e^{\frac{j2\pi 440n}{1000}} + e^{\frac{-j2\pi 440n}{1000}}) \end{aligned}$$

Since $\frac{2\pi 440}{1000}$ is between $-\pi$ and π, so for $\omega \in [-\pi, \pi]$

$$\begin{aligned} \mathcal{X}_1(\omega) &= \frac{1}{2}[2\pi\delta(\omega - 2\pi\frac{440}{1000}) + 2\pi\delta(\omega + 2\pi\frac{440}{1000})] \\ &= \frac{1000}{2}[\delta(\frac{1000}{2\pi}\omega - 440) + \delta(\frac{1000}{2\pi}\omega + 440)] \end{aligned}$$

For all w,

$$\mathcal{X}_1(\omega) = rep_{2\pi}\frac{1000}{2}[\delta(\frac{1000}{2\pi}\omega - 440) + \delta(\frac{1000}{2\pi}\omega + 440)]$$

The graph of $\mathcal{X}_1(\omega)$ is

Overall, the contents of this section could be better explained.
All of the mathematics in this section is correct.
Tag: (A4,B4,C3,D4)

Fig. 2. High value tag in a slecture

2) Statistical Model Building: For a more in-depth analysis, a statistical model that describes each student's individual Habits of Mind was built and later used to cluster the students. The statistical model represents a random process underlying the sequence of annotation tag vectors. The parameters of the statistical model are estimated from the annotated data. For simplicity, consecutive vector tags are assumed to be independent. The different elements (A,B,C,D,E) are also assumed to be independent. However, in other circumstances, another perhaps more complicated model could be more appropriate. For example, if the work was carried out over a long period of time during which an improvement was expected, the consecutive vector tags could be modeled by a time-dependent process.

The statistical model consists of the likelihood of the tag scores for each element in the proposed rubric. In other words, it is represented by the discrete probabilities [30]: $P(k)$ and $P(j|k)$, for $k \in \{A, B, C, D, E\}$ and $j = 1, 2, 3, 4$. These probabilities are estimated by the relative frequencies of each tag in the scored data as follows.

$$\bar{P}(A) = \frac{\sum_{i=1}^{N_s} \mathbb{I}\{student\ s\ gets\ tag\ A\ in\ annotation\ i\}}{N_s},$$

$$\bar{P}(1|A) = \frac{\sum_{i=1}^{N_s} \mathbb{I}\{student\ s\ gets\ tag\ A(1)\ in\ annotation\ i\}}{\sum_{i=1}^{N_s} \mathbb{I}\{student\ s\ gets\ tag\ A\ in\ annotation\ i\}}$$

where N_s is the number of annotations recorded for student s. A similar expression is used for the other score values 2, 3, and 4, for rubric element A. The probabilities for the other rubric elements B, C, D and E are computed in a similar fashion, except that the parameters N_s takes the value 26 for E (since the students could review a maximum of 26 slectures.)

Thus five model parameters for each of the five elements of the rubric were estimated, for a total of 25 parameters for each student. Since the number of tags is different for different students, the parameter N_s (number of annotations received by the student) is added to these parameters to highlight the difference between short and long slectures. Thus, 26 parameters are used to represent each student; these are stacked into a vector of dimension 26.

3) Clustering: Clustering a small number of points (27) in a high-dimensional space (26) is challenging and requires the use of an algorithm specially designed for small data. One such algorithm is "n-TARP" [31], [16], [17], which seeks good separation of the data after a projection onto a random line. The separation is obtained by projecting and thresholding the data n times, and picking the projection with the best separation among those n. This is a modification of the random projection approach developed in [21], [22], [18], empirically shown to work well for "real" high-dimensional data in general. TARP stands for "Thresholding After Random Projection." Instead of hierarchically clustering the data using a tree of thresholds after random projections (n-TARPs) as in [22], the method performs a single n-TARP on a fraction of

TABLE III
EXAMPLES OF HABITS OF MIND EXHIBITED BY STUDENTS

Habits of Mind	Example 1 (Low level)	Example 2 (High level)
Computation and Estimation (from slectures)	"$X(f) * f_s \sum_{k=-\infty}^{\infty} \delta(f - kf_s) = f_s \sum_{k=-\infty}^{\infty} X(f) * \delta(f - \frac{k}{f_s})$" −Basic Level	"$X_s(f) = F(\sum_{n=-\infty}^{\infty} x(nT)\delta(t - nT)) = \sum_{n=-\infty}^{\infty} x(nT)F(\delta(t - nT))) = \sum_{n=-\infty}^{\infty} x(nT)e^{-j2\pi f nT}$" −Advanced Level
Mathematical Rigor (from slectures)	"$X(2\pi f) = X(f)$" −Below Basic Level	"$\cdots X(f) = F(\delta(t - t_0)) = \int_{-\infty}^{\infty} \delta(t - t_0)e^{-i2\pi f t}dt = e^{-i2\pi f t_0} = e^{-i\omega t_0}$" −Advanced Level
Communication Skills (from slectures)	"Comb operator is used in time domain: $comb_T[x(t)] = \cdots = x(t) \cdot P_T(t) \cdots$" −Basic Level	"$\cdots x_s(t)$ is created by multiplying a impulse train $P_T(t)$ with the original signal $x(t)$ and actually $x_s(t)$ is $comb_T(x(t))$ where T is the sampling period $\cdots$" −Advanced Level
Critical Response Skills (from slectures)	"$x(t) = \int_{-\infty}^{\infty} \delta(t - \tau)d\tau$" −Below Basic Level	"$\cdots$ the minimum repeating period T has to be $> a + b$ (a is the left boundary of the curve and b is the right boundary of the curve)." −Proficient Level
Values and Attitudes (from peer reviews)	"I think specific outline is very helpful and make easy to follow the formula and graphs. Formulas and graphs are very clear to understand." −Basic Level	"I think an important aspect that you did not include in your final answer is that the DTFT of a DT signal must be periodic. Your answer must be"rep-ed" to denote it's periodicity. Otherwise your answer is only correct for $0 \leq \omega \leq 2\pi$. The DTFT of $x[n]$ is $rep_{2\pi}(2\pi\delta(\omega - \omega_0))$ Overall color coating was very helpful, and the slecture was concise and clear" −Advanced Level

the data given, and tests the statistical validity of any clustering identified using the remaining fraction of the data [17].

In general, clustering methods can be viewed as maps from the feature space (in high-dimensions for the data at hand) to one-dimensional space $\mathbb{R}$, followed by some thresholdings. Different methods have different ways of defining the "best" projection and thresholds. Projecting the data onto a line and thresholding corresponds to finding a linear separation between the clusters, which is the simplest form of clustering. Linear separations are well-suited for small data in high-dimensions because they can be found when only a small number of points are given. Previous work in [17], [18], [21], [22] has shown that many good linear separations can be found in real data by picking the line of projection at random. This is because real data often has a lot of hidden structure in high-dimensions that can be extracted through random projections [21], [22]. This observation, and there only being a small number of students considered in this study, are the reasons for employing n-TARP to cluster the data.

In these experiments, a binary clustering was performed using n-TARP with the parameter n set to 500 to divide the set of students into two groups, picking the best separation among the 500 projections performed. This was done in two phases: a training and a validity testing phase. Half of the data (randomly chosen every time) was used for each phase. Because the data size is so small (27 points), one would hardly expect to find any meaningful cluster in the original space. Looking for clusterings in a one-dimensional space addresses this issue because the projected points are closer together than in the original space. The extent to which the (training) projected points are clustered is measured using "normalized withinss (W)", a renormalized version of the within-class scatter of the data [32], [21]. More specifically the within-class scatter of [32] is divided by the number of points and the empirical variance of the projected data. This insures that the measure is independent of the number of points considered and invariant under a rescaling of the dataset [21]. The definition of "normalized withinss (W)" for a set of projected points $x_1, x_2, \ldots, x_m \in \mathbb{R}$ is given below [22]

$$W = W(x_1, \cdots, x_m) = \min_{C_1, C_2} \frac{\sum_{i \in C_1} (x_i - \mu_1)^2 + \sum_{i \in C_2} (x_i - \mu_2)^2}{\tilde{\sigma}^2 \cdot m},$$

where C_1 and C_2 are a disjoint partition of the set of indices $\{1, \ldots, m\}$, μ_1 and μ_2 are the (empirical) mean of the points x_i whose indices are in C_1 and C_2, respectively, and $\tilde{\sigma}$ is the (empirical) standard deviation of the set of points $x_1, \cdots, x_m \in \mathbb{R}$.

Training Phase:

1) For $i = 1$ to n:
2) Generate a random vector r_i in 26 dimensional space;
3) Project the training data onto this vector r_i to form 1D projection values;
4) Use k-means (set $k = 2$) to find 2 clusters in the 1D projection values;
5) Find the normalized withinss w_i for this cluster assignment;
6) End loop.
7) Pick lowest w_i among the n measurements and store the random vector r^* associated with it and determine a threshold t^* that separates the classes formed in the 1D projected space.

Validity Testing Phase

1) Import r^* and t^* from the training phase
2) Project the testing data onto the vector r^*
3) Use the threshold t^* to assign clusters to each of the testing samples
4) Perform permutation test with Monte-Carlo simulations [33] on the projected test data at statistical significance level of 99%.

4) Pattern (Binary Clustering) Analysis: Since the clustering is random, it can yield several different (and valid) binary clusterings. Each of these clusterings splits the students into two groups based on some distinctive Habits of Mind patterns. Although the pattern is described by the coefficients of the random projection vector r^* used for the projection, it is typically hard to make sense of the pattern directly from these coefficients. As an alternative, the histogram of Habits of Mind annotations for the two groups are considered and compared (i.e., the frequency of occurrence of each rubric annotation for both clusters). The distribution of the course grades for the two groups are also compared.

The number of different Habits of Mind patterns exhibited by students is quantified following the approach of [17], [21], [22]. Specifically, the distribution of the normalized withinss of the (random) projected data is plotted, and the area of the distribution below the value ≤ 0.36 (threshold value after which no clusters exist) is computed.

To quantify the relationship between Habits of Mind patterns and course grade, the empirical Cumulative Distribution Functions (CDF) [30] of the absolute difference between the average grades of both groups is constructed. To check the dependence of the different elements of the rubric and course grades, each element is removed one by one and a new CDF of absolute difference between average grades between groups is obtained: the resulting CDF curves are then compared.

5) Hypothesis Testing: Conceivably, randomly grouping the students into two clusters could result in different grade distributions for the two clusters due to chance rather than due to the Habits of Mind of the students. In order to test the statistical significance of the observations, independence of the grades on the patterns (groupings) of Habits of Mind is set as the null hypothesis, and statistical significance of the observations is tested by comparing the CDF curves of the previously obtained clusterings with the CDF curves for random clusterings. In other words, the CDF of grade differences for the binary clusterings previously obtained is compared with the CDF of grade differences that would be obtained with random division of the students into two groups.

IV. Results

In this section, the results of the data analysis are presented. After a brief comment on slectures, summary statistics like frequency of occurrence of the different levels of the elements of the rubric are presented. Following that, the results of the n-TARP clustering algorithm, which uses the feature vector formed through the model fitting described in the previous section, is presented. The frequency of occurrence of the rubric tags for the resulting groups are compared to identify the differences that led to the formation of the groups. Next, the results on the extent of clusterability of the data are presented. The various different clusters formed as a result of the random projection model underlying the n-TARP clustering algorithm are presented. Finally, the connections between 'Habits of Mind' patterns and the grades of students are examined.

A. Overall Patterns of Students' Habits of Mind

Table IV shows the relative number of times each element/level of the rubric was tagged in the study. The most frequent tag is Values at a Basic level (32.2%), followed by Values at a Proficient level (11.2%) and Computation at an Advanced level (7.7%). No Below Basic level was found with a frequency above 2%, and the only Habits of Mind element noted more than 3% of the time is Values. Overall, elements other than Values tend to be tagged more frequently at the Proficient or Advanced level. Overall, a majority (64.8%) of the tags were at the Proficient or Advanced level.

TABLE IV
PERCENTAGES OF EXHIBITED HABITS OF MIND AMONG ALL 27 STUDENTS

Element/Level	Below Basic	Basic	Proficient	Advanced
Computation	0	1.06	1.32	7.71
Rigor	1.59	2.65	5.85	3.98
Communication	0.26	2.92	6.11	3.98
Critical Response	1.32	1.06	3.98	5.58
Values	1.06	32.18	11.17	6.11

The clustering method was repeated more than 1000 times to form groupings; one such (statistically significant) grouping, found to have a significant effect on the grades, is analyzed in Tables V and VI. Observe that students in Cluster 2 have much larger numbers of high level tags for all the elements of the rubric than Cluster 1, indicative of a higher level of Habits of Mind performance. Indeed, a majority (64.8%) of annotation tags for Cluster 2 are at the "Proficient" or "Advanced" level. In contrast, a majority (63.3%) of annotation tags for Cluster 1 are at the "Below Basic" and "Basic" level. Thus, the members of Cluster 2 are identified as the "Habits Developed" students (Case 2), and the members of Cluster 1 as "Habits Developing" students (Case 1).

B. Case Comparison

As stated earlier, two cases were identified. Case 1 is called the "Habits Developing" group, and Case 2 is called the "Habits Developed" group. The groups were characterized on the basis of the overall distribution of levels (more Advanced level tags for Case 2 than for Case 1). As observed from the sums of the columns of Tables V and VI for each row element (Habit of Mind), the "Habits Developing" group (Cluster 1, Table V) is also distinguished by a higher probability of expressing the "Values" element, 55% vs 48% for the "Habits Developed" group (Cluster 2, Table VI). Further, the "Habits Developing" group also shows a slightly lower probability of expressing the "Communication" element, 11% vs 14% for the "Habits Developed" group. On the other hand, the likelihood of exhibiting the "Computation" (9% versus 11%), "Rigor" (14% versus 14%) and "Critical Response" (11% versus 12%) elements are somewhat similar for both groups (Cluster 1, Table V vs Cluster 2, Table VI).

C. Overall Course Performance and Performance by Case

The grade distributions for the clusters and the entire class are shown in Table VII. The grade differences between the

TABLE V
PERCENTAGES OF EXHIBITED HABITS OF MIND FOR CASE 1: HABITS DEVELOPING (10 STUDENTS)

Element/Level	Below Basic	Basic	Proficient	Advanced
Computation	0	1.66	3.33	4.16
Rigor	3.33	4.16	5.00	1.66
Communication	0	5.00	4.16	1.66
Critical Response	2.50	1.66	3.33	3.33
Values	3.33	41.66	9.16	0.83

TABLE VI
PERCENTAGES OF EXHIBITED HABITS OF MIND FOR CASE 2: HABITS DEVELOPED (17 STUDENTS)

Element/Level	Below Basic	Basic	Proficient	Advanced
Computation	0	0.78	0.39	9.37
Rigor	0.78	1.95	6.25	5.07
Communication	0.39	1.95	7.03	5.07
Critical Response	0.78	0.78	4.29	6.64
Values	0	27.73	12.10	8.59

clusters (e.g., difference of 1.45 between mean grades) indicates that having well developed Habits of Mind is associated with good course performance. Indeed, none of the Habits Developing students received an A in the course, whereas none of the Habits Developed students received an F or a D in the course, with a majority receiving As or Bs.

TABLE VII
GRADE DISTRIBUTIONS

Grade	All Students	Case 1: "Habits Developing"	Case 2: "Habits Developed"
A (4.0)	5	0	5
B (3.0)	10	2	8
C (2.0)	8	4	4
D (1.0)	2	2	0
F (0.0)	2	2	0
Mean Grade	2.51	1.60	3.05
Standard Deviation	1.12	1.07	0.74

However, there may be other patterns of Habits of Mind whose association to the course grade could be different. Fig. 3 shows the distribution of normalized withinss W for the dataset, which shows the very high clusterability of the dataset [21], [22], as approximately 80% of the clusters found have a value of $W \leq 0.36$ (cluster present). In other words, the data at hand is not homogeneous.

The connection between these patterns and the grade is shown to be very strong in Fig. 4. Specifically, the graph shows the CDF of the (absolute) difference in average grade between the two groups for a total of 1000 attempted binary groupings of which only valid statistically significant groupings are retained (about 90%). The lower the curve at a given point (grade value), the higher the proportion of patterns with an average grade difference at least as large as that grade value. For example, about 30% of the Habits of Mind patterns found were associated with an average grade difference of at least 0.5 (since the y-axis value for a difference in grades of 0.5 is about 0.7). The x-axis intercept is about 0.02 and thus no groups yield an average grade difference less than 0.02 (0.5%).

The elements of the rubric were removed one at a time: each time, similar to above, a new set of 1000 clusterings was obtained of which only valid statistically significant groups are retained, and the CDF of the absolute value of the average grade difference between the groups was computed. The resulting curves are also shown in Fig. 4. Observe that removing Element A shifts the CDF curve up (i.e. the new CDF curve is above the original CDF curve), and thus the relationship between Habits of Mind patterns not involving Computation are less strongly associated with different grade outcomes than Habits of Mind patterns involving Computation. This implies that Computation (element A) is related (dependent) to the final grade. The same is true, though to a lesser extent (less grade difference), with Values and Attitudes (element E), but not for Communication Skills (element C).

Note that this does not mean that there is an overall association between the level of performance with respect to these elements and the grade. Indeed, as shown in Figs. 5, 6 and 7, when collecting the data for all the students, no correlation/association is observed between the grade and the average level of performance in element E (Fig. 5) or with element A (Fig. 6). For comparison, the association graph for grade and level of performance for element C, Fig. 7, is also shown.

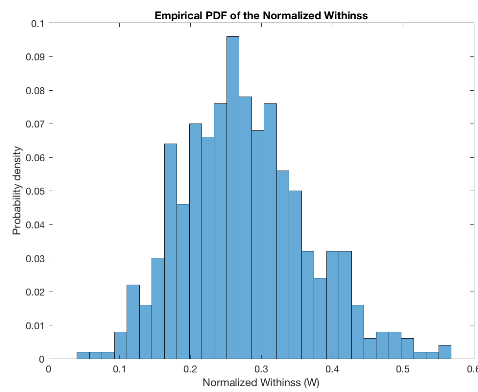


Fig. 3. Empirical probability distribution function of the normalized withinss W. The Clusterability of the data is measured by the pdf of Withinss.

Hypothesis Testing

The CDF of grade differences for the binary clusterings was compared with the CDF of grade differences that would be obtained with random division of the students into two groups, Fig. 4; three CDF curves were added to the plot, identified in the legend as mean and $\pm$ 5sigma.

To obtain these five curves, the students were randomly grouped into two clusters 10,000 times to get 10,000 differences in average grades of the resulting random clusters. Note that the Habits of Mind features were not utilized at any point in this process, just random divisions of the students into two groups. These 10,000 differences were used to generate a CDF curve based on the null hypothesis. This process was repeated 100 times in order to get 100 CDF curves, which were used to form the mean null hypothesis curve (in solid black in Fig. 4) along with the null hypothesis curves shifted five standard deviations away (in dashed black in Fig. 4).

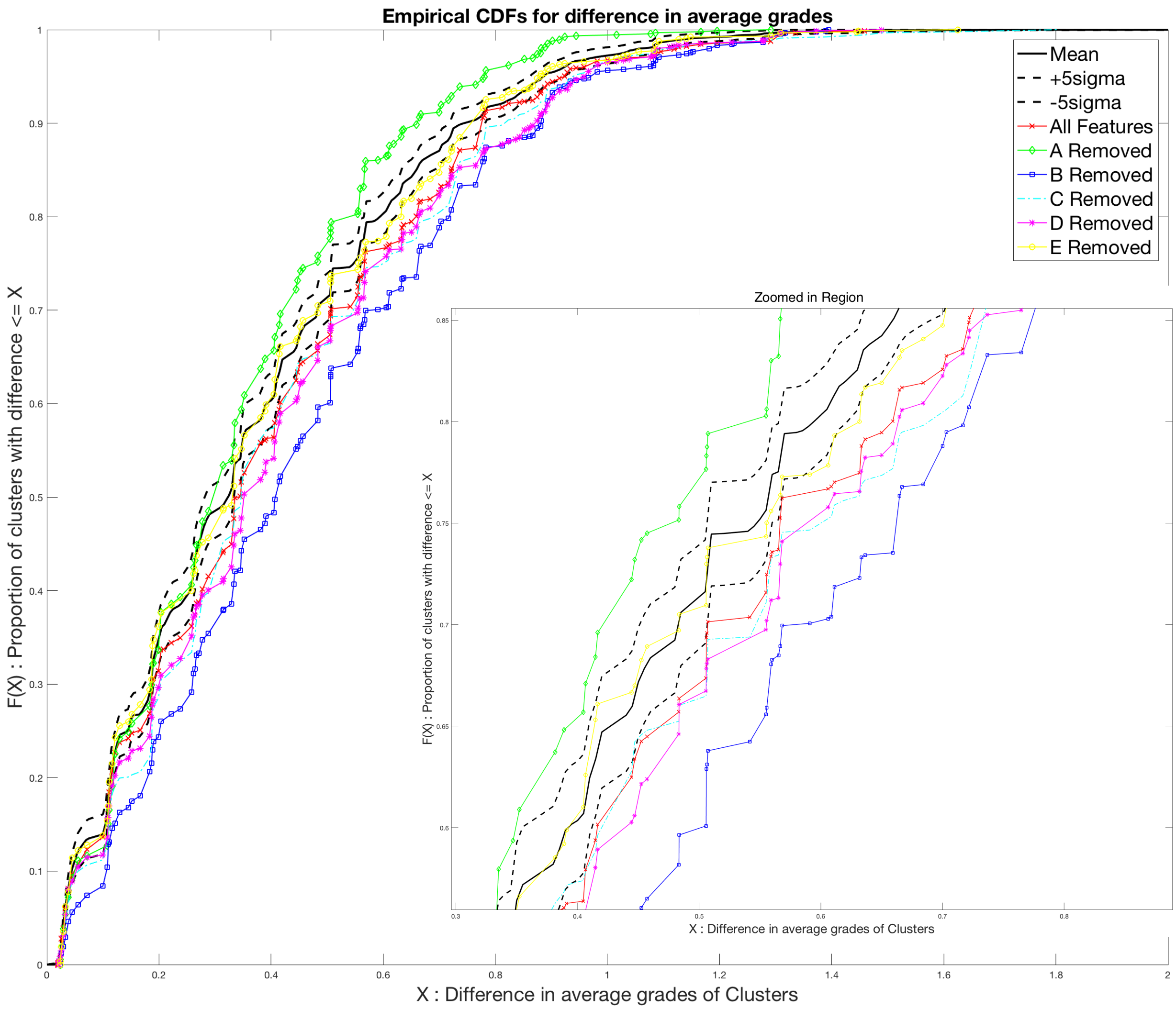


Fig. 4. Cumulative distribution functions (CDF) for absolute value of difference between average grades.

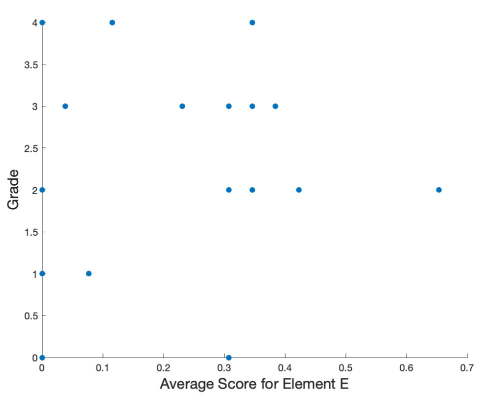


Fig. 5. Comparison of grade with average score for element E

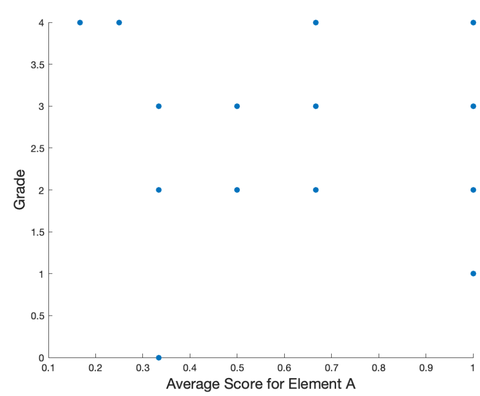


Fig. 6. Comparison of grade with average score for element A

The significance level estimated by the pair of curves for mean $\pm$ 5 sigma corresponds to a significance level of at least 96% based on Chebyshev's inequality [34], [35]. This states that for a random variable X with finite mean μ and finite non-zero variance σ^2 and any real number $k > 0$,

$$Pr(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$$

which for $k = 5$ means that there is at most 4% chance of a

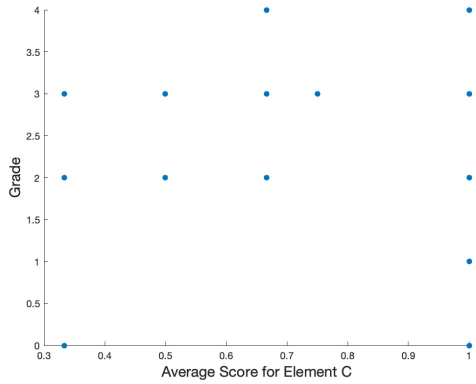


Fig. 7. Comparison of grade with average score for element C

sample lying more than five standard deviations away from the mean. This inequality does not make any assumptions on the underlying distribution of the random variable X and so is a conservative general bound, thereby guaranteeing a minimum significance level of 96% if a realization lies outside the $\pm\ 5\sigma$ boundary of the mean.

In order for the null-hypothesis to be rejected (i.e., to say that the grade differences observed are dependent on patterns of Habits of Mind), the CDF curve obtained with a certain set of Habits of Mind should lie above/below the $\pm\ 5\sigma$ curves at a given point. More specifically, if the CDF curve for a grade difference of at least X based on patterns of Habits of Mind is, say, above the $+5\sigma$ curve or below the -5σ curve at X, then the probability that the observed grade difference of at least X for the proportion of Habits of Mind patterns indicated by the value of the CDF curve at X is due to chance is below 4%.

Recall that removing element A (curve with diamonds on top in Fig. 4) not only shifted the Habits of Mind curve up, it also shifted it higher above the null-hypothesis curve (black curve). So in a statistically significant manner, removing element A reduces the association of the grades with the clusters. In other words, removing the "Computation and Estimation" Habit of Mind from the analysis results decreases the association of the grades with the patterns of the Habits of Mind. On the flip side, removing element E (curve with circles near the middle) results in the CDF curve being pushed up and significantly overlapping with the null-hypothesis curve (dashed curves). This implies that the grade cluster associations from this experiment are not statistically significant. Hence, removing the element "Values and Attitudes" results in patterns of Habits of Mind that are not associated with a statistically valid grade difference. Thus, this element is a pivotal component of the patterns formed by the Habits of Mind associated with a significant grade difference, since its removal results in statistically invalid patterns. Finally, one observes that curves corresponding to retaining all Habits of Mind (curve with crosses), removing element B (curve with diamonds), removing element D (curve with stars) and removing element C (dash-dot curve) one at a time result in curves that are below the null hypothesis curves (dashed curves) for a large range of grade difference values, indicating that the grade cluster associations displayed through these experiments are indeed statistically significant. Therefore, groups formed by either including all Habits of Mind, or all Habits of Mind except "Mathematical Rigor", or all Habits of Mind except "Communication Skills" or all Habits of Mind except "Critical Response Skills" yield patterns that are associated with significant differences in grades in a statistically valid manner.

V. Discussion and Implications for Research, Teaching and Learning

Results from this study suggest that the course grade was dependent on at least two Habits of Mind: (a) Computation and Estimation and (b) Values and Attitudes. The dependency of course grade on computation and estimation is consistent with previous work that suggest that students' ability to choose an appropriate computation method and accurately carry out a mathematical procedure is a critical skill in engineering professionals [1]. Similarly, as reported in previous work on student learning of signals and systems, strong mathematical knowledge is important to succeed in this course [6], [7]. A second dependency of course grade was on Values and Attitudes. In this study values and attitudes were operationalized as students' reactions and insights about others' work; that is, it was operationalized as peer-feedback. Student peer-feedback has been identified as a required skill to function properly in industry as well as educational settings [36]. It has also been identified as a critical form of effective communication skills, problem-solving skills, and professional responsibility to conduct the feedback. Although peer feedback has been widely implemented in engineering education as part of team performance [37], researchers have identified it as difficult to implement when the goal is improving students' answers to open-ended problems [36], [38], [39]. However, when successfully integrated, peer feedback can result in better course performance and higher level thinking skill display such as critical thinking, planning, monitoring, and regulation [40].

Implications for research relate to the use of clustering methods to supplement traditional approaches for data analysis in education research. For instance, if only traditional approaches for qualitative analysis were followed for this study, the investigators would have been limited to characterizing the Habits of Mind as identified in Table II. Specifically, following a traditional qualitative approach would have given understanding and a description of how students' Habits of Mind were enacted in the context of a signals and systems course. Taking this a step further by utilizing quantitative approaches to data analysis allowed the researchers to identify overall patterns of students' performance, Table IV. By utilizing the clustering method the investigators were able to identify several binary groupings (i.e. divisions of the students into two groups) that were found to be statistically significant. One particular grouping was highlighted. The patterns corresponding to the two groups (Habits Developed and Habits Developing students) were compared and contrasted based on their similarities and differences, both in terms of Habits of

Mind elements and levels exhibited, Tables V and VI, and course performance, Table VII. The final step tested whether the grade differences observed for all the different patterns (clusterings) of Habits of Mind were statistically significant, Fig. 4.

The implications for teaching and learning relate to the integration of pedagogies that not only focus on emphasizing the technical or mathematical elements of a course, but also those that integrate critical peer-feedback. The use of slectures appears to foster students' application of signals and systems knowledge along with other skills. That is, having students explaining the course material for a topic of their choice in their own words, as well as reviewing and commenting on the slectures prepared by their peers, may be an appropriate approach to help students develop Habits of Mind [24].

VI. Conclusions, Limitations and Future Work

This paper looked at how engineering students exhibited Habits of Mind in the context of student-generated content for a course on signal processing. The five Habits of Mind investigated were Computation and Estimation, Mathematical Rigor, Communication Skills, Critical Response Skills, and Values and Attitudes. A quantitative analysis based on random signal modeling and clustering was performed. The model assumed independence of the vector tags used to annotate the student slectures, which is a simplifying assumption for a first order model. A more complex model that relaxes this assumption and potentially models the data better, requires a larger number of data samples than were available.

Students were found to exhibit various different patterns of Habits of Mind (binary groupings). One such pattern (grouping) that was found to affect grade was analyzed: the main difference between these particular groups was found to be the level of proficiency of all the Habits of Mind elements. Thus the groups were designated as "Habits Developed" and "Habits Developing", respectively. Further analysis of the entire set of patterns (groupings) found by clustering revealed that many patterns of Habits of Mind affect grades, and that the grade is directly dependent on Computation and Estimation and Values and Attitudes. The main limitation of the study is the dependency of the proposed method on qualitative approaches to hand-scoring the data. The small sample size allows iterative scoring of the data by hand, and validation of such scoring by multiple raters. This step of the method will be harder to replicate with larger samples. While this study is limited in scope and size, it will be interesting to see if these results are confirmed in other electrical engineering core courses. It would also be interesting to conduct a comparative study between students who did slectures and those who did not. The analysis framework proposed is applicable in many other contexts and data types (e.g., video data or think-alouds) and could be used to study the relationship between other skills and educational outcomes.

VII. Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. EEC-1544244 and EEC-1826099.

References

[1] A. S. for Engineering Education, "Transforming Undergraduate Education in Engineering (TUEE)," *ASEE*, 2013.

[2] N. R. Council and Others, *Transforming undergraduate education in science, mathematics, engineering, and technology*. National Academies Press, 1999.

[3] E. A. C. ABET, "Criteria for Accrediting Engineering Programs Effective for Reviews during the 2017-2018 Accreditation Cycle," 2016.

[4] G. W. Clough and Others, "The engineer of 2020: Visions of engineering in the new century," *National Academy of Engineering, Washington, DC*, 2004.

[5] American Association for the Advancement of Science, "Project 2061: Science literacy for a changing future: A decade of reform." American Association for the Advancement of Science, 1995.

[6] K. Wage, J. Buck, C. Wright, and T. Welch, "The Signals and Systems Concept Inventory," *IEEE Transactions on Education*, vol. 48, no. 3, pp. 448–461, aug 2005. [Online]. Available: http://ieeexplore.ieee.org/document/1495653/

[7] M. D. Campbell, R. C. Houts, and E. A. Reinhard, "A Computer Utility Incorporating the FFT Algorithm for a Signal and System Theory Course," *IEEE Transactions on Education*, vol. 16, no. 1, pp. 42–47, 1973. [Online]. Available: http://ieeexplore.ieee.org/document/4320788/

[8] A. M. Goncher, D. Jayalath, and W. Boles, "Insights Into Students' Conceptual Understanding Using Textual Analysis: A Case Study in Signal Processing," *IEEE Transactions on Education*, vol. 59, no. 3, pp. 216–223, aug 2016. [Online]. Available: http://ieeexplore.ieee.org/document/7403926/

[9] S. Henderson and E. H. Segal, "Visualizing qualitative data in evaluation research," *New Directions for Evaluation*, vol. 2013, no. 139, pp. 53–71, 2013. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/ev.20067

[10] A. L. Costa and B. Kallick, *Habits of mind across the curriculum: Practical and creative strategies for teachers*. ASCD, 2009.

[11] S. B. Merriam, *Case study research in education: A qualitative approach.* Jossey-Bass, 1988.

[12] K. M. Eisenhardt, "Building Theories from Case Study Research," *Academy of Management Review*, vol. 14, no. 4, pp. 532–550, oct 1989. [Online]. Available: http://journals.aom.org/doi/10.5465/amr.1989.4308385

[13] R. K. Yin, *Case study research: Design and methods vol 5*. SAGE Publications, 2009.

[14] M. Q. Patton, *Qualitative Research and Evaluation Methods, 2nd ed.* SAGE Publications, 2002.

[15] M. B. Miles and A. M. Huberman, *Qualitative data analysis: An expanded sourcebook*. SAGE Publications, 1994.

[16] J. Hepp, Y. Tarun, and M. Boutin, "Code and Dataset for Pattern Recognition Benchmarks," dec 2016. [Online]. Available: https://purr.purdue.edu/publications/2030/2

[17] T. Yellamraju and M. Boutin, "Pattern Dependence Detection using n-TARP Clustering," *arXiv preprint arXiv:1806.05297*, June 2018. [Online]. Available: http://arxiv.org/abs/1806.05297

[18] T. Yellamraju, J. Hepp, and M. Boutin, "Benchmarks for Image Classification and Other High-dimensional Pattern Recognition Problems," *arXiv preprint arXiv:1806.05272*, June 2018. [Online]. Available: http://arxiv.org/abs/1806.05272

[19] T. Y. Taylor V. Williams, Kerrie A. Douglas and M. Boutin, "Characterizing MOOC learners from survey data using modeling and n-tarp clustering," in *2018 ASEE Annual Conference & Exposition*. Salt Lake City, Utah: ASEE Conferences, June 2018, https://peer.asee.org/30186.

[20] H. Zou and T. Hastie, "Regularization and variable selection via the elastic net," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 67, no. 2, pp. 301–320, apr 2005. [Online]. Available: http://doi.wiley.com/10.1111/j.1467-9868.2005.00503.x

[21] S. Han and M. Boutin, "The hidden structure of image datasets," in *2015 IEEE International Conference on Image Processing (ICIP)*, IEEE. IEEE, Sept. 2015, pp. 1095–1099. [Online]. Available: http://ieeexplore.ieee.org/document/7350969/

[22] T. Yellamraju and M. Boutin, "Clusterability and Clustering of Images and Other Real High-Dimensional Data," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 1927–1938, apr 2018. [Online]. Available: http://ieeexplore.ieee.org/document/8245810/

[23] A. W. Haddad and M. Boutin, "Rhea: a student-driven tool for enhancing the educational experience," *Journal of Computing Sciences in Colleges*, vol. 26, no. 1, pp. 59–66, 2010.

[24] M. Boutin and J. Lax, “Engaging graduate students through online lecture creation,” in *2015 IEEE Frontiers in Education Conference (FIE)*, IEEE. IEEE, oct 2015, pp. 1–4. [Online]. Available: http://ieeexplore.ieee.org/document/7344169/

[25] C. Steinkuehler and S. Duncan, “Scientific Habits of Mind in Virtual Worlds,” *Journal of Science Education and Technology*, vol. 17, no. 6, pp. 530–543, dec 2008. [Online]. Available: http://link.springer.com/10.1007/s10956-008-9120-8

[26] T. Yellamraju, A. J. Magana, and M. Boutin, “Board # 11 : Investigating Engineering Students Habits of Mind: A Case Study Approach,” in *2017 ASEE Annual Conference & Exposition*. Columbus, Ohio: ASEE Conferences, jun 2017. [Online]. Available: https://peer.asee.org/27686

[27] K. Pearson, “Note on regression and inheritance in the case of two parents,” *Proceedings of the Royal Society of London*, vol. 58, pp. 240–242, 1895.

[28] S. M. Stigler, “Francis Galton’s account of the invention of correlation,” *Statistical Science*, pp. 73–79, 1989.

[29] K. A. Hallgren, “Computing inter-rater reliability for observational data: an overview and tutorial,” *Tutorials in quantitative methods for psychology*, vol. 8, no. 1, p. 23, 2012.

[30] A. Papoulis and S. U. Pillai, *Probability, random variables, and stochastic processes*. Tata McGraw-Hill Education, 2002.

[31] T. Yellamraju, “n-tarp: A random projection based method for supervised and unsupervised machine learning in high-dimensions with application to educational data analysis,” Ph.D. dissertation, Purdue University, 2019.

[32] H. Wang and M. Song, “Ckmeans.1d.dp: Optimal k-means Clustering in One Dimension by Dynamic Programming.” *The R journal*, vol. 3, no. 2, pp. 29–33, Dec. 2011. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/27942416 http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC5148156

[33] M. D. Ernst, “Permutation Methods: A Basis for Exact Inference,” *Statistical Science*, vol. 19, no. 4, pp. 676–685, Nov. 2004. [Online]. Available: http://projecteuclid.org/euclid.ss/1113832732

[34] P. L. Chebyshev, “Des valeurs moyennes, Liouville’s,” *J. Math. Pures Appl.*, vol. 12, pp. 177–184, 1867.

[35] W. Feller, *An introduction to probability theory and its applications*. John Wiley & Sons, 2008, vol. 2.

[36] K. J. Rodgers, H. A. Diefes-Dux, and M. E. Cardella, “The nature of peer feedback from first-year engineering students on open-ended mathematical modeling problems,” in *American Society for Engineering Education*. American Society for Engineering Education, 2012.

[37] J. Smith, G. Hoffart, and T. O’Neill, “Peer Feedback on Teamwork Behaviors: Reactions and Intentions to Change,” in *American Society for Engineering Education*. American Society for Engineering Education, 2016.

[38] J. McGourty, P. Dominick, and R. Reilly, “Incorporating student peer review and feedback into the assessment process,” in *FIE ’98. 28th Annual Frontiers in Education Conference. Moving from ’Teacher-Centered’ to ’Learner-Centered’ Education. Conference Proceedings (Cat. No.98CH36214)*, vol. 1, IEEE. IEEE, 1998, pp. 14–18. [Online]. Available: http://ieeexplore.ieee.org/document/736790/

[39] R. Rada and Ke Hu, “Patterns in student-student commenting,” *IEEE Transactions on Education*, vol. 45, no. 3, pp. 262–267, aug 2002. [Online]. Available: http://ieeexplore.ieee.org/document/1024619/

[40] Eric Zhi-Feng Liu, S. Lin, Chi-Huang Chiu, and Shyan-Ming Yuan, “Web-based peer review: the learner as both adapter and reviewer,” *IEEE Transactions on Education*, vol. 44, no. 3, pp. 246–251, 2001. [Online]. Available: http://ieeexplore.ieee.org/document/940995/

Tarun Yellamraju received his Bachelors degree in electrical engineering with Honors from the Indian Institute of Technology - Bombay, India in 2015, and his Ph.D. in electrical and computer engineering from Purdue University in 2019, where he worked on high-dimensional machine learning methods.

Alejandra J. Magana received her B.E. in information systems and M.S. in technology from Tec de Monterrey, and a M.S. in educational technology and a Ph.D. in engineering education, both from Purdue University. She is an Associate Professor in the Department of Computer and Information Technology with a courtesy appointment in the School of Engineering Education at Purdue University. Her research program investigates how model-based cognition in Science, Technology, Engineering, and Mathematics (STEM) can be better supported by means of expert technological and computing tools such as cyber-physical systems, and computational modeling and simulation tools. In 2015 Dr. Magana received the National Science Foundations Faculty Early Career Development (CAREER) Award, to investigate modeling and simulation practices in undergraduate engineering education. In 2016 she was conferred the status of Purdue Faculty Scholar for being on an accelerated path toward academic distinction. Dr. Magana serves as Associate Editor for the Computer Applications in Engineering Education journal and Associate Editor for the Journal of Engineering Education.

Mireille Boutin received a B.SC. in physics-mathematics from the University of Montreal and a Ph.D. in mathematics from the University of Minnesota. She is an Associate Professor in the School of Electrical and Computer Engineering at Purdue University, with a courtesy appointment in the Department of Mathematics. Her research interests are in the areas of signal processing, machine learning and applied mathematics. She is a member of the Image, Video, and Multidimensional Signal Processing Technical Committee (IVMSP TC) of the IEEE Signal Processing Society. Dr. Boutin previously served as Associate Editor for the IEEE Signal Processing Letters and was on the editorial team of the Springer journal Applicable Algebra in Engineering, Communication and Computing. She is currently an Associate Editor for the IEEE Transactions on Image Processing.