A Psycholinguistic Analysis of Inter-Ethnic Views of Ethics

Roman Taraban
https://orcid.org/0000-0002-1815-4687
Scopus Author ID: 6603757476
roman.taraban@ttu.edu
Texas Tech University, USA

Sweta Saraff
https://orcid.org/0000-0001-9688-1612
Scopus Author ID: 57208392821
ssaraff@kol.amity.edu
Amity University Kolkata, India

Serhii Zasiekin
https://orcid.org/0000-0001-9453-3534
Scopus Author ID: 35112194200
sergiozasyekin@vnu.edu.ua
Lesya Ukrainka Volyn National University, Ukraine

Ramakrishna Biswal
https://orcid.org/0000-0003-3681-9790
Scopus Author ID: 32867511800
biswalrk@nitrkl.ac.in
National Institute of Technology in Rourkela, India

Abstract. In this paper we describe a collaborative, cross-cultural project whose focus is on developing an interactive pedagogy for undergraduate students that delivers a culturally relevant ethical activity, using an internet platform to widen the scope of learning at an international level. The ongoing collaborative project described here is implemented through a website, the Ethical Engineer (https://ethicalengineer.ttu.edu), that provides an open and unique online digital platform for developing a community of students committed to the discussion of ethics from a global perspective. A goal of this project was to gain insights into similarities and differences between Indian, Ukrainian, and USA respondents in their reasoning about a hypothetical ethical dilemma. Two complementary methods of analysis were applied to students' responses. First, a well-known method for text analysis, Linguistic Inquiry and Word Count (LIWC-22) was used to identify topics and perspectives in students' essays that distinguished the three groups. These analyses were followed by naïve Bayesian analyses that were used to further identify characteristic conceptual differences between the three ethnic groups. We discuss how the Ethical Engineer platform could be further developed by providing students with immediate substantive feedback to their written responses, in order to extend the

learning benefits to students from participation on the website. The analytic methods described here demonstrate ways to gain knowledge of globally-diverse students' thinking.

Keywords: ethnicity, culture, text analytics, ethics, Bayesian methods, LIWC-22.

Тарабан Роман, Сараф Свєта, Засєкін Сергій, Бісвал Рамакрішна. Психолінгвістичне дослідження міжетнічних поглядів на етику.

Анотація. У цій статті ми описуємо спільний міжкультурний проєкт, зосереджений на розробці інтерактивної педагогіки для студентів бакалаврату, яка забезпечує культурно зумовлену етичну поведінку, використовуючи інтернет-платформу для розширення сфери навчання на міжнародному рівні. Описаний тут поточний спільний проєкт реалізується на основі веб-сайту Ethical Engineer (https://ethicalengineer.ttu.edu), покликаного забезпечити відкриту та унікальну онлайнову цифрову платформу для розвитку спільноти студентів, відданих обговоренню етики в глобальному контексті. Мета цього проєкту полягала в тому, щоб зрозуміти подібності та відмінності між респондентами з Індії, України та США в їхніх міркуваннях щодо гіпотетичної етичної дилеми. До відповідей студентів було застосовано два взаємодоповнюючі методи аналізу. По-перше, добре відомий у психолінгвістиці метод аналізу тексту Linguistic Inquiry and Word Count (LIWC-22) був використаний для визначення тем і поглядів у студентських есе у трьох групах. Цей аналіз супроводжував наївний байєсів аналіз, застосований для подальшого визначення характерних концептуальних відмінностей між трьома етнічними групами. Автори обговорюють, як можна в подальшому розвивати платформу Ethical Engineer, надаючи студентам негайний змістовний зворотний зв'язок на їхні письмові відповіді, щоб розширити переваги навчання студентів від взаємодії з веб-сайтом. Описані аналітичні методи демонструють способи отримання знань про різноманітність мислення студентів із віддалених куточків нашої планети.

Ключові слова: *етнічність, культура, текстова аналітика, етика, байєсів метод, LIWC-22.*

> *Scandals tend to break not at the point people "find out" about bad behavior…but when they think a majority of others judge it to be wrong. People, after all, rarely make ethical judgments in a vacuum.*
> (Martha Gill, The New York Times, July 7, 2022)

Introduction

In the quote above, which appeared in a recent editorial essay in The New York Times newspaper, a simple yet powerful point is made regarding ethical thinking: ethical judgements of right and wrong are often made by groups and rarely by individuals. This suggests that personal growth in the discernment of what is right and just will likely take place in a group setting involving discourse, interaction, and consensus building. To the extent that this suggestion is valid, it implies a rather straightforward path in school curricula. However, there may be complications, particularly at upper levels of education. According to Muller (2009), there is a fundamental rift in college disciplines between the sciences and humanities. To Muller, this breach, which he traces back to the medieval university, has been a social threat across the ages, and is well captured in a quote from Muller where he states: "Science without the Humanities becomes a technicist Frankenstein" (p. 206).

Both the New York Times and Muller's (2009) technicist Frankenstein are especially poignant to pedagogical work that we have been conducting since 2017 through the development and launching of a globally accessible website for the discussion of ethical dilemmas by university undergraduates. Our work originated as

a collaborative project implemented through a website, the Ethical Engineer[1], that provides an open digital forum for interactive discourse across national boundaries. Participants read and reflect on case studies involving ethical dilemmas that are presented on the website. Students post their reactions to the case studies directly to the website. They also post reactions to other students' responses. Instructor-generated guidelines for developing a response to a case study are provided on the website to help students think critically and gain clarity on their position. Detailed descriptions of the Ethical Engineer can be found at Taraban et al. (in press, 2020).

The Ethical Engineer has been used primarily with engineering students in the USA and India as an instructional resource, typically implemented as a one-time homework or classroom activity. These students come from disciplines oriented towards the Sciences and mechanical applications, as described in Muller (2009). The primary research question in the present project was whether differences in conceptualizations about ethical issues would emerge when comparing those students to students more oriented towards the Humanities and social applications. To this end, we collected website responses from students in Ukraine in language sciences. A comparison of students' ethical thinking across national and disciplinary boundaries is relevant to an assessment of the impact of the Ethical Engineer as an instructional resource. Crossing disciplinary boundaries, as we do here, addresses Muller's concern regarding the fracture between science and humanities, or mechanical and liberal studies. Should a range of conceptual perspectives coalesce on the website, there would be suggestive evidence for the Ethical Engineer as a vehicle for constructive dialogue across national and disciplinary boundaries.

Method

Text Analytics as a Key to Students' Thinking

A fundamental theoretical principle underpinning an analysis of verbal communications comes from the work of Pennebaker and King (1999) who proposed that "the way people talk about things reveals important information about them" (p. 1297). Pennebaker et al. (2015) reasoned that it is possible to identify specific "beliefs, fears, thinking patterns, social relationships, and personalities" (p. 1) that characterize individuals based on the words that they use. In an instructional context, the Pennebaker and King principle potentially empowers researchers and teachers to uncover the cognitive and affective orientations of students in specific coursework and activities. Further, the ability to grasp key constructs that students are employing potentially aids instructors in understanding why a specific activity is effective, and in configuring and implementing beneficial instructional activities.

The present study applied the Pennebaker and King (1999) principle to writing samples from USA, Indian, and Ukrainian students that were submitted on the Ethical Engineer website. Two text analytic methods were applied to these samples. The methods incorporated the Pennebaker and King principle in different ways. One

---

[1] Available at https://ethicalengineer.ttu.edu

method used pre-defined word lists, termed dictionaries, for analysis of student responses; the second method created weighted word lists from the responses themselves and applied these to the analysis of student responses. These methods are described next.

LIWC-22: A Text Analytic Tool

A well-known method for text analysis, LIWC-22 (Linguistic Inquiry and Word Count) (Boyd et al., 2022), uses predefined dictionaries. LIWC-22 quantifies the frequencies of words (or stems) that fall into specific categories, like Cognition, Perception, Social, and Culture. These frequencies are converted into each word's percentage of the total number of words in the text. In that way, the percentages show the relative emphasis of some categories over others, adjusted for the total length of the text. The percentages can then be correlated with independent variables, like course grades or likelihood of academic success (Pennebaker et al., 2014; Robinson et al., 2013).

LIWC-22 includes four summary categories. These categories were constructed from a combination of specific categories by Boyd et al., (2022), were normed based on extensive sampling of data, and are expressed on a percentile scale. The four summary variables as summarized in Boyd et al. (2015) are:

- **Analytic Thinking -** Higher percentile scores suggest formal, logical, and hierarchical thinking, whereas, lower scores reflect more informal, personal, here-and-now, and narrative thinking.
- **Clout -** Higher percentile scores show confidence, whereas lower scores suggest a tentative and anxious style.
- **Authentic –** Higher scores are associated with honest, personal, and disclosing communication; lower scores indicate a more guarded, distanced form of discourse.
- **Emotional Tone** – Higher scores are associated with a positive, upbeat attitude; lower scores indicate greater anxiety, sadness, or hostility. A number around 50 suggests a lack of emotionality.

Recent research has used LIWC variables in order to assess communication styles in students' writing using a metric known as the Categorical-Dynamic Index (CDI). The CDI can be separated into two components: a categorical component and a dynamic component. The categorical component of the CDI assesses analytical thinking, i.e. deep reflection, and logical reasoning (Jordan & Pennebaker, 2017). The dynamic component of the CDI assesses narrative thinking, i.e. a more informal "here and now" communication style (Pennebaker et al., 2015). While analytical thinkers tend to scrutinize a problem, narrative thinkers prefer to share their own experiences and tell stories to understand the problem. Therefore, analytical thinking relies on facts, details, and objective treatment of events. By contrast, narrative thinking relies more on intuition, personal account of events, and snap judgements (Jordan & Pennebaker, 2016). Linguistically, the analytic style, as the authors confirmed, is marked by the high use of nouns, articles, and prepositions. Their use in text/speech indicates that the author/speaker is identifying conceptual categories and organizing

them in hierarchical ways. The narrative style is marked by the high use of personal pronouns, auxiliary verbs, common adverbs, and function words.

Thus, Categorical and Dynamic scores are calculated by applying Equations (1) and (2), respectively, using percentages provided by the LIWC-22 software for the variables shown:

$$\text{Categorical} = (\text{articles} + \text{prep})/2^1 \tag{1}$$

$$\text{Dynamic} = (\text{ppron} + \text{ipron} + \text{auxverb} + \text{conj} + \text{adverb} + \text{negate})/6 \tag{2}$$

In summary, LIWC-22 is a general purpose analytic tool and can be reliably applied across a variety of contexts. In addition to the summary variables, LIWC-22 includes the percentages of total text for specific predictors. The four types of measures summarized here were applied to the research data in this study in order to assess conceptual similarities and differences in student responses, by country.

Naïve Bayesian Text Analytic Method

An alternative to applying pre-compiled dictionaries to carry out text analysis uses computations based on an extension of Bayes theorem. The naïve Bayesian method is used to create classifiers that identify predictors that are able to classify old and new instances. The predictors are the actual words in the texts. For instance, after training on a set of newspaper editorials written from reactionary and liberal perspectives, a Bayesian classifier can be used to classify new editorials based on the discriminating predictors within the texts, that is, based on the words that authors use in the respective editorials. The appeal of naïve Bayesian classifiers is their capacity to be trained to classify content in specific subject domains. The strength of Bayesian classifiers is their ability to correctly classify new instances after being trained on a sample of known instances.

In the present study, a naive Bayesian classifier was trained on a sample of students' responses. The classifier was then tested on its ability to classify new instances by country. Successful classification would imply that students' responses contained conceptual (i.e., as indicated by the use of different words) differences reflective of their country of origin.

Case Study Applying LIWC-22 and Naïve Bayesian Analyses

Students from the USA, Ukraine, and India participated in this study. They were compensated with a small course credit. The conduct of research had approval from the respective university review boards. The primary material on the Ethical Engineer website that was used in this study is the case study: "Which Is More Important –

---

[1] *Note.* Abbreviations - prep: preposition; ppron: person pronouns; ipron: impersonal pronouns; auxverb: auxiliary verbs; conj: conjunctions; negate: negations

Environmental Concern or Economic Growth?" by Dr. Sudipta Majumdar (then Amity University Kolkata, India; currently, Faculty of Management Studies, ICFAI University, Jharkhand, India).[1] Participants logged on to the Ethical Engineer website, reviewed the case study, and posted a reaction. The case study and all student responses were posted in English.

Analyses of student responses were conducted through the applications of the LIWC-22 software[2] and the naïve Bayesian method, applied in the R language. The LIWC-22 analyses compared student responses to pre-compiled dictionaries and reported percentile and percent matches of words in student responses to words in the dictionaries, by country. The naïve Bayesian method weighted words across the responses in a way that reliably separated responses by country. A major difference between the two methods is that the LIWC-22 analysis functioned with advance specification of the country to which a student response belonged. The Bayes method used the country to which each student response belonged training (roughly 75% of available responses). On the critical test trials (i.e., the remaining 25% of the responses), the task of the Bayesian method was to classify the new responses, i.e., responses for which the Bayes method did not know the country in advance.

Results

LIWC-22 Results

Analysis of the four summary variables was conducted first and was used as a reference point for the subsequent analyses. Means and standard deviations for the summary variables are shown in Table 1.

Table 1
Mean Percentiles (standard deviations in parentheses) for LIWC-22 Summary Variables, by Country

LIWC-22 Summary Variables	Country		
	USA (n = 25)	Ukraine (n = 21)	India (n = 25)
Analytic	81.33 (9.61)[a]	67.32 (15.17)[b]	80.02 (14.19)[a]
Clout	40.49 (13.66)	39.02 (16.00)	40.77 (12.59)
Authentic	32.72 (15.10)[ab]	41.19 (15.34)[a]	28.18 (17.96)[b]
Tone	33.29 (20.51)	38.45 (18.85)	34.66 (20.84)

Note. Superscripts are used to signify significant differences between cells. When superscripts differ, there is a significant difference between Country for those LIWC-22 variables at $p < .05$. All tests are two-tailed.

---

[1] The case study can be viewed at https://ethicalengineer.ttu.edu
[2] Information regarding LIWC-22 software can be found at https://liwc.app

GLM-Univariate (IBM, 2022[1]) was applied to the variables shown in Table 1, with each LIWC-22 variable treated as a dependent variable and Country treated as the independent variable. An initial analysis showed the data were normally distributed. There were significant Country differences for Analytic and Authentic [$F(2.70) = 7.76$, $p < .001$; and $F(2.70) = 3.73$, $p = .029$, respectively], and non-significant differences for Clout and Tone (p-values $> .50$). Follow-up mean comparisons using Tukey HSD tests, showed that USA and India significantly exceeded Ukraine in the Analytic factor (p-values $< .005$), and Ukraine exceeded India in the Authentic factor ($p = .023$).

The results from these ANOVAs provide the first major results in this study. The USA and Indian participants expressed more formal, logical, and hierarchical thinking, compared to the Ukrainian participants; whereas Ukrainian participants expressed more informal, personal, here-and-now, and narrative thinking.

The significantly higher score on the Authentic variable for Ukrainian participants compared to Indian participants is consistent with the Analytic effects, since higher scores on the Authentic variable are associated more honest, personal, and disclosing communications; lower numbers suggest a more guarded, distanced form of discourse, which is more consistent with an Analytic perspective. Although not significant, the USA score for Authentic (32.72) is similar to the Indian score (28.18), and shows a notably lower score for Authentic compared to the Ukrainian participants.

There were no significant differences for Country for Clout and Tone. According to the LIWC handbook (Boyd et al., 2022), lower Clout numbers, as in the present case, suggest a tentative, humble, even anxious style. Using 50 as a neutral emotionality level for Tone, participants across countries expressed anxiety and sadness in their responses. From these results, we can infer that participants, across Country, responded similarly to the situation described in the ethical dilemma.

In order to further examine Country differences, additional tests were conducted for the specific LIWC-22 content variables, within LIWC-22 superordinate categories. Because a majority of data distributions for these variables were not normally distributed, the Kruskal-Wallis test (IBM, 2022) was applied, which is an appropriate test of significant differences when data distributions are not normally distributed. Table 2 shows the variables that reliably distinguished between categories. Sample words from those LIWC-22 categories are included in order to convey a sense of the gist of the category. Additional pairwise mean comparisons were carried out for each LIWC-22 variable in order to assess which countries significantly differed from one another. These differences ($p < .05$, two-tailed) are marked using superscripts.

---

[1] Information available at https://www.ibm.com

Table 2
Mean Percents (standard deviations in parentheses) for LIWC-22 Variables, By Country. Superordinate Categories Are Shown with Sample Variables and Words That Define the Category

	Country		
LIWC-22 Variables	USA (n = 25)	Ukraine (n = 21)	India (n = 25)
COGNITION			
all[vs]none: all kind, always, everyday, forever, never, no one, nothing, nowhere	.78 (053)[b]	1.28 (.82)[a]	0.68 (0.63)[b]
discrep[ancy]: deserve, desire, hope, ideal, lack, need, ought, try, undo, wish, yearn	2.70(.79)[ab]	2.19 (0.84)[b]	3.55 (1.85)[a]
differ: adjust, differ, disparity, distinct, equality, exception, identical, impossible, not really	2.65 (1.16)[b]	3.32 (1.26)[ab]	3.66 (1.50)[a]
SOCIAL PROCESSES			
polite: civility, courteous, decency, please, regard, respectful, sincerely, thank, welcome	.26 (.35)[a]	.35 (.34)[a]	.09 (.21)[b]
male: dad, father, he, him, his, man	.29 (.43)[a]	.37 (.85)[a]	.05 (.21)[b]
comm[unication]: answer, ask, call, confess, consult, discuss, inform, protest, question	.42 (.48)[b]	1.18 (.78)[a]	.40 (.59)[b]
CULTURE			
tech: automate, engineer, firewall, glitch, hardware, innovate, machine, multimedia	.62 (.59)[a]	.18 (.32)[b]	.34 (.37)[ab]
LIFESTYLE			
home: bed, broom, carpet, garden, house, kitchen, neighborhood, pets, room, window	.15 (.31)[b]	.41 (.41)[a]	.07 (.18)[b]
PHYSICAL			
reward: accumulate, benefit, bonus, earn, fortune, gain, lucrative, pay off, succeed, win	.80 (.48)[a]	.57 .35)[ab]	0.37 (.40)[b]
allure: beautiful, country, delicious, dream, easy, family, friend, fun, make, money, work	5.37 (1.42)[a]	6.06 (1.93)[a]	4.38 (1.86)[b]
wellness: cardio, diet, fitness, heal, healthy, mindfulness, organic, positive mood, self care	.11 0.33)[ab]	.06 (.29)[b]	.19 (.33)[a]
PERCEPTION			
attention: alarm, alert, cognizant, inspect, negligent, notice, oversee, vigilance	.32 (.35)[b]	.50 (.57)[ab]	.85 (.69)[a]
feeling: cold, dry, grasp, heat, hottest, hurt, numb, pain, sensation, tingle, touch, tremble	.08 (.16)[b]	.09 (.18)[b]	.35 (.50)[a]
focus past: advised, assumed, avoided, became, caused, didn't, lied, tricked, warned	2.65 (1.29)[a]	1.49 (1.00)[b]	2.21 (1.24)[a]

Note. All LIWC-22 variables shown in bold differ significantly by Country ($p < .05$). Superscripts are used to signify significant differences between specific countries. When superscripts differ, there is a significant difference between Country for those LIWC-22 variables ($p < .05$, uncorrected). All tests are two-tailed.

Of the eight superordinate content categories, six categories contained words that reliably distinguished between countries, as shown in Table 2. The prior results showed that USA and Indian participants expressed more formal, logical, and hierarchical thinking, compared to the Ukrainian participants; whereas Ukrainian participants expressed more informal, personal, here-and-now, and narrative thinking. The more detailed analyses here are consistent with those results. USA participants were high in LIWC-22 variables *tech, reward*, and *focus past*, which are associated with terms like *automate, engineer, machine, bonus, earn, lucrative, assumed, avoided,* and *tricked*. Indian participants were formal and logical as well, although they expressed that differently, through LIWC-22 variables like *discrepancy, differ,* and *attention*, which are associated with terms like *ideal, ought, yearn, adjust, disparity, impossible, cognizant, negligent,* and *vigilance*. Ukrainian participants expressed more informal and personal ideas consistent with LIWC-22 variables *communication, home,* and *allure*, and associated words like *discuss, inform, question, house, neighborhood, pets, dream, family,* and *friend*.

In a final application of LIWC-22 to participants' responses to the ethical dilemma, measures of Categorical and Dynamic thinking were calculated according to equations [1] and [2] described earlier. Mean values and standard deviations are shown in Table 3.

Table 3
Mean Percents (standard deviations in parentheses) for Categorical and Dynamic Indices, by Country

Categorical – Dynamic Indices	Country		
	USA (n = 25)	Ukraine (n = 21)	India (n = 25)
Categorical	12.44 (1.29)	11.31 (1.46)	11.83 (2.14)
Dynamic	4.81 (.62)[b]	5.39 (.80)[a]	4.61 (.67)[b]

Note: Superscripts are used to signify significant differences between specific countries. When superscripts differ, there is a significant difference between Country ($p < .05$). All tests are two-tailed.

GLM-Univariate (IBM, 2022) was applied to the Categorical and Dynamic data shown in Table 1, with the Categorical and Dynamic variables treated as dependent variables and Country treated as the independent variable. There were significant Country differences for the Dynamic Index [$F(2, 60) = 7.50$, $p < .001$]. The differences for Categorical were marginally significant ($p = .082$). Follow-up mean comparisons using Tukey HSD tests showed that Ukraine significantly exceeded USA and India in the Dynamic index (p-values < .02). Because the Dynamic index is associated with a more casual communication style, it is consistent with the earlier findings showing Ukrainian participants to be lower in Analytic and higher in

Authentic variables. Although the Country differences were statistically marginal for the Categorical variable, the means are higher for USA and India, compared to Ukraine, signaling a preference for more formal communication styles, which is consistent with the earlier finding showing USA and Indian participants' as being more Analytic than Ukrainian students.

Naïve Bayes Results

In order to carry out an analysis using naïve Bayesian methods to discriminate Ethical Engineer responses by country, a Bayesian classifier was trained on approximately 75% of the responses and was then tested on the remaining responses. In the training phase, the classifier learned the discriminating concepts (i.e., words) that distinguished responses by country. In the test phase, the classifier predicted whether an essay was from a USA, Ukrainian, or Indian participant.

Sample classification results are shown in Table 4. The confusion matrix shows 95% classification accuracy for new responses classified by the naïve Bayes classifier. Over multiple tests, the naïve Bayes method incorrectly classified no more than one out of twenty new responses incorrectly. In the present example, an Indian response was misclassified. In other runs of the classifier, a USA or Ukrainian response could be misclassified, depending on the responses that are randomly selected for training the classifier.

Table 4

Confusion Matrix Showing Frequencies (Percents in Parentheses) for Classification of 20 New Responses, by Country

Naïve Bayes Results				
Predicted	**Actual**			
	USA	Ukraine	India	Row Total
USA	7 (0.35)	0	1 (0.05)	8
Ukraine	0	6 (0.30)	0	6
India	0	0	6 (0.30)	6

A Psycholinguistic Analysis of Inter-Ethnic Views of Ethics

Naïve Bayes provides for the extraction of the Bayesian conditional probabilities associated with the individual predictors that are used for classification by country. The predictors are the words in students' responses. Combining the probabilities associated with predictors with the *a priori* probabilities of each country in the training trials, it is easy to compute the Bayesian probabilities for each of the predictors in the corpus. Knowledge of the strongest predictors for each country can give insights into how students from the respective countries differed in the most central concepts in their responses. The strongest predictors, by country, are shown in Table 5, with their associated probabilities.

Table 5
Naïve Bayes Rank-Ordered Predictors and Probabilities, by Country

Predictors	USA	Predictors	Ukraine	Predictors	India
mexico	1.00	geography	1.00	degradation	1.00
negatively	1.00	economics	.91	loss	1.00
perspectives	1.00	difficult	.86	proper	.88
positively	.88	authorities	.86	basic	.83
impacted	.85	everything	.86	effective	.83
taking	.83	main	.83	sustainable	.83
lack	.81	really	.83	increased	.83
regulations	.81	ecology	.81	research	.83
engineering	.77	course	.78	contamination	.80
innovative	.77	article	.75	greed	.80
impacts	.75	business	.75	needs	.80
serious	.75	lot	.74	importance	.77
engineers	.73	parties	.71	increasing	.77
affect	.71	positive	.71	organizations	.77
cancer	.71	state	.71	ecosystem	.75
decisions	.71	damage	.71	individuals	.75
gain	.71	farmer	.71	output	.75
lives	.71	generations	.71	scientific	.75
culture	.69	money	.68	climate	.71
ethical	.68	first	.67	heavy	.71
trombay	.67	say	.67	since	.71
appropriate	.67	solving	.67	without	.71
study	.65	view	.67	cost	.69
high	.65	side	.67	things	.69
needed	.65	fact	.64	costs	.67
used	.65	order	.64	individual	.67
surrounding	.63	polluted	.64	keep	.65
industries	.63	population	.64	level	.65
india	.61	residents	.64	related	.65
drilling	.61	rules	.64	social	.65
company	.60	position	.62	hand	.64
though	.60	mathur	.62	human	.63
unethical	.60	face	.62	taken	.61
complex	.60	production	.62	levels	.61
countries	.60	reduce	.62	potential	.61
implement	.60	future	.61	done	.61
living	.60	take	.61	part	.60
affected	.59	example	.60	may	.58
negative	.59	think	.60	society	.58
right	.58	citizens	.59	management	.57

The words most highly weighted by the naïve Bayesian classifier provide a sense of the conceptual differences that were used by the classifier to distinguish USA, Ukrainian, and Indian students. Prior analyses suggested that USA and Indian students were more analytic than Ukrainian students, who tended to be more personal and disclosing. The predictors in Table 5 support these characterizations. USA students were classified by words like *regulations, engineering, engineers, ethical, study, drilling, unethical,* and *complex*, which convey an analytic sense focused on engineering and oil drilling. Indian students were analytic as well, although focused more on the economics of oil drilling, expressed in words like l*oss, sustainable, greed, needs, output, cost, costs,* and *management.* In contrast, Ukrainian students' responses are characterized by words like *authorities, parties, farmer, generations, population, residents,* and *citizens*, and focus on the people who are affected by technology and economics. The Ukrainian students take a more personal and people-oriented approach in their discussion of the ethical dilemma.

Although the naïve Bayesian method carries out analyses of students' responses very differently compared to LIWC-22, the two methods appear to converge on similar characterizations of the three student groups.

Discussion

A major goal of this project was to gain insights into similarities and differences between Indian, Ukrainian, and USA students in their reasoning about a hypothetical ethical dilemma. Analyses applying the LIWC-22 software (Boyd et al., 2022) and naïve Bayesian text analytic methods provided evidence of clear differences between the respective students. Generally, USA and Indian students expressed more formal, logical, and hierarchical thinking, compared to the Ukrainian students; whereas Ukrainian students expressed more informal, personal, here-and-now, and narrative thinking. These general findings were supported by more detailed analyses involving specific words. USA participants were high in LIWC-22 variables *tech, reward*, and *focus past*, which are associated with terms like *automate, engineer, machine, bonus, earn, lucrative, assumed, avoided,* and *tricked*. Indian participants were formal and logical as well, expressed through LIWC-22 variables like *discrepancy, differ,* and *attention*, which are associated with terms like *ideal, ought, yearn, adjust, disparity, impossible, cognizant, negligent,* and *vigilance*. Ukrainian participants expressed more informal and personal ideas consistent with LIWC-22 variables *communication, home,* and *allure*, and associated words like *discuss, inform, question, house, neighborhood, pets, dream, family,* and *friend*. A third analysis involving the Dynamic index (Jordan & Pennebaker, 2017) showed Ukrainian participants to be lower in Analytic and higher in Authentic variables, which provided additional support for the previous analyses and overall consistency in outcomes. A naïve Bayesian method that was applied to the student responses produced outcomes consistent with the LIWC-22 results.

The general approach of this paper is similar to the work of Girnyk et al. (2021), who analyzed differences in the word *conflict* in Ukrainian and Indian university students. Girnyk et al. found clear differences between the groups, with Ukrainian

students associating the term *conflict* with words like *fight, war,* and *aggression*, whereas Indian students associated the term with emotion-laden terms, like *sadness, anger,* and *fear*. In part, Ukrainians may have provided those 'aggressive' response words due to the state of 'hybrid war' with Russia in 2021, now being full-scale since February 2022. The Girnyk et al. study is consistent with the Pennebaker and King (1999) principle that words reveal much about the way people think and feel, and it demonstrates the importance of acknowledging cultural differences in how individuals use words to convey ideas and feelings.

An ongoing goal in our work is to develop machine-based methods to provide students with instructive feedback immediately after they submit a response to the website. Prior work with naïve Bayesian methods suggests the feasibility of feedback tailored to individual student responses. Preliminary work in this area has shown promising results (Taraban et al., 2020).

Conclusions

We began this paper with two provocative suggestions: one, that ethical development and consensus building is a group process, and two, that universities may be susceptible to wide rifts between disciplines, especially Science, the Arts, and Humanities. The development and implementation of global interdisciplinary instructional resources, like the Ethical Engineer website, help in small ways to foster cross-disciplinary interactions and the exchange of ideas across national and cultural divides. The analyses showed that a range of conceptual perspectives coalesced on the website. These findings provide suggestive evidence for the Ethical Engineer as a vehicle for constructive dialogue across national and disciplinary boundaries. Further research including broader sampling and better experimental control is warranted, given the intriguing differences across the samples reported here. Future work should strive to better flesh out similarities and differences due to disciplinary orientation (i.e., science, art, humanities) and ethnic, cultural, and social backgrounds, and thereby deepen our capacity to build interactive platforms that bring students together in constructive and collaborative interactions.

References

Boyd, R. L., Ashokkumar, A., Seraj, S., & Pennebaker, J. W. (2022). The development and psychometric properties of LIWC-22. Austin, TX: University of Texas at Austin. http://dx.doi.org/10.13140/RG.2.2.23890.43205

Girnyk, A., Krylova-Grek, Y., & Khan, A. (2021). A Psycholinguistic cross-cultural study of the concept 'conflict' in India and Ukraine. East European Journal of Psycholinguistics, 8(2), 51-65. https://doi.org/10.29038/eejpl.2021.8.2.gir

Jordan, K. N., & Pennebaker, J. W. (2016). How the candidates are thinking: Analytic versus narrative thinking styles. Retrieved from: https://wordwatchers.wordpress.com/2016/01/21/how-the-candidates-are-thinking-analytic-versus-narrative-thinking-styles/

Jordan, K. N., & Pennebaker, J. W. (2017). The exception or the rule: Using words to assess analytic thinking, Donald Trump, and the American presidency. Translational Issues in Psychological Science, 3(3), 312-316. https://doi.org/10.1037/tps0000125

Muller, J. (2009). Forms of knowledge and curriculum coherence. Journal of Education and Work, 22(3), 205-226. https://doi.org/10.1080/13639080902957905

Pennebaker, J.W., Boyd, R.L., Jordan, K., & Blackburn, K. (2015). The development and psychometric properties of LIWC 2015. Austin, TX: University of Texas at Austin.

Pennebaker, J. W., Chung, C. K., Frazee, J., Lavergne, G. M., & Beaver, D. I. (2014). When small words foretell academic success: The case of college admissions essays. PloS One, 9(12), e115844. https://doi.org/10.1371/journal.pone.0115844

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. Journal of Personality and Social Psychology 77(6), 1296-1312. https://doi.org/10.1037/0022-3514.77.6.1296

Robinson, R. L., Navea, R., & Ickes, W. (2013). Predicting final course performance from students' written self-introductions: A LIWC analysis. Journal of Language and Social Psychology, 32(4), 469-479. https://doi.org/10.1177/0261927X13476869

Taraban, R., Marcy, W. M., LaCour, M. S., Prasad S. H. C., Zasiekin, S. (2020). Using the Web to develop global ethical engineering students. Advances in Engineering Education, 8(3), 1-21. Retrieved from: https://advances.asee.org/using-the-web-to-develop-global-ethical-engineering-students/

Taraban, R., Saraff, S., Iserman, M., Biswal, R., Marcy, W. M. (in press). Developing a global context for ethical reflection. In S. K. Sia, L. S. Crane, A. K. Jain, and S. Bano, (Eds.), Understanding psychology in the context of relationship, community, workplace and culture. Springer Nature Press.