sustainability

MDPI

*Article*

# Understanding Students' Perception of Sustainability: Educational NLP in the Analysis of Free Answers

**Hiroko Yamano [1,*], John Jongho Park [2], Nathan Hyungsok Choe [3] and Ichiro Sakata [1,4]**

1 Institute for Future Initiatives, The University of Tokyo, Tokyo 113-0033, Japan
2 School of Engineering Design, Technology, and Professional Programs, The Pennsylvania State University, University Park, PA 16802, USA
3 Department of Biomedical Engineering, The George Washington University, Washington, DC 20052, USA
4 Graduate School of Engineering, The University of Tokyo, Tokyo 113-8656, Japan
* Correspondence: yamano@ifi.u-tokyo.ac.jp; Tel.: +81-3-5841-8522

**Abstract:** This study explored undergraduate students' conceptions of sustainable development by asking about their definition of a sustainable world, current issues of sustainable development, and the necessary mindset and skillsets to build a sustainable world. We derived data from 107 participants' open-ended answers that we collected through an online survey at the beginning and the end of the sustainability class. Text mining with Natural Language Processing (NLP), principal component analysis (PCA), and co-occurrence network analysis were conducted to understand the changes in students' conception of sustainable development. In addition, we also conducted the Linguistic Inquiry and Word Count (LIWC) dictionary to investigate the psychometric properties of students' awareness and understanding related to sustainable development. This advanced analysis technique provided a rich understanding of university students' perceptions of sustainable development compared to what the UN initially defined as sustainable development goals (SDGs). The results showed imperative insights into the benefits of sustainability experiences and knowledge that generate motivation to develop students' competencies as change agents.

**Keywords:** natural language processing (NLP); sustainability education; sustainable development goals (SDGs); higher education; co-occurrence network; linguistic analysis



**Citation:** Yamano, H.; Park, J.J.; Choe, N.H.; Sakata, I. Understanding Students' Perception of Sustainability: Educational NLP in the Analysis of Free Answers. *Sustainability* **2022**, *14*, 13970. https://doi.org/10.3390/su142113970

Academic Editor: Changwoo Ahn

Received: 1 August 2022
Accepted: 3 October 2022
Published: 27 October 2022



## 1. Introduction

### 1.1. Definitions of Sustainability

The definition of sustainability has been elaborated to explain sustainability related to the human environment, referring to a technique or resource utilization [1]. In contrast, the resource is not exhausted or devastated in perpetuity. This inclusive definition brought about ambiguity, and vast flexibility applied to a broad spectrum of interests from various disciplinary fields. Thus, the interpretation of sustainability can vary depending on the purpose in a given area, however, the universal definition of sustainability applies to all scopes, disciplines, and facets of human effort. For instance, Moore et al. conducted a systematic literature review to conceptualize the definition of sustainability and fundamental concepts including the following categories: (1) after a defined period of time, (2) a program, clinical intervention, implementation strategies that continue to be delivered and (3) individual behavior change (i.e., clinician, patient) is maintained; (4) the program and individual behavior change may evolve or adapt while (5) continuing to produce benefits for individuals/systems [2]. This categorization across definitions reflects that there will be a desired sustained behavior maintained after a specified time an intervention/program is implemented.

Despite these varied definitions of sustainability, the clear focus is to develop solutions for sustainability issues, including growing concerns regarding pollution, social inequities, etc. Although these issues may seem daunting, sustainability goals have been developed to provide solutions to those crucial issues [3]. However, even though there has been a continuous effort to operationalize and conceptualize the concept of sustainability, the word "sustainability" remains ambiguous and multidimensional concept that is inherently difficult to incorporate into education [4].

Johnston et al. made an early attempt to operationalize the definition of sustainability while there were more than three hundred alternative and variously modified definitions of sustainability and sustainable development that exist broadly within the domain of environmental management [5]. If, however, true sustainability in human interactions with the biosphere is to be realized, a far more vital and more empirical interpretation of the original intent is urgently required. To be effective, such an interpretation must encompass and guide developments in political instruments and public policy as well as corporate decision-making and should focus increasingly on addressing the root causes of significant threats to sustainability rather than just their consequences.

Bilen et al. suggested the purpose of sustainability as (1) identifying a stable but inherently unjust equilibrium that causes the exclusion, marginalization, or suffering of a segment of humanity; (2) finding an opportunity in this unjust equilibrium, developing a social value proposition, and bringing to bear inspiration, thereby challenging the stable state's hegemony; and (3) forging a new, stable equilibrium that releases trapped potential or alleviates the suffering of the targeted group, and through imitation and the creation of a stable ecosystem ensuring a better future [6].

### *1.2. Education for Sustainable Development*

The old paradigm of education for sustainable development that began in the 1960s and 70s focused primarily on ecology and the natural world. Since this was not a holistic approach considering other dimensions of care and motivation or resilience, the topics may not have resonated with students as they may not have seen its immediate relevance. More recently, the emerging academic field of sustainability focuses on introducing complex anthropogenic issues including climate change, desertification, poverty, pandemics, and war [7,8]. To address and resolve these sustainability challenges, different academic fields integrate and link existing collective knowledge of problems and solutions to academic settings so that critical sustainability issues will be solved [9,10].

The introduction of education for sustainable development was at the UN Rio Summit in 1992 [11]. During this summit, instructors and curriculum development were identified as significant areas to improve education for sustainable development. After 10 years, the UN General Assembly initiated the Decade of Education for Sustainable Development (DESD) which focused on expediting implementations of the UN sustainable development goals by developing and disseminating education for sustainable development [12].

More recently, following the UN 2030 Agenda for Sustainable Development guideline, the Global Action Programme (GAP) on education for sustainable development was introduced and implemented in 2015. This program explains that sustainability education should be at the center of the plan for sustainable development, and accelerate the transformation of sustainability education in higher education through research and improved teaching and learning environments of sustainability, mainly in developing countries [13].

Further, in 2019, UNESCO adopted the new global framework for education for sustainable development that focuses on integrating education to sustainable development into policies, learning environments, capacity building of educators, empowerment and mobilization, and local level action. The UN General Assembly noted that implementing education for sustainable development will be the critical aspect of accomplishing the UN sustainable development goals [14].

Therefore, to address and resolve sustainability challenges, different academic fields integrate and link existing collective knowledge of problems and solutions to academic settings so that critical sustainability issues, many scholars and educators began to investigate and develop education for sustainable development with pedagogical techniques that promote system thinking, service learning, participatory learning, and project-based education [8–10].

Considering complex sustainability issues, Sterling et al. introduced a holistic approach to sustainability education, seeing it as a process of maintaining a healthy system that draws on qualities such as creativity, self-reliance, self-realization, wholeness, and resilience [15]. Based on Sterling's holistic approach, Fisher et al. introduced three imperative dimensions including cultivating an ethic of care, fostering resilience and regeneration, and advocating for life's flourishing [16]. These concepts have been widely introduced for comprehensive sustainability education. They also emphasized the concept of cultivating an ethic of care that expresses the concept of "Care for the community of life with understanding, compassion, and love." Fisher and colleagues claimed that this aspect is immensely imperative for education for sustainable development as if there is no genuine care for self, others, nature and the earth, then there will not be any intrinsic motivation to maintain them.

### *1.3. Instructional Strategies for Sustainable Development Education*

As for the instructional strategies of education for sustainable development, Remington and colleagues introduced the use of real-world problems in sustainability education [17]. This provides a vital context and relevance as well as allows students to solve real world problems and engage in a project in teams. In addition, project-based assignments that support students to work in teams to solve problems have become more common in recent times and have become more common in higher education [18,19]. Rulifson and Bielefeldt explored students' conceptualizations of sustainability by participating in sustainability projects [20]. They found that varied project experiences including social-learning shaped students' views of sustainability. These projects converged to develop students' understanding of socially responsible works to include safety, ethics, and various sustainability issues.

Smith et al. claimed that project-based learning empowers socio-technical thinking as an imperative element for engineers who will address sustainability issues [21]. They proposed that these sustainability-related projects should include integration of design thinking and interdisciplinary problem-solving into individuals' knowledge of global, societal, economic, and environmental contexts. Amadei and Wallace also mentioned that it is critical for academic programs to train students to develop their socio-technical perspectives so that they will be able to address global sustainability issues [22].

## **2. Research Rationale**

The concept of sustainable development was first introduced by the United Nations in 1992 and expedited by various nations and organizations to implement the UN sustainable development goals. Ever since, there have been continuous efforts to operationalize the concept of sustainability to develop the current education curriculum that incorporates existing knowledge of sustainable development and current issues. In order to properly integrate sustainable development goals into current education, it is necessary to understand students' conceptions of sustainability and current issues of sustainable development.

While limited in their ability to compass all variables and factors, the use of conceptual frameworks is vital in the training and development process for sustainability education. . According to previous operational definitions and categorizations of sustainability, Wiek et al. identified critical competencies for sustainability change agents including systems-thinking, anticipatory, normative, strategic, and interpersonal competencies to develop a framework for sustainability-related academic programs [23]. The authors proposed that education for sustainability should be based on those competencies. They highlight the concepts of sustainability should be different by period, characteristics of stakeholders, temporal phases, uncertainty and epistemic status, inertia and path dependency, consistency and plausibility of future developments, and risk and intergenerational equity. They addressed that higher education sustainability courses should be designed based on key concepts and encourage students to acquire competencies. Regardless of a continuous effort to conceptualize the concept of sustainability, the meaning remains complex and multidimensional that is inherently difficult to comprehend [4,24]. Aligned with Withycombe et al.'s comment on operationalizing the concept of sustainability, to properly integrate the sustainable development goals into current education, it is necessary to understand students' conceptions of sustainability and current issues of sustainable development.

### *2.1. Target of Research*

This study was conducted in a sustainability education class at a research-intensive university in Australia from February to November 2021. This class is designed to introduce students to the essential professional skills and mindsets for building a sustainable world. In this course, students learn how to lead and sustain themselves, and their relationships with others in our volatile, uncertain, complex, ambiguous (VUCA) world. The course has an emphasis on self-development and leadership based on a highly sophisticated framework extending awareness and understanding through empathy and systems thinking that leads to practical problem solving and successful project work. The focus areas for projects are based on the United Nations Sustainable Development Goals (SDGs). The course consists of weekly lectures covering essential topics on leadership and sustainability, tutorials where students engage in various activities supported by tutors and mentors, and mini courses led by domain experts to reflect on the points that can be applied to the team projects. The characteristic of this course is its focus on combining technical and non-technical skills for long-term effectiveness designed for high-achieving scholarship students. The course focuses on individual critical and creative thinking, reflection and contemplation, communication and collective creative problem-solving in a team, from a deep sense of awareness and understanding. This course seeks a shift in mindset, a transformation to see sustainable development problems differently, to relate to oneself, others, and the world in a new way of considering the root causes of issues with a holistic whole system perspective [25].

### *2.2. Quantitative Analysis of Students' Free Answers*

Course surveys are essential tools for evaluating and updating the class contents. The accessible answers of students, in particular, often contain valuable information and provide an essential basis for rebuilding current education. They are helpful in evaluating classes targeting diverse and multi-domain issues as seen in sustainability education. Previous researches that analyzed free answers to investigating the changes in students' perceptions and attitudes through the course have been primarily qualitative [26–28]. However, the summaries and extractions that human analysts make of these open-ended answers tend to rely on intuition and are not satisfactorily reliable [29]. Given the recent increase of massive open online courses during the pandemic, it has become easier to take surveys and receive feedback on a large scale, which requires precision and agility for the evaluation [30]. Therefore, applying computer-based quantitative analysis that can automatically mine useful information from open answers has become an important issue. In this study, we applied an integrated text mining approach with Natural Language Processing (NLP), principal component analysis (PCA), and co-occurrence network analysis to understand the changes in students' conception of sustainability from their free answers. We also used Linguistic Inquiry and Word Count (LIWC) dictionary to investigate the psychometric properties of students' awareness and understanding related to sustainability.

### *2.3. Research Purpose*

The purpose of this study is to verify whether the course would achieve expected effect as a sustainability education. Based on the characteristics of the course that focuses on the transformation in students' conception, this study sought to analyze the changes in the basic psychometric properties of students' awareness and understanding related to the UN sustainable development goals. On two occasions before and after the class, participants were asked to answer their standpoints toward sustainability. Of the 32 questions executed in the survey, we focused on the three descriptive questions and their open-ended answers as follows, which form the core of the survey.

Q1. What is your definition of a sustainable world?
Q2. What kind of sustainability issues are critical and should be addressed in creating a sustainable world? Please explain your standpoint based on your experience.
Q3. What are the essential mindset and skills required to build a sustainable world? Please explain why you would need such a mindset and skills.

While there were limitations to surveying all questions and responses, the use of free answers is vital for examining the effectiveness of the sustainability course that addresses complex issues as we discussed above. Three broad research questions around relative differences in students' understanding and language use in their answers guided the analyses: (a) How different is students' answer distribution before and after the class? (b) How distinctive is the word structure in students' answers? (c) How different is students' language use across questions before and after the class? Based on these research interests, we formulated the following hypotheses regarding the main trends in the changes in students' responses to each question after the class.

H1. Students obtained diverse aspects regarding the definition of sustainability. (Q1)
H2. Students shared consolidated views on the critical issues of sustainability. (Q2)
H3. Student opinions of the essential mindset and skills for a sustainable world remained diverse without common understanding. (Q3)

To shed light on the results, we used text mining techniques to investigate undergraduate students' conceptions of sustainability and necessary mindsets and skills. Text mining is the process of exploring and analyzing text data that includes various methods such as NLP, PCA, and co-occurrence network analysis, explained in the next section. This advanced analysis technique provided a rich understanding of university students' perspectives on sustainability compared to the UN originally defined sustainable development goals.

## 3. Materials and Methods

### *3.1. Participants*

Participants were scholarship students enrolled in a sustainability class at a research-intensive university in Australia who scored in the top 2% of Australian Tertiary Admission Rank scores. Students come from various disciplines including Engineering, Business, Science and Arts. A total of 107 out of 122 students completed the survey comprising 67 males (54.9%) and 55 females (45.1%). Most were in their first year (81, 66.4%), with 24 second-year (19.7%), 16 third-year (13.1%), and 1 fourth-year student. There were 33 Americans (27%), 23 Sub-Saharan Africans (18.9 %), 21 Asians (17.2%), 13 Oceanians (10.7%), 12 North African and Middle Eastern (9.8%). 14 Europeans (9), and others who participated in this study. Half of the students majored in Engineering (64, 52.4 %), 20 and 17 of them majored in Arts and Social Science (16.4%) and Architecture, Design & Planning (13.9%) as follows. When they were asked about their career intention in STEM, 84 students (68.9%) answered that they planned to pursue a STEM-related career. Data sources included responses to open-ended questions delivered via a Qualtrics survey about building a sustainable world. The student characteristics included demographic and academic information such as gender, ethnicity, major, and program year. Participation was voluntary, and students could decide to stop participating in the survey at anytime. Students did not have to answer any questions they did not want to answer. Their participation did not have any impact on their grades or engagement in the course. Their confidentiality was maintained to the degree permitted by the technology used and we did not capture identifiable information.

### 3.2. Data Analysis Methods

In this study, we conducted a comparison of students' free answers in before and after surveys using the techniques of NLP to prepare the free answers for analysis, as part of the pre-processing stage. NLP has been used for the text analysis of free-style answers, such as feature extraction [31,32], text classification [33], and sentiment analysis [34]. NLP is a set of methods to enable computers to understand the meaning of natural language, such as documents and sentences used by humans. It includes morphological analysis, syntactic analysis, semantic analysis, and contextual analysis. Morphological analysis is a method that divides sentences into morphemes, the minor linguistic units having meaning. It assigns information to morphemes such as parts of speech (POS tagging) using a machine-readable dictionary and a corpus of documents that stores the use of a language. In the pre-processing of documents, the stopwords unrelated to the topic are often removed beforehand. Stemming is a method that converts derivatives to the same feature. As many stemming methods use simple lexical processing, the converted result can have a different meaning. Therefore, the process of restoring words in their basic form is often used instead of stemming, which is called lemmatization. After the pre-processing, vector representation is often used to represent natural language in a mathematically tractable form that a computer can understand. Each element of a vector represents the features of a document or sentence. Each dimension of the vector is associated with a word, and its value is usually the frequency of words in the document. This kind of vector representation method of documents is called bag-of-words, as information about sentence structure and word order is lost. However, in many cases, the frequency of the word-based approach is sufficient to classify sentences according to their topics. In this study, since the scale of the data is relatively small, we used morphological analysis and vector representation based on the frequency of words.

To analyze the open-ended responses to the three questions, we executed basic pre-processing using Natural Language Toolkit (NLTK) [35] as the following: (1) Replace blank answers with empty space, (2) Lowercase all text, (3) Remove all blocks of digits, (4) Remove all punctuation and replace it with a space, (5) Remove all diacritics and accents, (6) Remove ubiquitous words using NLTK's English stopwords of 179 words, (7) Remove any extra whitespace, newline, tabs and any form of space. Next, to remove affixes and map various forms of a word to the canonical word in a dictionary (known as lemmatization), and to obtain the grammatical parts of speech tagging, we used a language model of en_core_web_sm provided by spaCy [36], the open-source library for advanced Natural Language Processing. en_core_web_sm is a trained pipeline for English that incorporates lexical databases from OntoNotes 5 [37] and WordNet 3.0 [38]. In the following analysis, we only used pre-processed words that are tagged as adjectives, verbs, nouns, proper nouns, or adverbs and excluded words tagged otherwise from the analysis data.

#### 3.2.1. Principal Component Analysis (PCA)

To extract principal information of the answers to each question, we performed PCA [39] and students' answers were transformed into vector representations with the following procedures. We used each pre-processed word in a response as a "term", and each response in the target question as a "document". We calculated the values of terms in each document using TF–IDF [40], which is the product of the term frequency (TF) and the inverse document frequency (IDF). This scoring method is empirically useful for a wide range of datasets and has been proven to represent "the amount of information of a term weighted by its occurrence probability" [41]. TF–IDF is defined as the following equation,

$$\mathrm{tfidf}(t,d) = \mathrm{tf}(t,d) \cdot \mathrm{idf}(t), \quad (1)$$
$$\mathrm{tf}(t,d) = \log(1 + f_{t,d}), \quad (2)$$
$$\mathrm{idf}(t) = \log \frac{N}{df_t}, \quad (3)$$

where $t$ and $d$ represent the term and the document, respectively, $\mathrm{tf}(t,d)$ is the log value of the TF, and $\mathrm{idf}(t)$ is the inverse ratio of the number of documents that include the term $t$. $N$ represents the total number of documents. When calculating TF–IDF we ignored terms that have a document frequency lower than the threshold of 1.

By concatenating the values of terms in each response to each question, we obtained the vector representation of each question, which is the multi-dimensional array that corresponds to the TF–IDF values of the terms in the documents. To investigate the difference in the vector representation of the questions, we applied PCA. PCA is the dimensionality-reduction method that transforms variables into lower-dimensional data while preserving as much of the data's variation as possible, by decomposing data into two orthogonal matrices with maximum variance and singular values known as an eigenvalue. In this study, we reduced data into two principal components to visualize the vector representation of the questions into X,Y two-dimensional coordinates.

#### 3.2.2. Co-Occurrence Network Analysis

A document is a set of words, and the meaning it contains can be regarded as being generated by the co-occurrence of words. Since the co-occurrence of words differs according to the documents, the semantic aspects of the topics they represent also change. In other words, topic context can be modeled using the co-occurrence relationships between the words in the topic [42]. A co-occurrence network is one of the primary methods of text mining, which visualizes such relationships between words in a document as a network, based on the similarity of word occurrence patterns in the sentences. The nodes in the network represent words, and the width of edges between nodes represents the strength of word relationships, i.e., the number of their co-occurrences. Many researchers have used a co-occurrence network of free answers, such as co-occuring emotions related to students dropping out during the Massive Open Online Course (MOOC) [43], self-reported labels of gender and sexual identities [44]. Co-word maps based on word occurrences are more effective than topic modeling for designating semantic components, especially for small and medium-sized data sets [45].

To capture the difference in the students' responses before and after the class, we counted the co-occurrences of words in every response, and the number of occurrences of all combinations of the pre-processed word pair except for pairs of the same word in response. Then we summed up the number of co-occurrences of pairs in the answers to each question, which corresponds to the weight of the edges in the co-occurrence network. For visualization, we selected the top 200 co-occurrence pairs and maximum connected components in the network to avoid complexity. The network was drawn using the Fruchterman-Reingold algorithm that simulates a force-directed representation of the network treating edges as springs holding nodes close while treating nodes as repelling objects. Simulation continues until the positions are close to an equilibrium. In this algorithm, nodes connected by edges are placed close together (attractive force), while nodes not connected by edges are placed farther apart (repulsive force). Therefore, in the case of a co-occurrence network, words with a high co-occurrence, which are the words students frequently use together in their responses, are placed closer and vice versa.

#### 3.2.3. Linguistic Analysis with LIWC Dictionary

To investigate the psychometric properties of students' understanding related to sustainability, we investigated the linguistic features of their free answers, which employed the LIWC2015 program (LIWC–Linguistic Inquiry and Word Count). This application was created to serve as an efficient and effective method for studying the various emotional, cognitive, and structural components present in individual's verbal samples. It has been used to analyse students' self-descriptions associated with their properties such as depression and depression-vulnerability [46], course performance [47], and personality differences [48]. The LIWC was firstly developed as part of an exploratory study of language and disclosure [49] and has been constantly updated over time [50–52]. LIWC2015 software relies on an internal default dictionary defining which words should be counted in the target text. In this study, LIWC2015 English dictionary data was provided by the authors under the particular license they offered for more efficient analysis. LIWC2015 dictionary is composed of almost 6400 English words and 73 categories including linguistic dimensions, psychological constructs, and personal concerns.

Among the psychological categories in LIWC2015, we focused on four categories of Affective, Cognitive, Drives, and Relativity that showed large differences in the number of word occurrences before and after class. The four categories contain sub-categories as follows: Affective (Positive emotion, Negative emotion, Anxiety, Anger, Sadness), Cognitive (Insight, Causation, Discrepancy, Tentative, Certainty, Differentiation), Drives (Affiliation, Achievement, Power, Reward, Risk), and Relativity (Motion, Space, Time). Therefore, we used 4 categories and 19 sub-categories in total, to extract the changes of students' perceptions before and after the sustainability class.

## 4. Results

Data summary for students' responses is provided in Table 1. Since each question was not mandatory, the number of valid answers differed between questions before and after the sustainability class. The average word per answer indicates the average number of pre-processed words (adjectives, verbs, nouns, proper nouns, or adverbs) in each questionnaire. Unique pairs are the number of combinations of pre-processed words in the questionnaires' responses, which were made for the co-occurrence network. Normalized unique pairs were calculated by dividing individual pairs by valid answers and average words per answer. There were fewer valid responses and fewer pre-processed words and unique pairs in the responses after the class. Q2 had the highest valid responses, average number of words, and unique pairs among the three questions. Individual pairs in Q2 and the rate of their decrease after class was higher than those in Q1 and Q3. However, the values of the normalized unique pairs were not significantly different between questions.

**Table 1.** Data Summary.

| | Valid Answers | | Average Words per Answer | | Unique Pairs | | Normalized Unique Pairs | |
|---|---|---|---|---|---|---|---|---|
| **Question** | **Before** | **After** | **Before** | **After** | **Before** | **After** | **Before** | **After** |
| Q1 | 167 | 115 | 13.98 | 13.78 | 15,864 | 11,385 | 6.795 | 7.184 |
| Q2 | 178 | 133 | 29.01 | 20.45 | 83,399 | 36,664 | 16.151 | 13.480 |
| Q3 | 169 | 124 | 23.20 | 19.31 | 43,435 | 26,142 | 11.078 | 10.918 |
| Average | 171 | 124 | 22.06 | 17.85 | 47,566 | 24,730 | 12.585 | 11.173 |

Figure 1 plotting points are calculated by PCA analysis for each response to questions Q1, Q2, and Q3, respectively. The blue dots show the distribution of participants' responses at the beginning of the semester. The orange dots provide the distribution of participants' answers at the end of the semester. In Q3, there is not much significant change in the structure of the latent space between the blue and orange dots, indicating that participants tended to provide dispersed answers without any common concepts. However, in Q1 and Q2, the distributions of the dots are particularly clearly differentiated, indicating that their answers tended to be semantically different after the semester. Regarding the content of the questions, the answers to Q2 about "most critical sustainability issues" tended to converge. In contrast, the answers to Q1 about the "definition of sustainability" grew to become diverse after the class. These results show that PCA analysis seemed useful for distinguishing the significant changes in the respondents' answers before and after the course, especially concerning the urgency and definition of sustainability. However, it may not help detect minor or diversified changes in mindsets and skills needed for a sustainable world.

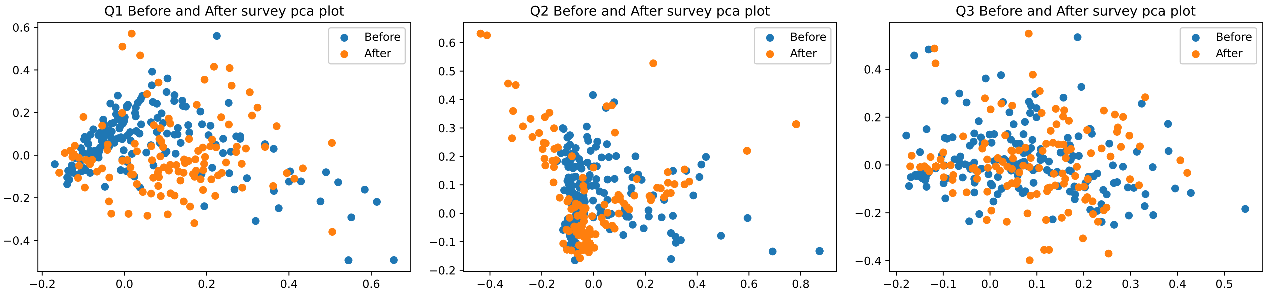


**Figure 1.** (**Q1.** ) The distribution of responses on "What is your definition of a sustainable world? Please explain your stand point based on your experience." in the latent semantic space, both for semester beginning (blue) and semester-end (orange). (**Q2.**) The distribution of responses on "What kind of sustainability issues are most critical and should be addressed in creating a sustainable world? Please explain your stand point based on your experience." in the latent semantic space, both for semester beginning (blue) and semester-end (orange). (**Q3.**) The distribution of responses on "What are the essential mindset and skills that you would need to build a sustainable world? Please explain why you would need such a mindset and skills. ? Please explain your stand point based on your experience." in the latent semantic space, both for semester beginning (blue) and semester end (orange).

Figure 2 represents the co-occurrence network of the words in the responses to each question. The structure of the co-occurrence networks differed between the six networks of the questions surveyed before and after the class. Q1 network had one center, i.e., almost every word was connected to the centered two words of "sustainable" and "world". On the other hand, Q2 and Q3 networks had broader centers and the centered words were dispersed after the class. Especially in Q2, the center was divided into two centers with the appearance of the new hub of the word "poverty".

Of the words used in the responses, the coverage rate of the terms registered in the LIWC2015 dictionary was about 60%. Table 2 shows the number of pre-processed words in each question, registered words, and registration rate in the LIWC2015 dictionary, both in all categories (73 categories) and psychological categories (46 categories).

**Table 2.** A number of words and their registrations in LIWC2015 dictionary.

| | | 73 Categories | | 46 Categories | |
|---|---|---|---|---|---|
| **Group** | **Words** | **Registered Words** | **Registration Rate** | **Registered Words** | **Registration Rate** |
| Q1 before | 2335 | 1384 | 0.593 | 1311 | 0.561 |
| Q2 before | 5164 | 2936 | 0.569 | 2728 | 0.528 |
| Q3 before | 3921 | 2366 | 0.603 | 2252 | 0.574 |
| Q1 after | 1585 | 1009 | 0.637 | 969 | 0.611 |
| Q2 after | 2720 | 1590 | 0.585 | 1482 | 0.545 |
| Q3 after | 2394 | 1511 | 0.631 | 1431 | 0.598 |

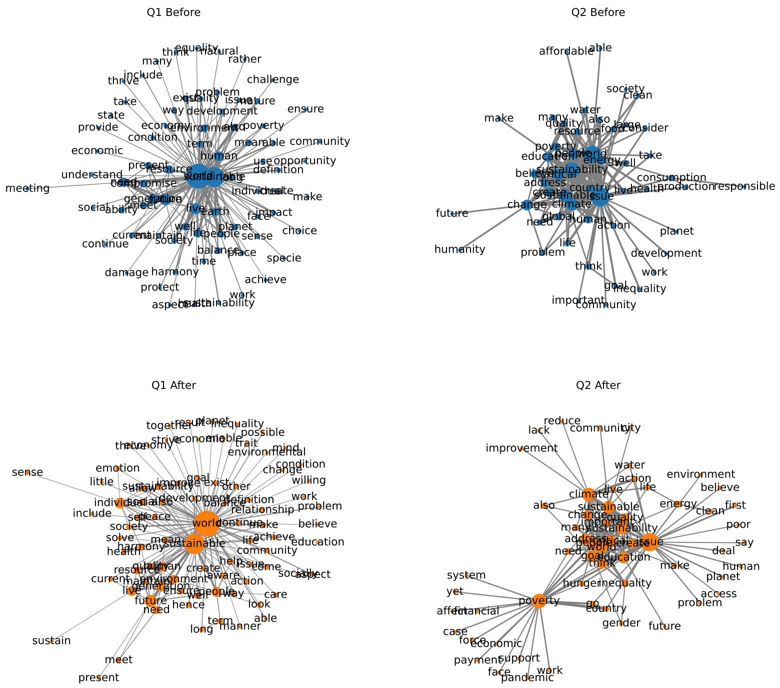


**Figure 2.** The co-occurrence network of the words in the responses to each question.

Figure 3 represents the Pearson correlation among the questions calculated by the distribution of word occurrences in psychological categories. We used 46 psychological categories, including sub-categories, in the LIWC2015 dictionary and compared the correlation of word distribution among the categories before and after the class. In total, highly correlated pairs were the group of the answers to the same questions before and after class, e.g., Q1_before and Q1_after.

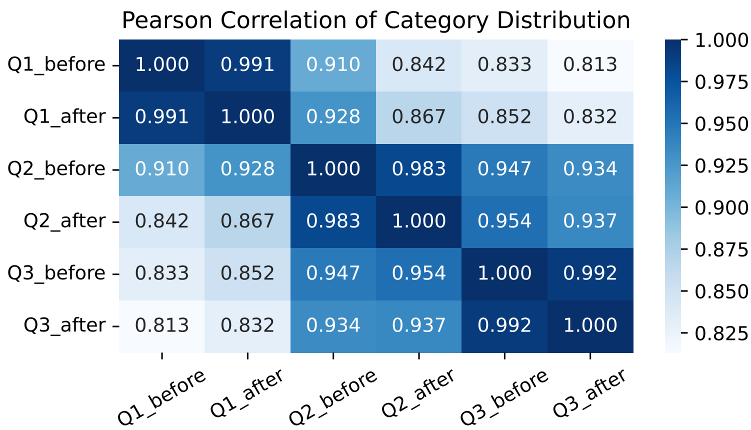


**Figure 3.** Pearson correlation of psychological category distribution.

The number of occurrences of the word in the targeted four psychological categories (Drives, Relativity, Affective, Cognitive), including 17 sub-categories in each question, is shown in Table 3. In LIWC2015, each word was assigned to approximately three categories on average, so the numbers were higher than the actual number of word occurrences. The values ranged from about 300 to 1650 with no significant bias in the distribution across categories.

**Table 3.** A number of word occurrences in psychological categories.

| Group | Drive | Relativity | Affective | Cognitive |
|---|---|---|---|---|
| Q1 before | 508 | 1012 | 331 | 504 |
| Q1 after | 426 | 731 | 303 | 358 |
| Q2 before | 1475 | 1639 | 982 | 1445 |
| Q2 after | 964 | 760 | 542 | 823 |
| Q3 before | 1446 | 1212 | 767 | 1777 |
| Q3 after | 872 | 739 | 523 | 1214 |

To compare the difference among categories, we calculated standard scores for each question before and after the class as the rate of the word occurrences in each of four categories with sub-categories, divided by their total number of occurrences in 43 psychological categories. Table 4 shows the representative word examples and their number of occurrences in the categories with significant differences in the standard scores before and after the class. The differences were calculated by the standard score before class minus those after class. If the difference was positive, before class word samples were displayed; if negative, after class samples were shown in Table 4. Among the categories, main categories were indicated with the asterisk "*". Achievement and Power were the sub-categories of Drive, and Work was the sub-category of Personal concerns. Therefore, we focused on the four categories of Drives, Relativity, Affective and Cognitive processes, and their sub-categories with high differences.

**Table 4.** Representative words in categories with significant difference of occurrences.

| | Category | Difference | Examples |
|---|---|---|---|
| Q1 | Affective * | −0.434 | (well, 15), (peace, 9), (create, 7), (problem, 6), (support, 5) |
| Q1 | Present focus | 0.305 | (live, 58), (need, 48), (meet, 21), (present, 15), (take, 10), |
| Q1 | Discrepancy | 0.296 | (need, 48), (would, 12), (problem, 5), (must, 4), (rather, 3) |
| Q1 | Social * | −0.277 | (people, 40), (human, 19), (individual, 15), (help, 9), (social, 9) |
| Q1 | Drives * | −0.277 | (able, 13), (help, 9), (social, 9), (allow, 9), (meet, 8) |
| Q2 | Relativity * | **0.814** | (world, 121), (change, 71), (global, 31), (action, 27), (environment, 22) |
| Q2 | Drives * | **−0.613** | (education, 61), (goal, 37), (important, 32), (poverty, 29), (create, 24) |
| Q2 | Power | −0.432 | (education, 61), (important, 32), (poverty, 29), (help, 7), (big, 6) |
| Q2 | Work | −0.424 | (education, 61), (goal, 37), (work, 12), (achieve, 11), (resource, 10) |
| Q2 | Time | 0.382 | (time, 17), (due, 12), (term, 11), (current, 11), (still, 10) |
| Q3 | Achievement | 0.433 | (skill, 90), (work, 28), (solution, 27), (achieve, 21), (able, 20) |
| Q3 | Work | 0.354 | (skill, 90), (work, 28), (achieve, 21), (learn, 14), (goal, 14) |
| Q3 | Cognitive * | −0.311 | (need, 56), (think, 31), (problem, 29), (change, 25), (other, 24) |
| Q3 | Insight | −0.292 | (think, 31), (idea, 19), (solution, 18), (understand, 15), (solve, 14) |
| Q3 | Drives * | 0.240 | (skill, 90), (important, 39), (work, 28), (problem, 28), (solution, 27) |

Figure 4 shows a scatter plot of the standard scores with the (X and Y) axes corresponding to (Drives, Relativity) with blue dots and (Affective, Cognitive) with orange dots, respectively. The red arrows indicate the most significant change in the after-class values compared to the before-class values, which corresponds to (Drives, Relativity) in Q2. On the contrary, the position of (Affective, Cognitive) in Q2 was almost the same before and after the class.

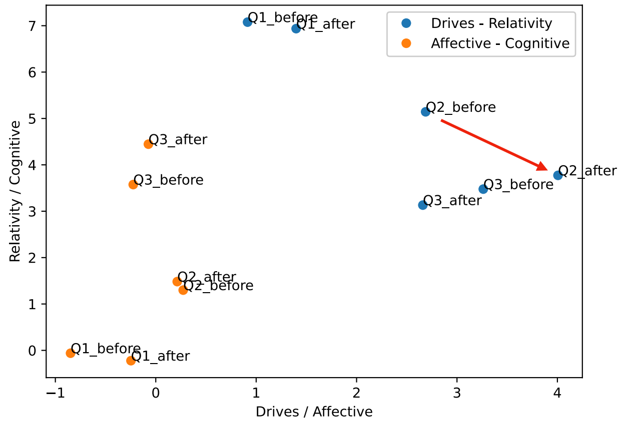


**Figure 4.** Changes in the psychology categories in the responses to each question before and after the class.

## 5. Discussion

This study aimed to determine the basic psychometric properties of students' understanding of sustainability as measured by their language use in their answers. This paper does not cover the full range of sustainability education. Instead, it focuses on the self-development and shift in mindset needed to address real-world sustainability concerns described in Section 2.1. Therefore, the primary purpose of this study was to evaluate the transformation in students' perception rather than memorizing knowledge about sustainability.

Due to the limited number of data and the variety of students' answers, we first focused on extracting major components of the data and their changes after the class. Then we confirmed the contents of the changes using network analysis and dictionary matching. Using network analysis in the text mining of free answers, we detected the unique perspective of the changes in students' responses after the course. Results show that students' understanding of critical issues has increased significantly, strengthening their motivation to build a sustainable world. According to the results, sustainability education is more than simply transferring knowledge to students. It is also driven by encouragement to acquire competencies for sustainability change agents, such as systems thinking, anticipatory, normative, strategic, and interpersonal competencies to cope with real-world problems [4].

### *5.1. Students' Understanding of Sustainability from Word Distribution*

#### 5.1.1. PCA Plotting

One of the goals of this study was to estimate fundamental changes in students' awareness and understanding of sustainability from the perspective of word distribution. Using TF-IDF values of words, we obtained the distributed representation in students' free answers and compared their differences using PCA plotting. Although PCA plotting is a naive method to visualize data points by summarizing word representation into two dimensions, it is essential to note that we found different trends among the answers. Since PCA extracts principal components with the most considerable variance in the data set, it reflects the main differences in the students' use of words. The distribution of the data points showed that students tended to give semantically different answers for Q1 and Q2 after class, with Q1 more dispersed and Q2 more aggregated. The Q1 is about the definition of a sustainable world; these dispersed Q1 PCA plots can be explained from class contents. This class provided UN 17 sustainable development goals to students. Thus, the students may expand their thoughts about definitions of sustainability related to the 17 goals to make a sustainable world. In addition, they gave continuously diverse answers over time in Q3, suggesting that students with different majors may provide different essential mindsets and skills to build a sustainable world. The results suggested that hypotheses H1, H2, and H3 were supported.

#### 5.1.2. Co-Occurrence Network Analysis

Before assessing the hypotheses, we first observed the consistency between network structure and word use in students' answers. We applied co-occurrence network analysis by linking frequently occurring words in response, the top 200 frequent word pairs in this study. The typical star structure of the co-occurrence network of Q1 was consistent with the grammatical characteristic of English that requires a subject to complete a sentence. Since Q1 required the same subject to explain its definition, almost all answers contained "sustainable world" as a subject, resulting in a hub of two words in the network. On the other hand, Q2, and Q3 allowed for a wider variety of subjects and words in their responses, consistent with the network structure with a broader center. This network characteristic was aligned with the more prominent average words and unique pairs of Q2 and Q3, as shown in Table 1. From these observations, we assumed co-occurrence network analysis is functional for assessing the basic features of the answers formulated in the hypotheses. The emergence of the new hub "poverty" and its neighbor words' low number of links in Q2 after the class showed the uniqueness of the neighbors, suggesting that quite a few students tended to use more specific terms related to poverty after the course. These findings support our hypothesis H2, indicating that a substantial number of students shared the perception of poverty as a critical issue of sustainability. This result aligns with a few empirical studies on college students' perceptions of sustainability and poverty. For example, Watson et al. reported that engineering students participating in learning-cycle-based sustainability class were able to connect relations between sustainability and poverty well [53]. However, we could not find particular patterns that strongly explain hypotheses H1 and H3 from the analysis based on the overall structure of the network. For a more detailed assessment, we investigated the word distribution using linguistic analysis based on psychological categories as the following.

### *5.2. Psychological Transfer in Students' Understanding of Sustainability*

#### 5.2.1. Assessing the Stability

In studies analyzing time-series changes in spoken words, the stability of the term used was often a key issue concerning the reliability of the analysis. Stability has been investigated by the correlation or overlap of people's language use over time [54,55]. Based on the preceding studies related to stability, we investigated the Pearson correlation of the students' word use before and after the class to assess the reliability of the analysis (Figure 2). As a result, the groups of answers to the same question were highly correlated compared to the other questions, indicating stability in the tendency of students to give similar answers to the same questions. The stability among questions was also confirmed by the scatter plot of the standard scores in Figure 4, i.e., the positions of each question were relatively close before and after class compared to those of the other questions.

#### 5.2.2. Assessing the Change

Among the relatively unchanged positions in the same questions, Q2, given Drive and Relativity, showed the most considerable shift before and after the class, contrary to Q2, given Affective and Cognitive process that showed the slightest shift (Figure 4). The LIWC2015 dictionary has five sub-categories within the Drive category: Affinity, Achievement, Power, Reward, and Risk, which represent the motivation factors that lead to behaviors. The new hub of "poverty" on the co-occurrence network in Figure 3 belongs to the sub-category of Power. Therefore, the hub "poverty" appears to emerge, accompanied by an increase in Drive-related words. Considering that the content of Q2 was "What are the essential issues for creating a sustainable world?" students' perceptions seemed to have deepened to produce more concrete actions rather than internal aspects such as emotion and cognition.

## 6. Conclusions

This study investigated students' perceptions of sustainability and the skills required to build a sustainable world from an educational perspective. The results showed imperative insights into the benefits of sustainability experiences and knowledge that generate motivation to develop students' competencies as change agents. In addition, this study's surveys were conducted at the beginning and the end of the sustainability class to explore changes in students' perceptions of building a sustainable world. The results indicated that the sustainability class impacts college students' perception of critical sustainability issues and these solutions. The word distribution PCA plots showed that class might assist students with understanding or identifying the core critical sustainability issues, given that the plots converged. Furthermore, a co-occurrence network analysis showed that students view poverty as an essential issue in their post-survey than in their pre-survey. In representative words in categories (Table 4), students used more drive words in the post-survey compared to the pre-survey and fewer relativity words in the post-survey compared to the pre-survey. Perhaps students have responded by providing solutions to build a sustainable world rather than raising sustainability issues. Lastly, the central phenomenon portrayed in Figure 2 shows that emerging poverty as a critical issue for sustainability tended to depend on motivations, as discussed above. These motivations were not limited to acquiring knowledge about sustainability. Thus, there were ways that the various interests of sustainability and envisioning a possible future self were connected to the sub-themes of sustainability. Based on our study, we recommend sustainability educators and practitioners include real-world-based projects that incorporate the United Nations' sustainable development goals and provide activities that address the fundamental issues with a holistic whole system perspective. This practice may help college students think comprehensively about sustainability issues and how to develop a sustainable world.

There are several caveats to our analysis. First, because the survey data used in this study were from the sustainability course designed for personal and leadership development from a holistic perspective and mutual understanding, it did not pursue domain-specific knowledge or techniques, as discussed in Section 5. Therefore, programs aimed at acquiring specific techniques that contribute to sustainability may not be applicable or produce different results for this study. Second, the visualizations of Figures 1 and 2 does not cover the entire data. Due to the limited number and variety of data in this study, it required more than 50 dimensions to meet the sufficient cumulative contribution rate above 0.6 for PCA validation. Therefore, the PCA plotting of 2-dimensional space does not represent the entire data distribution, and we used the visualization only to capture the major difference in the data. The characteristics of the overall structure of the network are also difficult to verify. Therefore, we visualized the network multiple times and confirmed its major characteristics did not change with the fluctuation caused by the visualization. PCA and co-occurrence networks only capture the principal differences among students' answers; however, the LIWC dictionary covers the majority of data, as shown in Table 2. Third, because we focused on the primary analysis based on word occurrences, the context of each student's answer was lost. In addition, because the data were not linked to individual names or IDs, the analysis result did not capture the individual student differences throughout the course. Applications of advanced NLP techniques, such as deriving domain models from free answers, are left for future research.

**Author Contributions:** Conceptualization, H.Y. and J.J.P.; methodology, H.Y.; software, H.Y.; formal analysis, H.Y.; investigation, J.J.P. and N.H.C.; resources, H.Y. and J.J.P.; data curation, H.Y.; writing—original draft preparation, H.Y. and J.J.P.; writing—review and editing, H.Y., J.J.P., N.H.C. and I.S.; visualization, H.Y.; project administration, N.H.C.; funding acquisition, H.Y. and I.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank Ehssan Sakhaee and Melissa Neighbour from the University of Sydney for data collection, and helpful discussion and revision of the manuscript.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Berkes, F.; Folke, C. Linking social and ecological systems for resilience and sustainability. *Link. Soc. Ecol. Syst. Manag. Pract. Soc. Mech. Build. Resil.* **1998**, *1*, 4.
2. Moore, J.E.; Mascarenhas, A.; Bain, J.; Straus, S.E. Developing a comprehensive definition of sustainability. *Implement. Sci.* **2017**, *12*, 110. [CrossRef]
3. Degirmenci, K.; Recker, J. Creating environmental sensemaking through green IS: An experimental study on eco-nudging paper printing behavior. In Proceedings of the 24th Americas Conference on Information Systems, Association for Information Systems (AIS), New Orleans, LA, USA, 16–18 August 2018; pp. 1–10.
4. Ngo, T.T.; Chase, B. Students' attitude toward sustainability and humanitarian engineering education using project-based and international field learning pedagogies. *Int. J. Sustain. High. Educ.* **2020**, *22*, 254–273. [CrossRef]
5. Johnston, P.; Everard, M.; Santillo, D.; Robèrt, K.H. Reclaiming the definition of sustainability. *Environ. Sci. Pollut. Res. Int.* **2007**, *14*, 60–66. [CrossRef] [PubMed]
6. Bilén, S.G.; Kisenwether, E.C.; Rzasa, S.E.; Wise, J.C. Developing and assessing students' entrepreneurial skills and mindset. *J. Eng. Educ.* **2005**, *94*, 233–243. [CrossRef]
7. Bäckstrand, K. Civic Science for Sustainability: Reframing the Role of Experts, Policy-Makers and Citizens in Environmental Governance. *Glob. Environ. Politics* **2003**, *3*, 24–41. [CrossRef]
8. Grunwald, A. Strategic knowledge for sustainable development: The need for reflexivity and learning at the interface between science and society. *Int. J. Foresight Innov. Policy* **2004**, *1*, 150. [CrossRef]
9. Bammer, G. Guiding Principles for Integration in Natural Resource Management (NRM) as a Contribution to Sustainability. *Australas. J. Environ. Manag.* **2005**, *12*, 5–7. [CrossRef]
10. Kerkhoff, L.v.; Lebel, L. Linking Knowledge and Action for Sustainable Development. *Annu. Rev. Environ. Resour.* **2006**, *31*, 445–477. [CrossRef]
11. Weiss, E.B. United Nations Conference on Environment and Development. *Int. Leg. Mater.* **1992**, *31*, 814–817. [CrossRef]
12. ESD Section, UNESCO. UN Decade of Education for Sustainable Development: The First Two Years. *J. Educ. Sustain. Dev.* **2007**, *1*, 117–126. [CrossRef]
13. Leicht, A.; Heiss, J.; Byun, W.J. *Issues and Trends in Education for Sustainable Development*; UNESCO Publishing: Paris, France, 2018; Volume 5.
14. Purnamasari, S.; Suhendi, F.A.F.; Zulfah, N.L.N. Implementasi Education for Sustainable Development (ESD) dalam pembelajaran IPA di Kabupaten Garut: Sebuah studi pendahuluan. *J. Kaji. Pendidik. IPA* **2022**, *2*, 7–12. [CrossRef]
15. Sterling, S.; Orr, D. *Sustainable Education: Re-Visioning Learning and Change*; Green Books for the Schumacher Society: Totnes, UK, 2001; Volume 6.
16. Fisher, J.R.; Wood, S.A.; Bradford, M.A.; Kelsey, T.R. Improving scientific impact: How to practice science that influences environmental policy and management. *Conserv. Sci. Pract.* **2020**, *2*, e210. [CrossRef]
17. Remington-Doucette, S.M.; Connell, K.Y.H.; Armstrong, C.M.; Musgrove, S.L. Assessing sustainability education in a transdisciplinary undergraduate course focused on real-world problem solving. *Int. J. Sustain. High. Educ.* **2013**, *14*, 404–433. [CrossRef]
18. Millis, B.J.; Cottell, P.G., Jr. *Cooperative Learning for Higher Education Faculty*; Series on Higher Education; Oryx Press: Phoenix, AZ, USA, 1997.
19. Bruffee, K.A. *Collaborative Learning: Higher Education, Interdependence, and the Authority of Knowledge*; Johns Hopkins University Press: Baltimore, MD, USA, 1993.
20. Rulifson, G.; Bielefeldt, A. Motivations to Leave Engineering: Through a Lens of Social Responsibility. *Eng. Stud.* **2017**, *9*, 222–248. [CrossRef]
21. Smith, J. Impacts of a humanitarian engineering education pathway on student learning and graduate outcomes. *Int. J. Serv. Learn. Eng. Humanit. Eng. Soc. Entrep.* **2019**, *14*, 1–20. [CrossRef]
22. Amadei, B.; Wallace, W.A. Engineering for humanitarian development. *IEEE Technol. Soc. Mag.* **2009**, *28*, 6–15. [CrossRef]
23. Wiek, A.; Withycombe, L.; Redman, C.L. Key competencies in sustainability: A reference framework for academic program development. *Sustain. Sci.* **2011**, *6*, 203–218. [CrossRef]
24. Cebrián, G.; Junyent, M. Competencies in education for sustainable development: Exploring the student teachers' views. *Sustainability* **2015**, *7*, 2768–2786. [CrossRef]
25. Sakhaee, E. The AAUL Framework of Leadership in Times of Crisis. In *Organization and Leadership in Disruptive Times*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 97–108.

26. Chaplin, G.; Wyton, P. Student engagement with sustainability: Understanding the value–action gap. *Int. J. Sustain. High. Educ.* **2014**, *15*, 404–417. [CrossRef]
27. Merritt, E.; Hale, A.; Archambault, L. Changes in Pre-Service Teachers' Values, Sense of Agency, Motivation and Consumption Practices: A Case Study of an Education for Sustainability Course. *Sustainability* **2018**, *11*, 155. [CrossRef]
28. Nousheen, A.; Zai, S.A.Y.; Waseem, M.; Khan, S.A. Education for sustainable development (ESD): Effects of sustainability education on pre-service teachers' attitude towards sustainable development (SD). *J. Clean. Prod.* **2020**, *250*, 119537. [CrossRef]
29. Yamanishi, K.; Li, H. Mining open answers in questionnaire data. *IEEE Intell. Syst.* **2002**, *17*, 58–63. [CrossRef]
30. Galhardi, L.B.; Brancher, J.D. Artificial Intelligence-IBERAMIA 2018. In Proceedings of the 16th Ibero-American Conference on AI, Trujillo, Peru, 13–16 November 2018; pp. 380–391. [CrossRef]
31. Mohler, M.; Bunescu, R.; Mihalcea, R. Learning to grade short answer questions using semantic similarity measures and dependency graph alignments. In Proceedings of the 49th Annual Meeting of the Association for Computational linguistics: Human Language Technologies, Portland, OR, USA, 19–24 June 2011; pp. 752–762.
32. Meurers, D.; Ziai, R.; Ott, N.; Kopp, J. Evaluating answers to reading comprehension questions in context: Results for German and the role of information structure. In Proceedings of the TextInfer 2011 Workshop on Textual Entailment, Edinburgh, UK, 30 July 2011; pp. 1–9.
33. Koufakou, A.; Gosselin, J.; Guo, D. Using Data Mining to Extract Knowledge from Student Evaluation Comments in Undergraduate Courses. In Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN), Vancouver, BC, Canada, 24–29 July 2016; pp. 3138–3142. [CrossRef]
34. Abd-Elrahman, A.; Andreu, M.; Abbott, T. Using text data mining techniques for understanding free-style question answers in course evaluation forms. *Res. High. Educ. J.* **2010**, *9*, 1.
35. Bird, S.; Klein, E.; Loper, E. *Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit*; O'Reilly Media, Inc.: Newton, MA, USA, 2009.
36. Available online: https://spacy.io/models/en#en_core_web_sm (accessed on 31 July 2022).
37. Weischedel, R.; Palmer, M.; Marcus, M.; Hovy, E.; Pradhan, S.; Ramshaw, L.; Xue, N.; Taylor, A.; Kaufman, J.; Franchini, M.; et al. *Ontonotes Release 5.0 ldc2013t19*; Linguistic Data Consortium: Philadelphia, PA, USA, 2013; Volume 23.
38. Fellbaum, C. WordNet and wordnets. *Encyclopedia of Language and Linguistics*, 2nd ed.; Elsevier : Amsterdam, The Netherlands, 2005; pp. 665–670.
39. Peason, K. On lines and planes of closest fit to systems of point in space. *Philos. Mag.* **1901**, *2*, 559–572. [CrossRef]
40. Jones, K.S. A statistical interpretation of term specificity and its application in retrieval. *J. Doc.* **1972**, *28*, 11–21. [CrossRef]
41. Aizawa, A. An information-theoretic perspective of tf–idf measures. *Inf. Process. Manag.* **2003**, *39*, 45–65. [CrossRef]
42. Liu, J.; Chua, T.S. Building semantic perceptron net for topic spotting. In Proceedings of the 39th Annual Meeting on Association for Computational Linguistics-ACL '01, Toulouse, France, 6–11 July 2001; pp. 378–385. [CrossRef]
43. Dillon, J.; Bosch, N.; Chetlur, M.; Wanigasekara, N.; Ambrose, G.A.; Sengupta, B.; D'Mello, S.K. Student Emotion, Co-Occurrence, and Dropout in a MOOC Context. *Int. Educ. Data Min. Soc.* **2016** .
44. White, A.E.; Moeller, J.; Ivcevic, Z.; Brackett, M.A. Gender identity and sexual identity labels used by US high school students: A co-occurrence network analysis. *Psychol. Sex. Orientat. Gend. Divers.* **2018**, *5*, 243. [CrossRef]
45. Leydesdorff, L.; Nerghes, A. Co-word maps and topic modeling: A comparison using small and medium-sized corpora (N < 1,000). *J. Assoc. Inf. Sci. Technol.* **2017**, *68*, 1024–1035. [CrossRef]
46. Rude, S.; Gortner, E.M.; Pennebaker, J. Language use of depressed and depression-vulnerable college students. *Cogn. Emot.* **2004**, *18*, 1121–1133. [CrossRef]
47. Robinson, R.L.; Navea, R.; Ickes, W. Predicting Final Course Performance From Students' Written Self-Introductions. *J. Lang. Soc. Psychol.* **2013**, *32*, 469–479. [CrossRef]
48. Syah, T.; Nurhayaty, A.; Apriyanto, S. Computerized Text Analysis on Self-Description Text to Get Student's Prevailing, Confidence, and Drives. *J. Phys. Conf. Ser.* **2021**, *1764*, 012056. [CrossRef]
49. Pennebaker, J.W. Putting stress into words: Health, linguistic, and therapeutic implications. *Behav. Res. Ther.* **1993**, *31*, 539–548. [CrossRef]
50. Pennebaker, J.W.; Francis, M.E.; Booth, R.J. Linguistic inquiry and word count: LIWC 2001. *Mahway Lawrence Erlbaum Assoc.* **2001**, *71*, 2001.
51. Pennebaker, J.W.; Booth, R.J.; Francis, M.E. LIWC2007: Linguistic inquiry and word count. *Austin Tex. Liwc. Net* **2007**.
52. Pennebaker, J.W.; Booth, R.J.; Boyd, R.L.; Francis, M.E. *Linguistic Inquiry and Word Count: LIWC2015*; Pennebaker Conglomerates: Austin, TX, USA, 2015. Available online: www.LIWC.net (accessed on 31 July 2022).
53. Watson, M.K.; Pelkey, J.; Noyes, C.; Rodgers, M.O. Using Kolb's learning cycle to improve student sustainability knowledge. *Sustainability* **2019**, *11*, 4602. [CrossRef]
54. Mehl, M.R.; Pennebaker, J.W. The Sounds of Social Life: A Psychometric Analysis of Students' Daily Social Environments and Natural Conversations. *J. Personal. Soc. Psychol.* **2003**, *84*, 857–870. [CrossRef]
55. Schnurr, P.P.; Rosenberg, S.D.; Oxman, T.E.; Tucker, G.J. A Methodological Note on Content Analysis: Estimates of Reliability. *J. Personal. Assess.* **1986**, *50*, 601–609. [CrossRef]