2023

The Use of Artificial Intelligence to Detect Students Sentiments and Emotions in Gross Anatomy Reflections

Krzysztof J. Rechowicz
Old Dominion University, krechowi@odu.edu

Carrie A. Elzie

Original Publication Citation

Rechowicz, K. J., & Elzie, C. A. (2023). The use of artificial intelligence to detect students' sentiments and emotions in gross anatomy reflections. *Anatomical Sciences Education*, 13 pp. https://doi.org/10.1002/ase.2273

Received: 11 March 2022 | Revised: 9 February 2023 | Accepted: 14 March 2023

DOI: 10.1002/ase.2273

RESEARCH REPORT

Anatomical Sciences Education WILEY

The use of artificial intelligence to detect students' sentiments and emotions in gross anatomy reflections

Krzysztof J. Rechowicz[1] | **Carrie A. Elzie**[2]

[1]Virginia Modeling, Analysis, and Simulation Center, Old Dominion University, Suffolk, Virginia, USA

[2]Department of Pathology and Anatomy, Eastern Virginia Medical School, Norfolk, Virginia, USA

Correspondence
Dr. Carrie A. Elzie, Department of Cell Systems and Anatomy and the Department of Medical Education, University of Texas Health San Anatonio, 7703 Floyd Curl Drive, San Antonio, TX 78229. USA.
Email: celzie@gmail.com

Abstract

Students' reflective writings in gross anatomy provide a rich source of complex emotions experienced by learners. However, qualitative approaches to evaluating student writings are resource heavy and timely. To overcome this, natural language processing, a nascent field of artificial intelligence that uses computational techniques for the analysis and synthesis of text, was used to compare health professional students' reflections on the importance of various regions of the body to their own lives and those of the anatomical donor dissected. A total of 1365 anonymous writings (677 about a donor, 688 about self) were collected from 132 students. Binary and trinary sentiment analysis was performed, as well as emotion detection using the National Research Council Emotion Lexicon which classified text into eight emotions: anger, fear, sadness, disgust, surprise, anticipation, trust, and joy. The most commonly written about body regions were the hands, heart, and brain. The reflections had an overwhelming positive sentiment with major contributing words "love" and "loved." Predominant words such as "pain" contributed to the negative sentiments and reflected various ailments experienced by students and revealed through dissections of the donors. The top three emotions were trust, joy, and anticipation. Each body region evoked a unique combination of emotions. Similarities between student self-reflections and reflections about their donor were evident suggesting a shared view of humanization and person centeredness. Given the pervasiveness of reflections in anatomy, adopting a natural language processing approach to analysis could provide a rich source of new information related to students' previously undiscovered experiences and competencies.

KEYWORDS

dissection, gross anatomy, health professions education, machine learning, natural language processing, reflections, reflective writing, sentiment analysis

INTRODUCTION

The importance of reflection and reflective practice is regarded by many as an essential characteristic for professional development.[1] Within medicine, the explicit awareness of one's own experiences deepens the capacity to respond empathically to patients and is the crux of relationship-centered care.[2,3] Exploring one's narrative through reflection and self-awareness allows better listening to and accepting of another's narrative.[3] Reflection also allows for integrating concepts or a combination of skills, knowledge, attitudes, and values with the learner's cognitive framework.[4-6] This integration usually requires an understanding or exploration of one's personal beliefs and experiences to be interpreted and expanded to form new knowledge in an active, experiential manner. Meaning can then be constructed within a community of professional discourse, encouraging learners to achieve and maintain critical control over the more intuitive aspects of their experience.[7] It is considered an essential aspect of lifelong self-learning and core to professional competency.[8] Because of this, reflective activities are becoming integral to curricula at all levels of health professional and medical education with positive impacts.[7] Reflection by undergraduate medical students has been shown to increase self-reported measures of self-awareness, professional thinking skills, and the skills required for intimate examinations.[9] Additionally, reflection has been shown to have a positive effect on cross-cultural understanding,[10] diagnostic accuracy,[11] empathy,[12] feedback integration,[13] and well-being.[14]

Since the cadaver–student relationship is thought to mirror the patient–physician relationship,[15] gross anatomy education has often been intentionally humanistic.[16] As one of the first courses in medical and health professional curricula, gross anatomy provides a rich setting for the practice of reflection to promote professionalism, integrate clinical concepts, and process emotions around death and dying.[17] Examples of reflective activities in anatomy have included anatomical donor reports,[18] attitudinal questionnaires,[19] creative projects,[20] journals,[21] personal or fictional narratives,[22] and problem-oriented dissection.[23] Students' reflective writings related to anatomical donors have revealed complex and difficult emotions, as well as an appreciation for the dissection experience.[24,25] In this study, a reflective exercise was used to garner self-awareness and to shift the thinking from "specimen-minded" to "person-minded," as this is associated with empathy, emotional engagement, and better performance within the course.[26] While the use of reflective writings in anatomy to this end is not new, this study used reflective writings in a novel way to explore the contribution of various anatomical body regions to one's life.

While it is widely acknowledged that reflective practices are beneficial, assessing reflection can be challenging, often requiring tremendous resources and time.[27-30] Some researchers have employed a quantitative system for evaluation such as questionnaires,[31] scales,[32] and rubrics.[33] While these are efficient to grade, they provide a limited, structured, and often shallow reflective analysis. Assessment of narratives has also been done by stating judgments about the learners' abilities or engagement with the exercise.[34] Numerous models and frameworks with varying complexity have been developed to analyze reflections.[35] However, many of these mechanisms require interrater and internal consistency reliability to develop scores, often lack validity, and are demanding in terms of resources.[35-37] Even when thematic analysis is utilized, there can be a tension between the reliability of coding schemes and their ability to discriminate between learners.[30]

Natural language processing is a promising and nascent way to potentially overcome some of the challenges with analysis of student reflections, especially for large sets of open-ended text and limited time. Natural language processing is one of the applications of artificial intelligence which enables computers to understand natural language as humans do. Whether the language is spoken or written, natural language processing takes real-world input, processes it, and makes sense of it in a way a computer can understand.[38] While this does not replace the richness of information that can be extracted using qualitative analysis, the use of natural language processing to analyze text data can be a valuable means of discovering underlying sentiments and themes.[39] Within medical education, natural language processing has been used to map clinical notes to school-identified learning objectives,[40] track medical students' clinical experiences,[41] and assess clinical exposure to training competencies.[42] It has also been used to determine if gender bias exists in evaluation of residents.[43] A widely used technique in natural language processing is sentiment analysis. Sentiment analysis is an automated process in which natural language processing, text analysis, computational linguistics, and biometrics are used to systematically identify, extract, quantify, and study affective states and subjective information.[44] Also known as opinion mining, it is the automated task of determining what feelings a participant expresses in text, typically framed as the binary distinction of positive and negative sentiments.[45] Sentiment analysis can be applied to a document, sentence, or aspect (levels) and produce a score, rating, or polarity (positive, negative, or neutral). For example, sentiment analysis has been used to analyze the association of narrative feedback given on clerkships with assigned grades of medical students[46] and to assess the sentiment associated with competency evaluations of medical residents.[47] While sentiment analysis aims to detect positive, neutral, or negative feelings from text, emotion detection aims to detect and recognize types of feelings through the expression of texts, such as anger, disgust, fear, happiness, sadness, and surprise.[48] Emotion detection is often performed using two main techniques: machine learning, which extracts features using a criterion or a combination of criteria, and a lexicon-based approach involving dictionaries with mapped words to emotions.[49] Thus, leveraging sentiment analysis and emotion detection, underlying attitudes associated with open-ended response data can be determined, or feelings of groups of people can be assessed without directly asking for purposes of formative or summative assessments. Applying these principles to students' reflective writings can produce a rich data source. For example, Lin et al.[50] used Linguistic Inquiry and Word Count to investigate overall language patterns and gender differences in reflective writings of medical students regarding pediatric patients and the psychosocial challenges faced by the patients and their family members. Another study used a process-oriented text mining approach to better understand meanings of learner experiences within a community placement by connecting key concepts in extended student reflective essays.

Students' sentiments in an anatomy course can have lasting impacts on the future of their medical practice and professionalism[15,52,53]; therefore, it is essential to capture and analyze them early in training. Given the vast emotional experiences of students within gross anatomy, we aimed to quantitatively compare students' reflective writings on the importance of their anatomy to their life and the proposed importance of anatomy to their anatomical donors. It was hypothesized that writings about their anatomical donors would be similar as writings about themselves—lending toward a humanistic view of the cadaver. Writings about various regions of the body were expected to elicit varying emotions, both positive and negative. To test this, natural language processing was employed to analyze students' writings for sentiment and emotions using a lexicon-based approach. Specifically, the goals were to (1) determine if particular emotions were correlated with specific body parts/organs and (2) determine if students shared similar sentiments and emotions about their donor's anatomy compared to their own.

METHODS AND MATERIALS

Course description

The gross anatomy course was a 16-week, semester long course running from January to April of 2021 for health professional students (n = 132). The course was comprised of physician assistant (n = 82), pathologist assistant (n = 25), and surgical assistant (n = 25) students of which 39 were males. Other demographic information was not collected. The course comprised 124 contact hours but was delivered as a hybrid because of the COVID-19 pandemic. All lectures (n = 36) were pre-recorded, nonsynchronous and approximately 50 minutes. There were 22, 50 minute long, synchronous active learning sessions that included cases, living donor presentations, distributed virtual reality, and gamification reviews. Students performed whole body dissection in groups of three every other laboratory session using a published dissector. There were two prosection laboratory sessions (hand and foot). Students were assessed with four block exams that included multiple choice questions to assess didactic material and practical identification questions tagged on the anatomical donors within the laboratory.

Emphasis on the "anatomical donor," as opposed to "cadaver" was an important part of the humanistic element of this course. Students were introduced to each donor by their first names at the beginning of the course. In lieu of dissecting during the first laboratory, students instead completed a whole body examination of the donor which encouraged them to use observational skills to document changes associated with aging and or disease. During this time students were provided with the donor's assigned gender at birth, age at the time of death, occupation, cause of death, and relevant medical history (if provided by the donor). Students also watched a video from a living anatomical donor who detailed his memories of the importance of anatomy to his own life, similar to the reflection prompts asked of the students.[54]

Reflective writing

At the beginning of the course (week 1), students were asked to reflect on what exactly their anatomy has meant to them thus far in their life: "I would like you to reflect on what exactly your anatomy has meant to you thus far in your life. You must at least choose 5 of the following 17 regions within the body and write a short paragraph about how/why it is important to you. There is an 'other' option in case you want to write about something not listed on the form." The body regions included the arms, back, brain, ears, eyes, face, feet, gastrointestinal system, gluteals, hands, heart, knees, lungs, mouth, nose, pelvis, and skin. Students were given the same writing prompt at the end of the semester and completion of a whole-body dissection (week 16) but asked to speculate how anatomy had contributed to their anatomical donors' lives using the same body regions. The specific prompt read "I would like you to reflect on what your anatomical donor's story could have been. You must at least choose 5 of the following 17 regions within the body and write a short paragraph about how/why it might have been important to your donor. There is an 'other' option in case you want to write about something not listed on the form." Students were also asked to provide their donor table number, as groups of 10–12 students shared an anatomical donor. There was a total of 12 donors dissected, of which four were males and eight were females according to their death certificates. Course participation points (1% of their grade) were awarded for the completion of each reflective writing indicated by an attestation within the learning management system. However, reflective writings were gathered in a separate link using Google Forms (Google HQ Mountain View, California, USA) to provide anonymity. A total of 1365 anonymous reflective writings (677 about a donor, 688 about self) were collected. All data collected were utilized for this study (no exclusion or inclusion criteria). This study was deemed exempt by the Institutional Review Board of Eastern Virginia Medical School IRB # 21-12-XX-0279.

Data pre-processing

Reflective writings were downloaded into R statistical package, version 4.1.2 with supplementing packages (R Studio, Boston, Massachusetts, USA). Data were reviewed and corrected for typographical and grammatical errors. Data were formatted to ensure it was correctly encoded in UCS Transformation Format 8, a Unicode standard best suitable for representing text for mining. The preprocessing transformed the running text into a structured format that allowed computations. In addition to the narratives, data were categorized as about self or the donor and the body part referred to in the reflective writings.

Additionally, all reflective writings were broken into sentences. Tokenization, lemmatization, and stop words removal were all performed. Tokenization divides a document into characters grouped by a functional semantic unit. A token could be a word or a set of words, called n-grams. Lemmatization is a subset of stemming that reduces words to their roots (e.g., "am" becomes "be," and "better" becomes "good"). Subsequently, stop words were removed that did not carry much discriminative content using three lexicons (SMART, onix, snowball) containing 1149 commonly recognized stop words.[55]

To illustrate, it was expected that the word "and" would appear in all reflective writings with a similar frequency and, therefore, would not allow inference about a reflection's characteristics or content. Lastly, a custom dictionary of words had to be generated to remove the body regions' names (arms, back, etc.) since they frequently appeared (Appendix 1).

Data analysis

All analyses were conducted using R statistical software, version 4.1.2 (R Studio, Boston, Massachusetts, USA) with supplementing packages.[56-58] Word counts were initially calculated for each reflection without taking into account stop words. Then, the average reflection length and standard deviation were calculated for all reflective writings. Subsequently, median word counts of donor-related and self-reflections were compared. A Wilcoxon test was performed to establish if the difference between the medians was statistically significant at a p-value <0.05. Median word counts were also compared for all body regions. The relative frequency of what body region the students chose to reflect about was calculated. Additionally, a comparison was made between of regions chosen in self and donor reflective writings. The top 10 most common words in all reflective writings were identified. Like the previous step, 10 words with a high occurrence rate were broken down by the reflection type, that is, self versus donor. Stop words and words included in the custom dictionary were not counted. Word frequencies across reflective writings about self and donors were compared. Pearson's correlations between word occurrences in distinct sets of reflective writings were calculated.

Next, sentiment and emotion detection analyses were performed on the students' reflections. For a general review of the process see Nandwani and Verma[59]. Sentiment analysis was performed utilizing an approach that considers valence shifters when the polarity is calculated.[60] Valence shifters, such as negators (e.g., not, no, never), amplifiers (e.g., really, very, precisely), de-amplifiers (e.g., hardly, maybe, probably), and adversative conjunctions (e.g., but, still, yet) can affect words' sentiment.[61] This method was implemented in R's `sentimentr` package.[57] Each reflective writing was broken down into sentences and subsequently into words in this method. Words, excluding stop words, were tagged according to their lexical polarity. The lexicon used was the Jockers and Rinker[62] polarity lookup table. The dictionary used was an augmented combination of the lexicon developed by Jockers[63] and Riker's augmentation of the lexicon developed by Hu and Liu[64]. The dictionary contained 11,710 words identified as polarized positively or negatively. Around each polarized word, a cluster including four preceding and two following words was created. The algorithm[60] searched that neighborhood for valence shifters, and, if found, the sentiment value was recalculated. Punctuation (comma, colon, semicolon) was used to delimit the effect of the valence shifters on the word. Cluster's value resulted from applying weights based on the valence shifter type and their number in the cluster. For example, an amplifier increased the base polarity score by a number greater than 1, whereas a de-amplifier by a number greater than 0 and smaller than 1. Negation was determined by raising −1 to the power of the number of negators plus 2. This resulted from a belief that an even number of negatives equal a positive, whereas an odd number gives a negative. The adversative conjunction up-weighted or down-weighted the context cluster depending on their location, before or after the polarized word. Similarly to negators, it was essential to consider the number of preceding and subsequent conjunctions. To obtain an *unbounded polarity score* for each sentence in a reflection, the weighted context clusters were summed and divided by the square root of the word count. In the end, a reflection's sentiment was calculated as a mean of sentences' polarity scores. The top 10 words contributing to positive and negative sentiment were identified for all reflective writings and separately for those about self and a donor. Again, the lexicon developed by Rinker was utilized in this step.[62]

Lastly, lexicon-based emotion classification was performed utilizing the National Research Council Emotion Lexicon (EmoLex)[65] which lists 14,182 English words and their associations with eight basic emotions (anger, fear, anticipation, trust, surprise, sadness, joy, and disgust) and two sentiments (negative and positive). The annotations for this were initially manually done by crowdsourcing. The results of the tokenization process were used for this analysis, that is, a set of tokens corresponding to all words used in reflective writings without the stop words. Then, an inner join operation was performed on the list of words and the EmoLex dictionary[65] to assign an emotion to each word in the reflective writings, unless the particular word was not part of the lexicon. Those words were omitted. The resulting list of words with assigned emotions was used to characterize what feelings were present when the student reflected on a particular body region about themselves or their donor.

RESULTS

Word count

The length of student writings varied across body regions and type of reflection (about self compared to donor). The average word count for all reflective writings was 34.6 (SD ±18.8) words. The shortest reflection was four words, and the longest was 169 words long. On average, reflective writings about the donor averaged 29.4 (±14.9) words, whereas self-reflections averaged 39.6 (±20.7) words. Median word counts of donor and self-reflections showed significant difference ($p < 0.05$). The difference between the word count is also visible when the data were broken down by the type of reflection and the body regions. Reflective writings about donors' noses were the least elaborated. In contrast, reflective writings about students' knees were the most elaborated based on median word count.

Body regions

Students were compelled to write about different body regions; however, every option for a body region was reflected on by students

(Figure 1). Overall, students most frequently chose the hands, heart, and brain to write about. When reflecting on self, the students chose hands most of the time (12%), eyes (11%), and brain (10%). When reflecting on the donor, the students chose to write most frequently about the heart (13%), hands (12%), and brain (11%). In the case of donor and self-reflections, gluteals accounted for 1% and 1.3% reflective writings, respectively. However, the least selected body region was the donor's nose (0.9%). The gastrointestinal system and pelvis were chosen more often when the reflective writings were about the donor than the self. The opposite held valid, for the back. Other areas (0.02%, $n = 14$) chosen by students to write about included their appendix, breast, hair, kidneys, larynx, legs, and nails (data not shown). Only two students chose "other" for their donors (larynx and kidneys).

Word frequency comparison

Students used similar words when reflecting about themselves as they did when reflecting about their donors. The most frequently written words in all reflective writings were "life," followed by "body" and "allowed." The order of those words changed when the reflective writings were separated into donor-related and self-reflections. Words "body" and "life" were the most common when the students wrote about themselves. The word "allowed" was not present in the list of the 10 most common words in those reflective writings. When the donor-related reflective writings were considered, "life" and "allowed" were the most occurring words, followed by the word "loved." Figure 2 illustrates the comparison between frequencies of words appearing in reflective writings about self and donors. Words closer to the diagonal line appear in both types of reflective writings with similar frequency. Terms further from the line are found more in one set of texts than another. Words that appear in only one type of reflective writings were omitted. The data show a more extensive set of words that often appear in reflective writings about self than about donors. Pearson's correlation coefficient was calculated between word frequency for reflective writings about self and donor. The coefficient of 0.76 indicated a strong positive correlation between the two datasets.

Sentiment analysis

In general, students wrote positively about themselves and their donors. Using binary sentiment analysis, over 80% of reflective writings were positive for self and donor types. The mean positive sentiment was 0.22 ($SD = 0.14$) for donor-related reflective writing and 0.22 ($SD = 0.13$) for self-reflection, while the mean negative sentiment was −0.14 ($SD = 0.13$) and −0.10 ($SD = 0.09$), respectively. There was no statistically significant difference between the means of sentiment for each polarity level when the subject of the reflective writing was considered (Figure 3A).

For three levels of polarity (negative, neutral, positive), where "neutral" polarity was assigned to sentiment values between −0.2 and 0.2, most reflective writings were neutral, over 57% for self-reflection and almost 53% for reflections about donors, followed by positive sentiment, almost 40% and 43%, respectively. The mean positive sentiment was 0.32 ($SD = 0.11$) for donor-related reflective writing, and 0.33 ($SD = 0.11$) for self-reflections, while the mean negative sentiment was −0.31 ($SD = 0.11$) and −0.26 ($SD = 0.06$), respectively. The mean neutral sentiment was 0.06 ($SD = 0.10$) for both types of reflective writing. There was no statistically significant difference between the means of sentiment for each polarity level when the subject of the reflective writing was considered (Figure 3B).

Words "pain" and "issues" were the most often occurring words contributing to negative sentiment for the donor-related

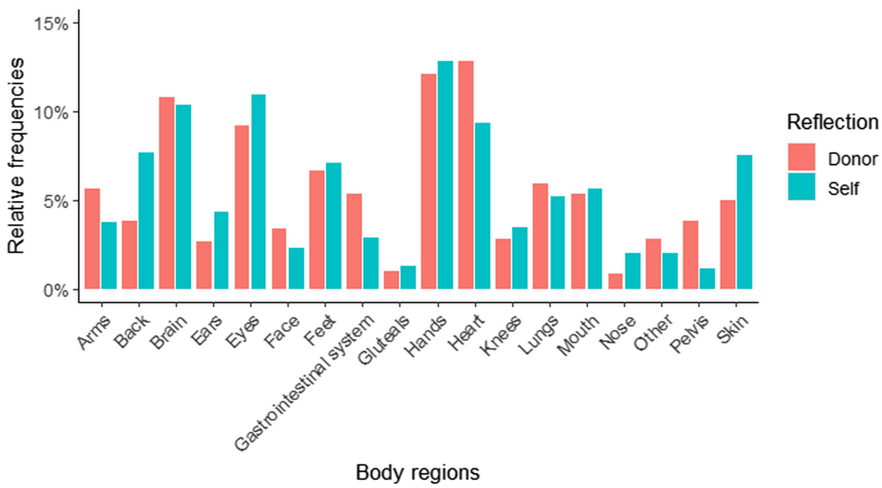

FIGURE 1 Comparison of relative frequencies of body regions students ($n = 132$) wrote about in reflections about donors ($n = 677$) and self ($n = 688$). Frequencies for each body region were calculated by dividing the number of reflections the students chose to reflect on by the number of reflections of the particular type. While self-reflecting, students chose to reflect mostly on hands ($n = 88$), in contrast to reflections about the heart ($n = 87$) when reflecting about the donor.

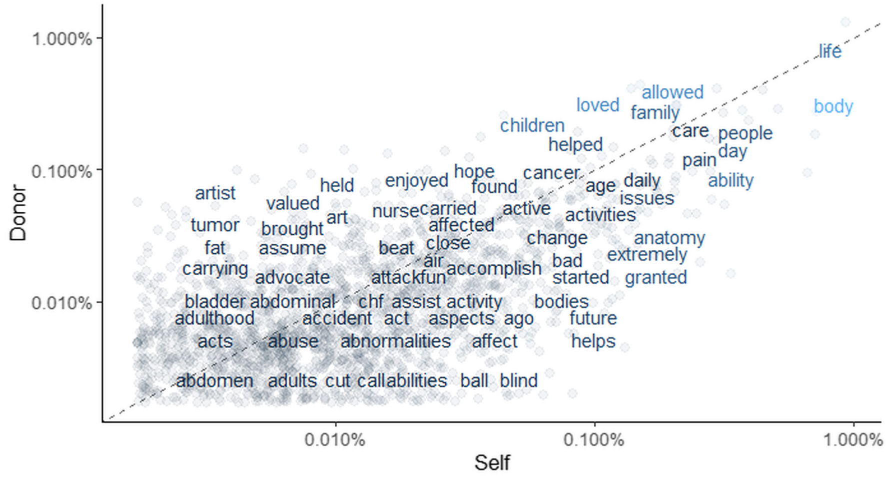

FIGURE 2 Comparison of word frequency in reflections about self ($n = 688$) and donors ($n = 677$) in the log–log scale. Only words appearing in both types of reflections were plotted ($n = 1952$).

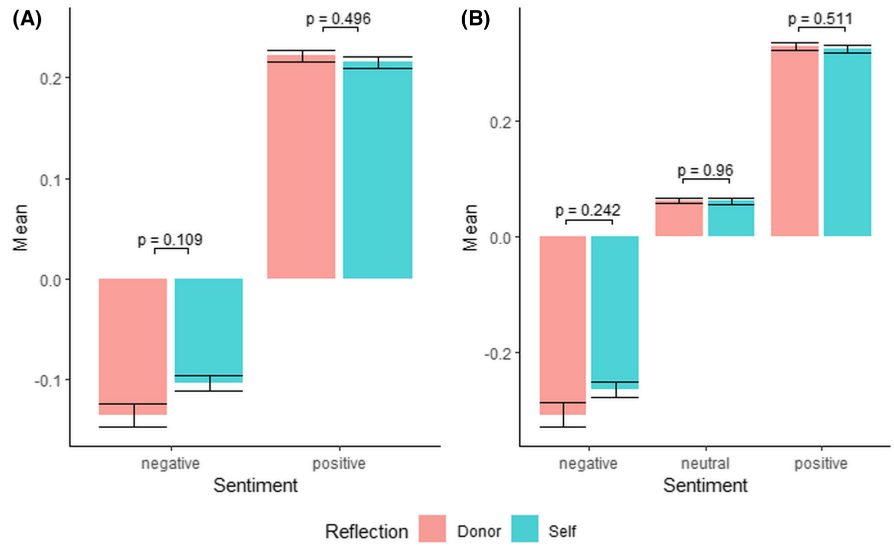

FIGURE 3 (A) Comparison of two-level (positive and negative) sentiment mean value between donor-related ($n = 677$) and self-reflections ($n = 688$). There was no statistically significant difference between mean sentiment value ($p > 0.05$); (B) Comparison of three-level (positive, neutral, and negative) sentiment mean between donor-related ($n = 677$) and self-reflections ($n = 688$). There was no statistically significant difference between mean sentiment value ($p > 0.05$).

reflective writings, whereas the words “love” and “loved” to positive sentiment. In general, more words contributed to the positive than negative sentiment, which is explained by the overall positive sentiment of the reflective writings. When the words contributing to positive and negative sentiment were broken down by the reflection type, that is, about self or a donor, “pain” was still the most often occurring word with negative polarity. At the same time, “cancer” moved down to the eighth position for reflective writings about self. On the positive side, changes to the order of words were smaller. “Love” and “ability” were the top two words with positive polarity. When the words contributing to the positive and negative sentiment of reflective writings about donors were considered, “pain” and “cancer” were the top two words. At the same time, “hard” was the third most often occurring word with negative polarity. “Loved” and “care,” followed by “helped,” were high-occurring positive words. Differences in the top 10 terms, negatively and positively influencing sentiment, likely reflect differences in the demographics of medical students and donors and their respective medical histories.

Emotions

A variety of emotions were detected in student writings and with certain regions of the body conveying differing emotions. To extract emotions from reflective writings, a lexicon-based approach was utilized. An emotion lexicon is a specific type of linguistic resource that maps the emotive or affective vocabulary to a fixed set of emotion labels. Each entry in the dictionary associates a word with zero or more emotion labels. The EmoLex lexicon[65] applied here was based on Plutchik's[66] classification of emotions into eight categories: anger, fear, sadness, disgust, surprise, anticipation, trust, and joy. When emotions were analyzed, the levels of emotional sentiment per emotion were similar when reflective writings about the donor and self were compared. The most considerable difference was for anger 8.2% (self) versus 5.9% (donor) (Figure 4). Regardless of the reflection type, the top three emotions were trust, joy, and anticipation. Disgust, anger, and surprise were the bottom three emotions. In general, trust, joy, and anticipation are the dominating emotions for almost all body regions. Trust was present for all body regions, with lungs, nose, and pelvis being the lowest. Surprise was relatively low for all body regions except the mouth. Sadness varied depending on the body part, with the nose having none. Joy was high for the arms, mouth, and ears and consistent for all body regions except reflective writings about the back, knees, and lungs. Fear was higher for several body regions, such as the lungs, knees, pelvis, and back. It was low for the arms, ears, hands, and nose. Disgust was relatively low for all body regions except the nose, which was likely associated with words related to smells. The high proportions of disgust decreased the ratios of other emotions. Anticipation was relatively high for almost all body regions, with the arms being the highest and the nose the lowest. Anger was the highest for the nose, which is already skewed toward negative emotions. When the reflective writings were broken down by type (self vs. donor) (Figure 5), the emotion distributions follow similar patterns. In student self-reflections, the nose had a higher proportion of "sadness" and "fear" when compared to the same donor's body region.

DISCUSSION

The mobilization of personal experience in academic writing is essential for professional identity formation.[12] As such, this study asked students to reflect on how their anatomy has contributed to their own lives and to speculate how anatomy mattered to their anatomical donor's life. The latter prompt encouraged students to imagine their donor's lives and ascribe a greater purpose to the anatomy studied throughout the course. In essence, it encouraged a more "person-minded" view of the donor as opposed to a "specimen-minded" view.[67,68] This illuminated specific body areas that resonated most with students, such as the heart, hands, and brain, consistent between self and donor reflective writings. This aligns with previously published studies where students also commonly wrote about these body regions.[22,25,26,69] Interestingly, many students wrote about their love/hate relationship with their own skin and its importance to beauty and identity, which was much less common in donor reflective writings. Other differences in regions chosen to write about may have been influenced by the fact that students were provided with their donor's occupation at the introduction of the course. The use and hindrance of their donor's anatomy on donors' careers were commonly mentioned (data not shown).

Sentiment analysis

Prior survey and thematic analyses of sentiments related to anatomy courses have revealed both positive and negative responses. While a positive experience for many students,[52,70,71,72,73] gross anatomy can also be distressing for some.[74-79] However, these previous reports were primarily about the overall course experience or dissection process. Our data are the first (to our knowledge) to report various sentiments conjured by distinct body regions and their impact on one's life (self and donor). Most body regions elicited a neutral response related more to their function and utility. However, when using a dual scale (positive or negative), most self and donor reflective writing sentiments were

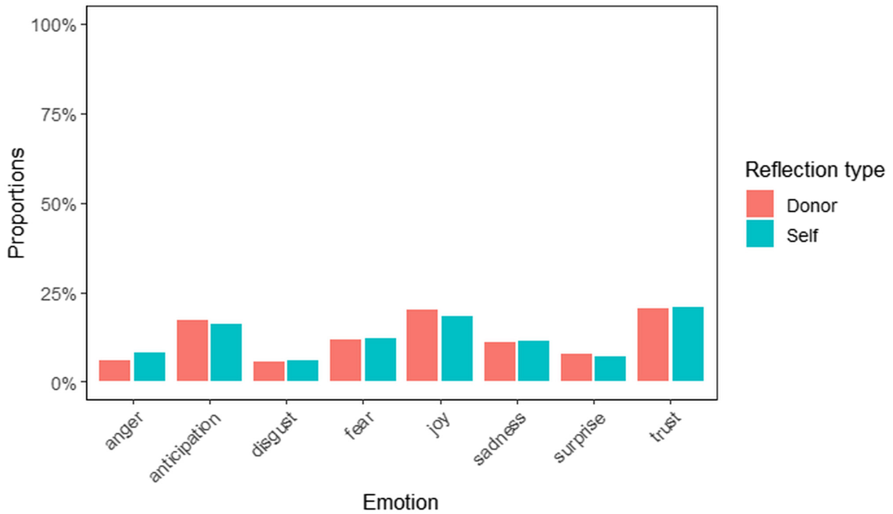

FIGURE 4 Proportions of emotions per reflection type. Proportions of emotions were a ratio of the sum of occurrences of words associated with a particular emotion to the sum of occurrences of words associated with any out of eight emotions. Out of all occurrences in reflections about a donor of words associated with an emotion (n = 5307), words related to "trust" (n = 1350), "joy" (n = 1027), and "anticipation" (n = 941) were the most common. Out of all occurrences in reflections about self of words associated with an emotion (n = 7140), words related to "trust" (n = 1796), "joy" (n = 1250), and "anticipation" (n = 1.161) were also the most common.

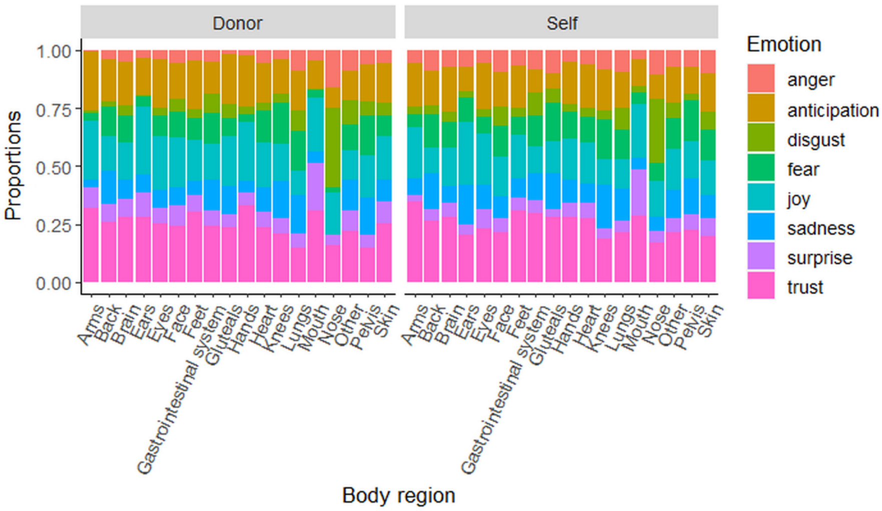

FIGURE 5 Comparison of emotions' proportions per body region between self- ($n = 688$) and donor-related ($n = 677$) reflections. The proportions were calculated as a ratio of the sum of occurrences of all words in reflections, about self or a donor, regarding a particular body region associated with an emotion to the sum of occurrences of all words in reflections about the body region related to any out of eight emotions.

positive, with subtle differences in word contributions. While both expressed words of love, self-reflections focused on the "ability" of one's anatomy as a positive contributor to their life. This was reversed when looking at the contribution of negative words which were more disease focused (words such as "cancer" and "issues") in donor reflective writings. This was not unexpected given the advanced age of the donors and disease accumulated throughout a lifetime compared to students' relative youth and healthiness. Student self-reflections with negative sentiments were mostly related to previous personal "pain" and "issues." These personal illness stories are often central to students' motivations for becoming healthcare professionals. They may critically inform the nature of students' professional caregiving.[12] Additionally, a detailed analysis of the types of language used in reflective writings can indicate the reflectors' preferred mode of processing their experience and constructing their narratives.[50] Thus, further analysis of what factors may influence the language style used by students in their anatomy reflective writings is appealing. It could inform students' psychological well-being and ability to empathize with patients by enabling them to access and accept their feelings.

Emotions

Humanistic anatomy curricula, such as those practiced in this course, can engage students by validating the full range of emotional experiences of dissection and acknowledging the protective role of emotional detachment in medical practice.[16] Numerous studies have used a qualitative approach to describing students' complex emotional and moral experiences in gross anatomy.[22,77,80,81,82] While the reflective writings reported here were not directly related to the anatomical dissection, they still prompted a similar range of emotional responses experienced in the class as a whole. Furthermore, the general similarity of the words and emotions used in writings about self compared to the donor which suggests a shared view of humanization and person centeredness toward the anatomical donor. The higher levels of positive emotions align with the positive nature of the sentiment analysis reported here and the appreciation for one's body, a common theme in anatomy reflective writings.[24] While very little has been reported on emotions related to specific body regions, particular patterns have been studied in East Asian Medicine and surprisingly coordinate with the data presented here.[83] Overall, the most frequently extracted emotions in this study were trust, joy, and anticipation. According to the broaden-and-build theory positive emotions can broaden thought–action mechanisms such as playing, exploring, and integrating which contribute to resiliency and well-being.[84] For example the emotion of "joy" is linked with an urge to play, push the limits, and be creative not only socially but also intellectually. A previous study indicated that students exhibit a spectrum of emotions within a gross anatomy course and that positive emotions are aroused while negative emotions may be suppressed.[85]

The advantage of using natural language processing to extract emotions is that it may uncover a wider range and more authentic spectrum of emotions because it does not rely on potentially skewed self-reporting. This type of understanding of the emotions of students could also provide insights into their cognitive processes,[86,87] learning,[88] perceived self-efficacy,[89] and professional identity formation.[90,91] Such emotional influences should be carefully considered in educational course design to maximize learner engagement, as well as improve learning and long-term retention of the material.[92] Since attentional and motivational components of emotion have been linked to heightened learning and memory,[93,94] it would be interesting to determine if a correlation exists between emotions reported here and outcomes on laboratory examinations.

Because this was not a direct longitudinal study (donor vs. donor over time), it is unclear if viewing the donor as a person over time declined over the course length as previously reported.[52] However, the data still support a high level of personhood toward the donors after 16 weeks of full dissection. Furthermore, insights gained here about students' emotional responses to various regions of the body may allow anatomists to respond appropriately to create a positive impact on the learning process as each region is discussed. Assessment of emotional responses over time may also provide a holistic picture of the students' emotional trajectory throughout the course and beyond.

Natural language processing

As the use of reflective writings in medical and health professional education rises, the challenge will remain in analyzing large sets of text and providing valuable feedback to learners. The true advantage of using natural language processing lies in an evaluator's ability to use a quantitative approach in analyzing text. It is especially valuable where the breadth of a dataset or constraints associated with time and available data scientists make it very difficult for practitioners to perform a comprehensive analysis. Since this method is purely quantitative, comparing sentiments across and between groups and longitudinally can be accomplished easily. The quantitative extraction of emotions may be helpful, considering evidence exists that self-reports about feelings may be misleading, as students may not be aware of the depth of their own emotions.[95] This type of objective measure of emotions from writings may provide educators with better insight for providing student feedback about self-awareness.

Building larger datasets related to student reflections could make deep learning more accessible to educators. Over the last decade, deep learning has become a powerful method for creating sentiment analysis models.[96,97] Deep learning builds on multi-layer artificial neural networks trained to perform the sentiment classification task. In short, deep learning mimics how the human brain works. Given enough examples, an artificial neural network can identify distinct features in text and use them to differentiate between various sentiments carried by, in this case, reflective writings. Once the model is trained, it can infer the sentiment of reflective writings not presented during the training process with a certain level of accuracy. When applied to education, deep learning can allow for more personalized learning approaches and enable educators to tailor learning pathways to individual students. Deep learning environments can also intelligently analyze data across all personalized training instances to recommend improvements and highlight inefficiencies that would not be possible otherwise.[98] For example, applying deep learning approaches to this dataset on reflections could generate new clinical vignettes or cases related to specific organs that could be used in teaching and assessment. It might also be used to predict future behaviors of students (academic performance or professionalism).

Limitations of the study

This study is not without limitations. To encourage personal and honest reflective writings, no demographic data on the students was collected; however, this may have limited the results. While previous work suggests that women and men have similar emotional responses to dissection,[99] differences have been reported concerning the need for detached concern[74] and person centeredness.[26,80] Additionally, the lexicon used for this data analysis was built for English syntax; it is unclear if the data would change for individuals whose primary language was not English. Furthermore, previous experience with dissection, which was not recorded, may also have influenced students' emotions, especially toward their donors. Future studies might gather more information about students to classify the data further.

The use of natural language processing for analysis also has some inherent restrictions. Although the selected approach overcomes some of the challenges associated with this method,[60] this study had limitations inherent to the sentiment analysis approach. The collected reflective writings were relatively short, 35 words on average. Although sentiment analysis is often applied to social media messages, such as tweets that are 280 characters long,[100] a few polarized words can decide the sentiment of the whole reflection, especially if the remaining sentences have neutral polarity. Requiring a minimum reflection length may help minimize this possibility. Additionally, the neighborhood where valence shifters was searched for was four words before and two after a polarized word. Sensitivity analysis could generate insight into the most optimal text cluster size for students' reflective writings.

Another constraint is the use of general-purpose lexicons created by crowdsourcing polarity assignment. As with any natural language processing, that selection significantly impacted the results. It has to be mentioned that there are also several biomedical domain-related lexicons.[101] Medical WordNet (MWN)[102] and WordNet of Medical Events (WME)[103,104] are two lexicons based on WordNet, which is a domain-free extensive dictionary developed at Princeton University.[105] MWN consists of medically relevant terms used by and intelligible to non-expert subjects and supplemented by natural-language sentences that are designed to provide medically validated contexts. MWN primarily focuses on relationships between sets of synonyms, does not provide polarity, and is not publicly available yet. WME, in its 2.0 version, concentrates on conveying information, such as event definition and polarity, to medical experts and laypeople. Its main objective is to support expert systems in extracting medically relevant insights to identify diseases and their relationships to various populations.[106] However, neither MWN nor WME have been constructed to reflect emotions while becoming a medical professional and being exposed to the humanistic side of the profession. The existence of such a lexicon is a prerequisite to furthering the use of stories and reflective writings in medical education.

Analogically, it also applies to the lexicon-based emotion classification. In the EmoLex, used for this analysis, words were assigned one out of eight basic emotions (anger, anticipation, disgust, fear, joy, sadness, surprise, trust). Recent studies showed a different model of emotions, such as anxiety, apprehension, confusion, fear, uncertainty, upset, calmness, curiosity, enthusiasm, excitement, fascination, and interest,[107] which may also be applicable to a gross anatomy course setting. Since the students had not reported sentiment or emotions accompanying the reflection process, it is also challenging to verify and validate if the reported emotions were actually felt by the students. A mixed research design, combining qualitative and quantitative methods, might provide further insights into the current results.

CONCLUSIONS

The emotionally charged experience of anatomical dissection provides a rich opportunity for students to reflect thoughtfully on their experiences and those of their "silent" teachers. The automated analysis of such writings is a valuable tool for student feedback, educators, and researchers, especially considering the importance of reflective writing to foster professional development. Our data revealed an overall positive sentiment and shared emotional responses in self-reflections compared with reflections about donors. These similarities suggest a common view of humanization and person-centeredness of the donor. While still a maturing tactic, assessing reflective writings through natural language processing is a promising method for uncovering themes, the connectedness of student responses, and determining what areas warrant future investigations. In this study, natural language processing on student reflective writings successfully presented information in an easy-to-understand manner about the sentiment and emotions experienced while writing about anatomical contributions to self and donors' lives. Given the pervasiveness of reflective writings in anatomy, adopting a natural language processing approach to analysis could provide a rich source of new information related to students' previously undiscovered experiences and competencies. Not without impact will be the concurrent development of machine learning approaches to sentiment analysis and domain-specific lexicons, which will increase the accuracy of the results.

ACKNOWLEDGMENTS

The authors acknowledge the anatomical donors, donor families, and donor programs for their pivotal contributions to anatomical education.

CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest or competing interests.

ORCID

Carrie A. Elzie https://orcid.org/0000-0001-5017-3936

REFERENCES

1. Schön DA. *The reflective practitioner: how professionals think in action*. 1st ed. New York: Basic Books, Inc.; 1983. 384 p.
2. Beach MC, Inui T. Relationship-centered care. A constructive reframing. *J Gen Intern Med*. 2006;*21*:S3–8.
3. Charon R. Narrative medicine: a model for empathy, reflection, profession and trust. *JAMA*. 2001;*286*:1897–902.
4. Kolb DA. *Experiential learning: experience as the source of learning and development*. 1st ed. Englewood Cliffs, NJ: Prentice-Hall; 1984 416 p.
5. Smith CS, Irby DM. The roles of experience and reflection in ambulatory care education. *Acad Med*. 1997;*72*:32–5.
6. Westberg J, Jason H. Fostering learners' reflection and self-assessment. *Fam Med*. 1994;*26*:278–82.
7. Mann K, Gordon J, MacLeod A. Reflection and reflective practice in health professions education: a systematic review. *Adv Health Sci Educ Theory Pract*. 2009;*14*:595–621.
8. Wald HS, Reis SP, Monroe AD, Borkan JM. 'The loss of my elderly patient': interactive reflective writing to support medical students' rites of passage. *Med Teach*. 2010;*32*:e178–84.
9. Sandars J. The use of reflection in medical education: AMEE guide no. 44. *Med Teach*. 2009;*31*:685–95.
10. Lie D, Shapiro J, Cohn F, Najm W. Reflective practice enriches clerkship students' cross-cultural experiences. *J Gen Intern Med*. 2010;*25*:S119–25.
11. Mamede S, Schmidt HG, Penaforte JC. Effects of reflective practice on the accuracy of medical diagnoses. *Med Educ*. 2008;*42*:48–75.
12. DasGupta S, Charon R. Personal illness narratives: using reflective writing to teach empathy. *Acad Med*. 2004;*79*:351–6.
13. Sargeant JM, Mann KV, Vander Vleuten CP, Metsemakers JF. Reflection: a link between receiving and using assessment feedback. *Adv Health Sci Educ Theory Pract*. 2009;*14*:399–410.
14. Rabow MW, McPhee SJ. Doctoring to heal: fostering well-being among physicians using personal reflection. *West J Med*. 2001;*174*:66–9.
15. Gustavson N. The effect of human dissection on first-year students and implications for the doctor-patient relationship. *J Med Educ*. 1988;*63*:62–4.
16. Hildebrandt S. Thoughts on practical core elements of an ethical anatomical education. *Clin Anat*. 2016;*27*:839–43.
17. Lachman N, Pawlina W. Integrating professionalism in early medical education: the theory and application of reflective practice in the anatomy curriculum. *Clin Anat*. 2006;*19*:456–60.
18. Wisner BP, Cravens GT, Holley KE, Carmichael SW, Pawlina W. Autopsy report: a valuable tool for developing observational skills. *Clin Anat*. 2004;*17*:197.
19. Crow SM, O'Donoghue D, Vannatta JB, Thompson BM. Meeting the family: promoting humanism in gross anatomy. *Teach Learn Med*. 2012;*24*:49–54.
20. Shapiro J, Nguyen VP, Mourra S, Boker JR, Ross M, Thai TM, et al. Relationship of creative projects in anatomy to medical student professionalism, test performance and stress: an exploratory study. *BMC Med Educ*. 2009;*9*:65.
21. Lazarus L, Sookrajh R, Satyapal KS. Reflective journals: unmasking student perceptions of anatomical education. *Folia Morphol (Warsz)*. 2017;*76*:506–18.
22. Coulehan JL, Williams PC, Landis D, Naser C. The first patient: reflections and stories about the anatomy cadaver. *Teach Learn Med*. 1995;*7*:61–6.
23. Chan LK. Promoting reflection in the laboratory. In: Chan LK, Pawlina W, editors. *Teaching anatomy: a practical guide*. 1st ed. New York: Springer International Publishing; 2015. p. 203–11.
24. Abrams MP, Eckert T, Topping D, Daly KD. Reflective writing on the cadaveric dissection experience: an effective tool to assess the impact of dissection on learning of anatomy, humanism, empathy, well-being, and professional identity formation in medical students. *Anat Sci Educ*. 2021;*14*:658–65.
25. Wu A, Yu AC, Chang CS, Goel R, Sagoo MG, Sakurai T, et al. Initiating students' reflections on life's passing in the anatomy course—an international observation at 14 universities. *Ann Anat*. 2021;*237* : 151741.
26. Goss AL, Viswanathan VB, DeLisser HM. Not just a specimen: a qualitative study of emotion, morality, and professionalism in one medical school gross anatomy laboratory. *Anat Sci Educ*. 2019;*12*:349–59.

27. Hatton N, Smith D. Reflection in teacher education: towards definition and implementation. *Teach Teach Educ*. 1995;*11*:33–49.
28. Kember D. Determining the level of reflective thinking from students' written journals using a coding scheme based on the work of Mezirow. *Int J Lifelong Educ*. 1999;*18*:18–30.
29. Williams RM, Sundelin G, Foster-Seargeant E, Norman GR. Assessing the reliability of grading reflective journal writing. *J Phys Ther Educ*. 2000;*14*:23–6.
30. Wong FK, Kember D, Chung LY, Yan L. Assessing the level of reflection from reflective journals. *J Adv Nurs*. 1995;*22*:48–57.
31. Wittich CM, Pawlina W, Drake RL, Szostek JH, Reed DA, Lachman N, et al. Validation of a method for measuring medical students' critical reflections on professionalism in gross anatomy. *Anat Sci Educ*. 2013;*6*:232–8.
32. Aukes LC, Geertsma J, Cohen-Schotanus J, Zwierstra RP, Slaets JP. The development of a scale to measure personal reflection in medical practice and education. *Med Teach*. 2007;*29*:177–82.
33. Wald HS, Borkan JM, Taylor JS, Anthony D, Reis SP. Fostering and evaluating reflective capacity in medical education: developing the REFLECT rubric for assessing reflective writing. *Acad Med*. 2012;*87*:41–50.
34. Learman LA, Autry AM, O'Sullivan P. Reliability and validity of reflection exercises for obstetrics and gynecology residents. *Am J Obstet Gynecol*. 2008;*198*:461.e1–8.
35. Préfontaine C, Gaboury I, Corriveau H, Beauchamp J, Lemire C, April MJ. Assessment tools for reflection in healthcare learners: a scoping review. *Med Teach*. 2022;*44*:394–400.
36. McMullan M. Students' perceptions on the use of portfolios in pre-registration nursing education: a questionnaire survey. *Int J Nurs Stud*. 2006;*43*:333–43.
37. Pee B, Woodman T, Fry H, Davenport ES. Appraising and assessing reflection in students' writing on a structured worksheet. *Med Educ*. 2002;*36*:575–85.
38. Kumar E. *Natural language processing*. 1st ed. New Delhi: I.K. International Pvt. Ltd; 2011 224 p.
39. Dattu BS, Gore DV. A survey on sentiment analysis on twitter data using different techniques. *Int J Comput Inf*. 2015;*6*:5358–62.
40. Denny JC, Spickard A, Speltz PJ, Porier R, Rosensteil DE, Powers JS. Using natural language processing to provide personalized learning opportunities from trainee clinical notes. *J Biomed Inform*. 2015;*56*:292–9.
41. Denny JC, Bastarache L, Sastre EA, Spickard A. Tracking medical students' clinical experiences using natural language processing. *J Biomed Inform*. 2009;*42*:781–9.
42. Chen Y, Wrenn J, Xu H, Spickard A, Habermann R, Powers J, et al. Automated assessment of medical students' clinical exposures according to AAMC geriatric competencies. *AMIA Annu Symp Proc*. 2014;*2014*:375–84.
43. Andrews J, Chartash D, Hay S. Gender bias in resident evaluations: natural language processing and competency evaluation. *Med Educ*. 2021;*55*:1383–7.
44. Liu B. Sentiment analysis and opinion mining. *Synth Lect Hum Lang Technol*. 2012;*5*:1–67.
45. Zhang L, Liu B. Sentiment analysis and opinion mining. In: Sammut C, Webb GI, editors. *Encyclopedia of machine learning and data mining*. 2nd ed. New York: Springer Science+Business Media, LLC; 2017. p. 1152–61.
46. Rice DR, Rice GN, Baugh C, Cloutier RL. Correlation of narrative evaluations to clerkship grades using statistical sentiment analysis. *Med Sci Ed*. 2022;*2022*:1–7.
47. Zhang R, Pakhomov S, Gladding S, Aylward M, Borman-Shoap E, Melton GB. Automated assessment of medial training evaluation text. *AMIA Annu Symp Proc*. 2012;*2012*:1459–68.
48. Medhat W, Hassan A, Korashy H. Sentiment analysis algorithms and applications: a survey. *Ain Shams Eng J*. 2014;*5*:1093–113.
49. Balakrishnan V, Martin MC, Kaur W, Javed A. A comparative analysis of detection mechanisms for emotion detection. *J Phys Conf Ser*. 2019;*1339* : 012016.
50. Lin CW, Lin MJ, Wen CC, Chu SY. A word-count approach to analyze linguistic patterns in the reflective writings of medical students. *Med Educ Online*. 2016;*21*:29522.
51. Lebowitz A, Kotani K, Matsuyama Y, Matsumura M. 2020. Using text mining to analyze reflective essays from Japanese medical students after rural community placement. *BMC Med Educ*. *20*:1–7.
52. Arráez-Aybar LA, Castaño-Collado G, Casado-Morales MI. Dissection as a modulator of emotional attitudes and reactions of future health professionals. *Med Educ*. 2008;*42*:563–71.
53. Shalev A, Nathan H. 'Medical students' stress reactions to dissections. *Isr J Psychiatry Relat Sci*. 1985;*22*:121–33.
54. Elzie CA, Tenaglia HA, Eagan M, Rechowicz K. Development of a cinematographic life story of anatomical donor to generate student appreciation and responsibility. *FASEB J*. 2022;*36*:S1.
55. Aggarwal CC. *Machine learning for text*. 1st ed. Cham: Springer International Publishing, AG; 2018 493 p.
56. Feuerriegel S, Proellochs N. *Package 'SentimentAnalysis.'* 2021st ed. San Francisco, CA: GitHub, Inc.; 2021. 51 p. Available from: https://cran.r-project.org/web/packages/SentimentAnalysis/SentimentAnalysis.pdf. Accessed 20 Jul 2022.
57. Rinker T. *Trinker/lexicon*. San Francisco, CA: GitHub, Inc.; 2021a. Available from: https://github.com/trinker/lexicon. Accessed 20 Jul 2022.
58. Ripley BD. The R project in statistical computing. *MSOR Connect*. 2001;*1*:23–5.
59. Nandwani P, Verma R. A review of sentiment analysis and emotion detection from text. *Soc Netw Anal Min*. 2021;*11*:81.
60. Rinker T. *Package 'sentimentr.'* 2021st ed. San Francisco, CA: GitHub, Inc.; 2021b. 62 p. Available from: https://cran.r-project.org/web/packages/sentimentr/sentimentr.pdf. Accessed 20 Jul 2022.
61. Polanyi L, Zaenen A. Contextual valence shifters. Chapter X. AAAI Spring Symposium on Attitude. 2006.
62. Rinker T. *Trinker/sentimentr*. San Francisco, CA: GitHub, Inc.; 2021c. Available from: https://github.com/trinker/sentimentr. Accessed 20 Jul 2022.
63. Jockers ML. *Syuzhet: extract sentiment and plot arcs from text*. Ottawa, Ontario: National Research Council Canada; 2020. Available from: https://rdrr.io/cran/syuzhet/. Accessed 20 Jul 2022.
64. Hu M, Liu B. Mining and summarizing customer reviews. *Proceedings of the tenth ACM Special Interest Group on Knowledge Discovery and Data Mining International Conference on Knowledge Discovery and Data Mining; Seattle, WA, 2004 August 22–25*. New York: Association for Computing Machinery; 2004. p. 168–77.
65. Mohammad SM, Turney PD. Crowdsourcing a word–emotion association lexicon. *Comput Intell*. 2013;*29*:436–65.
66. Plutchik R. A general psychoevolutionary theory of emotion. In: Plutchik R, Kellerman H, editors. *Emotion: theory, research and experience. Volume 1. Theories of emotion*. 1st ed. New York: Academic Press; 1980. p. 3–33.
67. Olejaz M. When the dead teach: exploring the postvital life of cadavers in Danish dissection labs. *Med Anthropol Theor*. 2017;*4*:125–49.
68. Tseng WT, Lin YP. "Detached concern" of medical students in a cadaver dissection course: a phenomenological study. *Anat Sci Educ*. 2016;*9*:265–71.
69. Shiozawa T, Glauben M, Banzhaf M, Griewatz J, Hirt B, Zipfel S, et al. An insight into professional identity formation: qualitative analyses of two reflection interventions during the dissection course. *Anat Sci Educ*. 2020;*13*:320–32.

70. Dinsmore CE, Daugherty S, Zeitz HJ. Student responses to the gross anatomy laboratory in a medical curriculum. *Clin Anat*. 2001;*14*:231–6.
71. McGarvey MA, Farrell T, Conroy RM, Kandiah S, Monkhouse WS. Dissection: a positive experience. *Clin Anat*. 2001;*14*:227–30.
72. O'Carroll RE, Whiten S, Jackson D, Sinclair DW. Assessing the emotional impact of cadaver dissection on medical students. *Med Educ*. 2002;*36*:550–4.
73. Penney JC. Reactions of medical students to dissection. *J Med Educ*. 1985;*60*:58–60.
74. Dickinson GE, Lancaster CJ, Winfield IC, Reece EF, Colthorpe CA. Detached concern and death anxiety of first-year medical students: before and after the gross anatomy course. *Clin Anat*. 1997;*10*:201–7.
75. Druce M, Johnson MH. Human dissection and attitudes of preclinical students to death and bereavement. *Clin Anat*. 1994;*7*:42–9.
76. Evans EJ, Fitzgibbon GH. The dissecting room: reactions of first year medical students. *Clin Anat*. 1992;*5*:311–20.
77. Finkelstein P, Mathers L. Post-traumatic stress among medical students in the anatomy dissection laboratory. *Clin Anat*. 1990;*3*:219–26.
78. Horne DJ, Tiller JW, Eizenberg N, Tashevska M, Biddle N. Reactions of first-year medical students to their initial encounter with a cadaver in the dissecting room. *Acad Med*. 1990;*65*:645–6.
79. Quince TA, Barclay SI, Spear M, Parker RA, Wood DF. Student attitudes toward cadaveric dissection at a UK medical school. *Anat Sci Educ*. 2011;*4*:200–7.
80. Hafferty FW. *Into the valley: death and the socialization of medical students*. 1st ed. New Haven, CT: Yale University Press; 1991. p. 234.
81. Lempp HK. Perceptions of dissection by students in one medical school: beyond learning about anatomy. A Qualitative Study. *Med Educ*. 2005;*39*:318–25.
82. Sinclair S. *Making doctors: an institutional apprenticeship*. 1st ed. New York: Berg; 1997. 320 p.
83. Lee YS, Ryu Y, Jung WM, Kim J, Lee T, Chae Y. Understanding mind–body interaction from the perspective of east Asian medicine. *Evid Base Compl Alternative Med*. 2017;*2017* : 7618419.
84. Fredrickson BL. The broaden–and–build theory of positive emotions. *Phil Trans R Soc Lond B*. 2004;*1449*:1367–77.
85. Chiou RJ, Tsai PF, Han DY. Effects of a "silent mentor" initiation ceremony and dissection on medical students' humanity and learning. *BMC Res Notes*. 2017;*10*:483.
86. Keltner D, Oatley K, Jenkins JM. *Understanding emotions*. 4th ed. Hoboken, NJ: Wiley; 2018 544 p.
87. LeBlanc VR, McConnell MM, Monteiro SD. Predictable chaos: a review of the effects of emotions on attention, memory and decision making. *Adv Health Sci Educ Theory Pract*. 2015;*20*:265–82.
88. Pekrun R, Goetz T, Titz W, Perry RP. Academic emotions in students' self-regulated learning and achievement: a program of qualitative and quantitative research. *Educ Psychol*. 2002;*37*:91–105.
89. Bandura A. *Self-efficacy: the exercise of control*. 1st ed. New York: W.H. Freeman and Co; 1997. p. 604.
90. Dornan T, Pearson E, Carson P, Helmich E, Bundy C. Emotions and identity in the figured world of becoming a doctor. *Med Educ*. 2015;*49*:174–85.
91. Helmich E, Bolhuis S, Laan R, Dornan T, Koopmans R. Medical students' emotional development in early clinical experience: a model. *Adv Health Sci Educ Theory Pract*. 2014;*19*:347–59.
92. Shen L, Wang M, Shen R. Affective e-learning: using "Emotional" data to improve learning in pervasive learning environment. *Educ Technol Soc*. 2009;*12*:176–89.
93. Pekrun R. The impact of emotions on learning and achievement: towards a theory of cognitive/motivational mediators. *Appl Psychol*. 1992;*41*:359–76.
94. Seli P, Wammes JD, Risko EF, Smilek D. On the relation between motivation and retention in educational contexts: the role of intentional and unintentional mind wandering. *Psychon Bull Rev*. 2016;*23*:1280–7.
95. Grochowski CO, Cartmill M, Reiter J, Spaulding J, Haviland J, Valea F, et al. Anxiety in first year medical students taking gross anatomy. *Clin Anat*. 2014;*27*:835–8.
96. Agarwal B, Nayak R, Mittal N, Patnaik S, editors. *Deep learning-based approaches for sentiment analysis*. 1st ed. Singapore: Springer Nature Singapore Pte Ltd; 2020 332 p.
97. Zhang L, Wang S, Liu B. Deep learning for sentiment analysis: a survey. *Wiley Interdiscipl Rev Data Min Knowl Discov*. 2018;*8* : e1253.
98. Muniasamy A, Alasiry A. Deep learning: the impact on future elearning. *Int J Emerg Technol*. 2020;*15*:188–99.
99. Snelling J, Sahai A, Ellis H. Attitudes of medical and dental students to dissection. *Clin Anat*. 2003;*16*:165–72.
100. Agarwal A, Xie B, Vovsha I, Rambow O, Passonneau RJ. Sentiment analysis of Twitter data. In: Nagarajan M, Gamon M, editors. *Proceedings of the Workshop on Language in Social Media (LSM 2011): Portland, Oregon, 2011 June 23*. Stroudsburg, PA: Association for Computational Linguistics (ACL); 2011. p. 30–8.
101. Satapathy R, Cambria E, Hussain A. *Sentiment analysis in the biomedical domain: techniques, tools, and applications*. 1st ed. Cham: Springer International Publishing AG; 2017. 158 p.
102. Smith B, Fellbaum C. Medical WordNet: a new methodology for the construction and validation of information resources for consumer health. *Proceedings of the 20th International Conference on Computational Linguistics (COLING '04); Geneva Switzerland. 2004 August 23–27*. Stroudsburg, PA: Association for Computational Linguistics; 2004. p. 371–82.
103. Mondal A, Chaturvedi I, Das D, Bajpai R, Bandyopadhyay S. Lexical resource for medical events: a polarity based approach. *Proceedings of the IEEE 15th International Conference on Data Mining Workshop (ICDMW 2015); Atlantic City, NJ, 2015 November 14–17*. Piscataway, NJ: Institute of Electrical and Electronics Engineers, Inc.; 2015. p. 1302–9.
104. Mondal A, Das D, Cambria E, Bandyopadhyay S. WME: sense, polarity and affinity based concept resource for medical events. In: Fellbaum C, Vossen P, Mititelu VB, Forascu C, editors. *Proceedings of the 8th Global WordNet Conference (GWC 2016); Bucharest, Romania, 2016 January 27–30*. Amsterdam: Global WordNet Association; 2016. p. 243–8.
105. Fellbaum C. WordNet. In: Poli R, Healy M, Kameas A, editors. *Theory and applications of ontology: computer applications*. 1st ed. New York: Springer Science+Business Media, M.V.; 2010. p. 231–43.
106. Mondal A, Cambria E, Das D, Hussain A, Bandyopadhyay S. Relation extraction of medical concepts using categorization and sentiment analysis. *Cognit Comput*. 2018;*10*:670–85.
107. Chiou RJ, Tsai PF, Han DY. Impacts of a gross anatomy laboratory course on medical students' emotional reactions in Taiwan: the role of high-level emotions. *BMC Med Educ*. 2021;*21*:489.

How to cite this article: Rechowicz, K. J. & Elzie, C. A. (2023). The use of artificial intelligence to detect students' sentiments and emotions in gross anatomy reflections. *Anatomical Sciences Education*, *00*, 1–13. https://doi.org/10.1002/ase.2273

APPENDIX 1

Custom dictionary:

back
arms
hands
gluteals
knees
feet
heart
lungs
gastrointestinal
system
skin
pelvis
brain
face
eyes
ears
nose
mouth
other
donor
donor's

AUTHOR BIOGRAPHIES

Krzysztof J. Rechowicz, Ph.D., is a research assistant professor in the Virginia Modeling, Analysis, and Simulation Center of Old Dominion University in Suffolk, Virginia. His research interest is in applying machine learning, virtual reality, and human–computer interfaces to create inclusive and accessible environments for people of all abilities.

Carrie A. Elzie, Ph.D., M.A., is a professor of Cell Systems and Anatomy and Medical Education at the University of Texas Health Science Center San Antonio. Her research interests include humanistic approaches such as narrative medicine and virtual embodiment to anatomical education, as well as, teaching innovations such as gamification, art, and adaptive learning.