

## **Sentiment Analysis on Autism Content in College-Level Textbooks**

Alexander Maksiaev

Cognitive Science Distinguished Majors Program: University of Virginia

Vikram Jaswal, Ph.D., Department of Psychology

Brad Pasanek, Ph.D., Department of English

May 2, 2024

### **Abstract**

Both autistic and transgender people face social stigma. The way in which college-level course textbooks describe autistic and transgender people may contribute to or help reduce stigma towards these individuals. The type of language used plays a large role — more specifically, its valence, or how positive or negative the language sounds. For example, autistic people can be characterized as exhibiting “problematic behavior” and social and motor “deficits”, or they can be characterized as having “normal variations” in cognitive functioning that are different but not deficient. Transgender people may be characterized as living in a way that “does not conform” to their biological sex, or they may be recognized as “naturally expressing” their gender identities. While charged language may be marked by hand, sentiment analysis is a well-established computational method for determining sentiment (usually positive or negative valence) within text – one that eliminates the need for manual coding for sentiment in a text and is therefore a much faster and consistent method of evaluating it. Using sentiment analysis, 296 textbooks from eight fields were tested for how positive or negative the autism and transgender content in them was. Both lexicon-based (a VADER-like dictionary and NLTK’s SentiWordNet module) and document-based (ChatGPT API) sentiment analysis approaches were used. Both the individualized VADER-like dictionary and ChatGPT API’s sentiment analysis for autism search terms showed a significant positive trend across publication years for all textbooks, suggesting a general lessening of stigma in college-level education across time for autism — but not for transgender people.

## **Sentiment Analysis on Autism Content in College-Level Textbooks**

The goal of this project is to perform exploratory sentiment analysis on a set of textbooks with content about autism within them. In this opening section, first, a brief definition and history of autism will be given. Then, the comparison case, transgender people, will also be discussed. Afterwards, the concept of textbooks in education will be introduced, along with how the language used within textbooks affects perception of outgroups, and some examples of sentiment within textbooks concerning autism and transgender people. Finally, sentiment analysis will be introduced and defined, along with several different sentiment analysis methods.

### **Autism**

Autism spectrum disorder (ASD) is a neurodevelopmental disorder categorized by difficulties with social communication and restricted, repetitive behaviors and interests (Hodges et al., 2020). Originally thought to be a form of childhood schizophrenia, both autism as a diagnosis and autism as a social concept have radically changed meaning over the years (Evans, 2013).

Leo Kanner (1943) published one of the first papers written describing what today is now known as autism. Kanner described 11 children who all showed similar symptoms – most prominently, an “inability to relate themselves in the ordinary way to people and situations from the beginning of life”, and an inherent “extreme autistic aloneness” in which they deliberately ignored their social surroundings (1943, p. 242). Kanner also described the group of children as having “few really warmhearted fathers and mothers,” although he also insisted that the condition itself was “innate” (1943, p. 250). Kanner further suggested that these children may have previously been misdiagnosed as being “feeble-minded” or “schizophrenic”, which suggests a social concept of autism which at the time did not differentiate from the social concept of intellectual disability or psychosis (1943, p. 242).

Indeed, autism was not differentiated from the aforementioned two labels, even after Kanner’s landmark paper. In 1948, *Time* magazine published the article “Frosted Children”,

which was the first popular depiction of autism and its possible cause (“Frosted Children”). *Time* described Kanner’s patients as “diaper schizoids”, and while the article did use the word “autism”, it also repeatedly and explicitly stated that the children were either developing or already had “schizophrenia”, making no distinction between the two diagnoses (“Frosted Children”, 1948). This was also the first instance of the belief that autism was being caused by “refrigerator mothers” — mothers who were mechanical, educated, and devoid of warm parental feelings towards their children (Silberman, 2015, p. 198).

It wasn’t until 1980 when autism became formally recognized as its own diagnosis, in the third edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-III) (Rosen et al., 2021). Autism’s popularity as a research subject sky-rocketed with its inclusion in the DSM; however, the concept of autism as a social category continued to be pathologized and its prognosis believed to be bleak, with multiple longitudinal studies showing relatively few autistic adults achieving full independence (Rosen et al., 2021).

Following the addition of Asperger’s syndrome in the DSM-IV in 1994, a different social picture of autism began to emerge (Barahona-Corrêa & Filipe, 2016). Children with Asperger’s syndrome held similar qualities to those with autism, but ultimately were “higher-functioning” than those given the autism label (Barahona-Corrêa & Filipe, 2016). Asperger’s syndrome patients were still considered mentally ill, but the label itself was much less stigmatizing, as Asperger’s quickly became public shorthand for a person who had symptoms that looked like autism but was “intelligent and educable” (Grinker, 2020). Asperger’s syndrome became extinct as a diagnostic category in the DSM-5 in favor of Autism Spectrum Disorder (ASD), essentially blending the two diagnoses together (Barahona-Corrêa & Filipe, 2016). Additionally, the discovery of genetics’ role in autism with monozygotic twins decreased mother blame and deprecated the “refrigerator mother” stereotype (Rosen et al., 2021). The combination of these developments led to autism becoming “increasingly desirable as a replacement for other diagnoses,” and thereby reduced overall stigma in the public eye (Grinker, 2020).

The prevalence of ASD diagnoses generally have significantly increased in recent years worldwide, as has public awareness (Zeidan et al., 2022). Along with public awareness of autism generally, awareness and public portrayals of the concept of neurodiversity has also increased (Lewin & Akhtar, 2020). Neurodiversity is the concept that neurological difference is, like other forms of diversity, a natural and necessary variation of the human experience, the same way biodiversity is natural and necessary in the biosphere (Mcgee, 2012). Neurodiversity and its increasing presence in popular culture reflects an overall social change in how autism is perceived — no longer is the popular perception that autistic people are all loner imbeciles; rather, the concept of neurodiversity shows that autism is a different, and not lesser, form of cognition (Lewin & Akhtar, 2020). This change from different as deficient to different as desirable is also being reflected in fictional and non-fictional media representations of autistic people, especially in literature, social media, and video games (Mittmann et al., 2023).

However, awareness alone does not eliminate bias (Redfield, 2017, p. 295). Although sentiment towards autism is changing for the better, there is still room for improvement. For example, neurotypical (non-neurodivergent) college students' positive feelings towards autistic people may have increased over the years, but this is mostly reflecting a change from negative sentiment to more neutral (Lewin & Akhtar, 2020). Stigma against autistic people still exists, often with debilitating consequences — only 38.8% of autistic undergraduate students graduate from their respective institutions, with stigma often being cited as a significant barrier to success, even when these students are academically capable (Kitchin & Karlin, 2021). Campaigns designed to educate neurotypical college students about autism have also seen mixed results in terms of changes in stigmatizing attitudes (Kitchin & Karlin, 2021). Therefore, awareness alone seems to be inadequate.

### **Transgender People**

An interesting comparison case to consider of another stigmatized group of people would be transgender people. More people today are publicly identifying as transgender, and

public awareness about being transgender has increased in recent years (Mocarski et al., 2019). Socially and culturally, stigma against transgender people has decreased — although it has not completely been removed, and transgender people still face large amounts of discrimination today (Verbeek et al., 2020). This phenomenon of increased saliency and acceptance parallels that of autistic people, and so it would be interesting to compare the two in terms of overall public sentiment.

### **Language and Education**

Language matters. Valence — that is, how positively or negatively a subject is perceived or described — influences stigma and the likelihood of discriminatory behavior (Turnock et al., 2022). When undergraduate students were asked to identify traits associated with autism, 80% of the most common traits named were negative, illustrating negative stereotypes (Turnock et al., 2022). Medical professionals who express negative attitudes towards transgender people are also more likely to show discriminatory behavior (Cutillas-Fernández et al., 2023). How positive or negative a person's thoughts are towards a group influences how that person treats them.

Textbooks in education reflect current social sentiments; for example, conservative areas tend to purchase textbooks with less female and black representation (Lucy et al., 2020). K-12 textbook purchases in the United States are often decided at the state level, and are influenced by political attitudes of the committees which choose them (Schmidt, 2022). At the same time, textbooks also have the power to change sentiment — students' views on people different from themselves reflect the language of textbooks used in their instruction (Lucy et al., 2020). The way autistic and transgender people are written about in textbooks may both reflect and influence popular perception. Therefore, it can be illuminating to investigate sentiments towards autistic and transgender people in textbooks in order to track overall public sentiment. To give some examples, below are excerpts from published textbooks which talk about either autism or transgender people positively or negatively.

A positive depiction of autism is as follows:

Neurodiversity is the concept that autism, as well as other disabilities, are normal variations of functioning of the human mind, not pathological conditions or deficits. The neurodiversity movement calls for appropriate supports and inclusiveness for all individuals and may refer to those without ASD or other disabilities as neurotypicals. (Friend, 2018, p. 297)

This quote presents a positive valence in its emphasis on autism *not* being a pathology or something to be “fixed”. It advocates for inclusion and understanding for autistic and other neurodivergent people, and says that these disabilities are “normal variations” of human functioning as opposed to abnormal or pathological.

A contrasting negative characterization of autism is as follows: “[Autism] is a developmental disorder that involves a wide range of problematic behaviors including deficits in language, and perceptual and motor development; defective reality testing; and an inability to function in social situations” (Hooley, 2006, p. 575). This quote presents a negative valence in its description of autism in its charged vocabulary: “disorder”, “problematic”, “deficits”, “defective”, and “inability” are all terms with negative connotations. Autism is here described as a collection of deficits, with all autistic behaviors evaluated as being “problematic”, without any redeeming qualities.

For still further contrast, consider the following positive depiction of transgender people and gender identity:

For transgender people, wearing the clothes of the gender with which they identify is a natural expression of their gender identity. Trans people do not typically wear these clothes for purposes of erotic arousal or as a parody. They are only “cross” dressing with respect to their natal (birth) sex, not their gender identity or (if they have physically transitioned) their current sex. (LeVay et al., 2015, p. 406)

This quote communicates a positive sentiment because it dispels myths cisgender (non-transgender) people may believe about transgender people. By means of negation, the language explicitly states that a transgender person's clothing choices are most informed by considerations of comfort and not for sexual gratification or self-amusement, as well as clarifying that transgender people do not “crossdress” because crossdressing would mean wearing clothes of a different gender identity. Additionally, the quote emphasizes that wearing these clothes is a “natural” expression of gender identity, with the word “natural” having a positive connotation.

To complete the four-square comparison, consider a final negative depiction of transgender people and gender identity:

Transgenderism includes people “whose appearance and/ or behaviors do not conform to traditional gender roles” (Crooks & Baur, 2008, p. 62), “who live full- or part-time in the other gender's role” — that is, their non biological gender — “and derive psychosocial comfort in doing so” (Carroll, 2007, p. 86). (Zastrow et al., 2010, p. 375)

Although the negative valences in this passage may be subtle, the quote above is informed by and communicates prejudiced sentiments. First, it uses the word “transgenderism”, which implies that being transgender is a condition that can be changed or an ideology one can disagree with (“Transgenderism”, 2024).<sup>1</sup> Secondly, the phrase “do not conform to traditional gender roles” invalidates transgender identity by implying that the transgender person's gender is their birth sex and they're simply not “conforming” to what is expected, which is further reinforced by the later phrase “non biological gender”, conflating gender with sex. The phrase “other gender's role” also implies that the transgender person is not actually the gender they say they are. Although both passages mention “comfort”, this second passage communicates a subtly negative evaluation.

---

<sup>1</sup> Although the term may have been less obviously offensive in 2010, it was still, even at that time, objected to by trans activists.



With practice, a reader will become attuned to subtle valences. Sentiment, however, is difficult to systematically quantify. Writing is subjective, and different readers may disagree on the perceived emotion of a text. In order to collect sentiment as objectively as possible, it would be ideal for all text analyzed to be read by one careful reader. However, an adequate sample size of text may render this impossible, if the reader is human. In this case, instead of relying on one human reader, computational methods may be employed, such as sentiment analysis.

### **Sentiment Analysis**

Sentiment analysis (also called opinion analysis or opinion mining) is the computational processing of sentiment (opinions, thoughts, feelings) found in text (Wankhade et al., 2022). Sentiment analysis is often used by corporations and governments to analyze sentiments of interest, including product reviews and general public opinion (Wankhade et al., 2022). This analysis normally uses natural language processing (NLP) and machine learning techniques.

Most sentiment analysis involves an evaluation factor and potency (Taboada et al., 2011). One of the most common evaluation factors used to evaluate a text is positive or negative valence (Taboada et al., 2011). Potency is the strength of that evaluation factor — e.g. just how positively or negatively a subject is rated — in regards to that particular text (Taboada et al., 2011). What counts as a “subject” varies, and a subject may be as broad as an entire idea or topic, or it may be as specific as a phrase or individual word (Taboada et al., 2011).

### ***Lexicon-Based Sentiment Analysis***

There are two main approaches to sentiment analysis. The first approach is “lexicon-based”, in which sentiment is constructed via the individual sentiments of individual words or phrases (Taboada et al., 2011). This approach often involves dictionaries which contain semantically meaningful words, “scored” by the potency of their evaluation factor (Taboada et al., 2011). For example, the word “happy” may appear in a lexicon-based sentiment analysis dictionary with a score of “+3” on a scale from -4 (most negative) to +4 (most positive).

The VADER (Valence Aware Dictionary for sEntiment Reasoning) lexicon is a popular dictionary used for lexicon-based sentiment analysis (Hutto & Gilbert, 2014). VADER is an open-source lexicon with around 7,500 words and a number of other lexical features (e.g. emojis) with meaningful valence (Hutto & Gilbert, 2014). All of these words have been individually scored by trained humans on a scale from -4 (most negative) to +4 (most positive) (Hutto & Gilbert, 2014). Using VADER is simple — words in the text that happen to be in the lexicon are given scores according to the lexicon, added together, and are turned into a “normalized, weighted composite score” from -1 (most negative) to +1 (most positive) (Hutto, 2021). Most often, this score is then split into a separate positive score, negative score, and neutral/objective score which together equal 1 (Hutto, 2021). Despite this simplicity, however, it has been shown to be very accurate when analyzing tweets, movie reviews, newspaper articles, and other texts (Hutto & Gilbert, 2014).

SentiWordNet is a lexicon available in a module of the same name in the Python library NLTK (Natural Language ToolKit) (“Sample Usage”, 2023). It has a total of 155,287 words (Kumar et al., 2018). Similarly to VADER, SentiWordNet performs sentiment analysis by giving each word a separate positive, negative, and objective score which together equal 1 (Baccianella et al., 2010). The sentiment scores for these words were generated automatically through semi-supervised learning, and then evaluated for accuracy, in which these ratings were evaluated by comparison to a very small subset of words in the lexicon which had previously been rated by five different humans (Baccianella et al., 2010). SentiWordNet’s accuracy has improved over its different iterations, with the most recent version being SentiWordNet 3.0 (Baccianella et al., 2010). However, despite being a smaller lexicon, VADER has been shown to be more accurate (Al-Shabi, 2020).

### ***Machine-Learning-Based Sentiment Analysis***

The other main approach to sentiment analysis is “machine-learning-based”, in which sentiment is constructed via the overall sentiments of the sentence, paragraph, or even full

document using textual context (Truică et al., 2021). As the name of the approach suggests, machine learning is employed in order to complete this task, often using natural language processing (NLP) methods. Machine-learning based sentiment analysis has traditionally involved both supervised and unsupervised machine learning models (Srivastava et al., 2022). However, more recently neural networks and deep learning have also been employed (Truică et al., 2021).

ChatGPT is a powerful generative AI chatbot built by the company OpenAI (Hetler, 2023). It uses deep learning and transformer neural networks to produce human-like text in a conversational manner, and also uses supervised reinforcement learning to improve responses (Hetler, 2023). It can solve a wide range of complex problems and perform a wide array of tasks, including tasks like sentiment analysis.

Despite ChatGPT's strengths, it has a number of limitations. One of its most troubling limitations is the tendency to spontaneously create false information, also called "hallucinations" (de Wynter et al., 2023). Although its output may have correct English syntax and seemingly make sense semantically, the information it presents may be incorrect or fabricated (de Wynter et al., 2023). For example, when asked to generate medical papers with references, ChatGPT did so, but only 7% of the references made were both accurate and authentic (Bhattacharyya et al., 2023).

ChatGPT also has its own API (Application Programming Interface), where programmers may embed ChatGPT into their own code and use it to solve problems (Brockman et al., 2024). ChatGPT has several versions of its API to choose from, including GPT-3.5 Turbo, GPT-4, and GPT-4 Turbo, the latter of which can analyze images as well as text ("GPT-4 Turbo and GPT-4"). ChatGPT's API works by using tokens — small pieces of words which make up the input and output that ChatGPT receives and gives ("What Are Tokens"). Tokens are required to make ChatGPT run inside of a program, and programmers are required to pay for the tokens they use ("Pricing", 2024).

### **The Current Study**

The present study aims to enlist these three sentiment analysis techniques (two lexicon-based methods and also ChatGPT) in order to score many textbooks from various fields in terms of their sentiment towards autism and transgender subjects. While this project is mostly exploratory, there is one hypothesis: based on the history of autism presented, the prediction is that older textbooks may have more negative autism scores, while newer textbooks may have neutral scores, showing a positive trend over time from negative to neutral.

## Materials and Methods

### Textbooks

Word documents containing excerpts from textbooks were supplied by Dr. Gernsbacher and colleagues at the University of Wisconsin-Madison. These documents were created in advance of the current research. The documents were divided into two corpuses: one containing autism search terms, and the other containing transgender search terms. They were also divided by field, author, and edition. The process in which these documents were created is outlined below. For a list of textbooks used, see Table A1 in the Appendix.

The sub-disciplines or “fields” selected for analysis were as follows: Abnormal Psychology (Abn), Developmental or Lifespan Psychology (Devo), Gender Studies (GS), Human Sexuality (HS), Neuro-/physio-/cognitive psychology (Neuro), Introductory Psychology (Intro), Social Psychology (Socl), and Special Education (Spcl).

### ***Criteria for selecting the most prolific current textbooks***

First, as many textbooks as possible were identified for each of the eight fields, using publishers’ websites, sellers’ websites (e.g. Amazon, Barnes & Noble), publicly available syllabi, etc. Next, textbooks that had published at least three editions with their most recently published edition being “current” were identified. “Current” editions were defined as appearing between 2015 and 2018 (inclusive). The “top ten” textbooks in each field were selected – these were the textbooks with the largest number of editions, i.e. the most “prolific”.

If more than one textbook had the same number of editions, the textbook that was selected was: the textbook whose current edition was published most recently, then the textbook with the most immediate future edition, then the textbook whose first author’s last name came first alphabetically. Only one version of a textbook was selected if there were multiple versions (e.g. AP version, other versions from the same author, etc.), and this was the textbook with the highest rank. If two textbooks within a field were written by the same author (or group of

authors), the textbook selected was the one with the highest rank, and the other lower-ranked same-authored textbooks were omitted.

All selected current textbooks must have been available electronically. This was because searching for key terms (see procedure for searching textbooks for the list) and extracting the needed pages was done electronically. All textbooks were also required to have an index, as did their previous versions.

### ***Procedures for accessing textbooks***

76 current editions were obtained. 72 electronic versions were obtained through VitalSource. One electronic version was obtained through its publisher. One electronic version was obtained through RedShelf. Two electronic versions were obtained through other unknown sources. Two hundred fifty-two previous editions were obtained. 37 electronic versions were obtained through VitalSource, and 215 hard copy versions were obtained through InterLibraryLoan.

### ***Procedure for searching textbooks***

Search terms are words or phrases that were determined to be relevant to autism or transgender topics. Capitalization did not matter.

Autism search terms were as follows: *autism, autism disorder, autism disorders, autistic, autistic disorder, Asperger, Asperger disorder, Asperger syndrome, Asperger's, Asperger's disorder, Asperger's syndrome, ASD, autism spectrum disorder, autism spectrum disorders, autism spectrum disorder (ASD), autism spectrum disorders (ASD).*

Transgender search terms were as follows: *gender, genders, gendered, transgender, transgenders, transgendered [sic.], transgenderism, cisgender, cisgenders, cisgendered [sic.], transsex, transexed, transsexual, transsexuals, transexuated [sic.], transsexualism, transsexuality, intersex, intersexed, intersexual, intersexuals, intersexuated [sic.], intersexualism, intersexuality, genderqueer, genderqueers, genderqueered [sic.], gender queer, gender queers, gender queered [sic.].*

### ***Extracting Textbooks***

If the textbook was a physical copy, pages with search terms were found using the indices. For each unique page listed, the page was scanned and saved as a PDF. If the textbook was an eBook, pages with search terms were found using word searches. For each unique page listed, the page was “saved as PDF” using the print function. Each PDF was manually read by humans and double-checked. Adobe’s optical character recognition made the PDF text readable.

### ***Full text search data***

A full text search was completed for current textbooks, all of which were eBooks. Full text searches were performed by using the search bar to search the entire textbook for each search term. Then, each search result was visited and confirmed that the search term appeared on the page somewhere besides the chapter header (i.e. the body of the text, index entries, glossary entries, figure captions, text boxes, tables, figures, and references). Textbooks without either autism or transgender search terms were omitted from analysis. Overall, 296 unique textbooks were analyzed.

### ***Procedure for excerpting***

Key words in the extracted PDFs were captured and entered into a Word document for each textbook edition. Each textbook edition had an extracting Word document for autism search terms and an extracting Word document for transgender search terms. For each textbook edition, the PDFs were opened and searched using the search function to find any autism or transgender search terms. Any paragraphs containing at least one search term were copy-pasted into the textbook editions’ excerpt document with page numbers added before the paragraph, and each paragraph was separated with a blank line. Text in picture and figure captions, text boxes, tables, figures, footnotes, and glossary entries were also captured – when capturing this content, only the one sentence or line that contained one or more search terms were captured. Search terms in references, citations, indices, or chapter outlines were not

copy-pasted. PDFs were also manually read through to find any search terms that the search function missed, and these instances were also copy-pasted. The quotes were ordered in the order they appeared in the textbook. Typos made by the textbook were not corrected, but computer misinterpretations from copy-pasting were corrected.

Each Word document was saved into a folder named by an author and edition (e.g. one folder named “Abn\_Brown\_04\_eBook\_Excerpts” contained only Word documents which used Brown’s 4th edition Abnormal Psychology textbook, one folder named “Abn\_Brown\_05\_eBook\_Excerpts” contained only Word documents which used Brown’s 5th edition Abnormal Psychology textbook, etc). These folders were each saved into a larger folder representing a common field (e.g. all folders with Abnormal Psychology textbooks were saved into one larger folder, all folders with Social Psychology textbooks were saved into one larger folder, etc). However, for analysis, all textbook Word documents with autism search terms were placed into a single folder for autism search term textbooks, and all textbook Word documents with transgender search terms were placed into a single folder for transgender search term textbooks.

Textbook documents generally had three lines at the beginning of the document to form a header, followed by a break between the header and the text. The first line of the header was the title of the textbook. The second line of the header was the edition and year of the textbook. The third line of the header was the name of the author. However, some textbooks did not have this header. To ease processing, “dummy headers” (e.g. “DUMMY TITLE”) were given to these textbooks without original headers — this was done because exact titles and years for these header-less textbooks were lost. Headers needed to exist because of how the documents were cleaned — analysis began at the fourth line of text in the document, so if a textbook did not have exactly three lines of text at its beginning, then three lines of “dummy text” representing what would have been the original headers were supplied.



## **Sentiment Analysis**

Sentiment analysis was conducted on both the autism and the transgender corpuses for each textbook. Three methods were employed: a) lexicon-based sentiment analysis using the NLTK Python library's module SentiWordNet ("Sample Usage", 2023); b) lexicon-based sentiment analysis using a dictionary built from the autism corpus' unique words; and c) document-level sentiment analysis using OpenAI's ChatGPT API (Katrekar).

For the lexicon-based approaches, each Word document had to be "cleaned" and prepared for analysis. Headers were removed from the beginning of the document, punctuation was removed, all words were converted to lower-case, and stop words were removed from the body of the document.

Stop words are common words in a language with little semantic value ("Dropping Common Terms", 2009). These are words that commonly do not contribute to the meaning of a sentence — for example: "the", "a", "for", etc ("Dropping Common Terms", 2009). Removing stop words allows the program to focus on more meaningful words when conducting lexicon-based sentiment analysis. The words listed in an online English stop words list were removed from all of the Word documents in the autism and transgender textbook corpuses before using the lexicon-based approaches ("Stop Words List"). Additionally, the following words were also removed from the documents used in the lexicon-based approaches: "autism", "autistic", "asperger", "transgender", "transsexuality", "transgenderism", and "transsexual". These were removed because inclusion would severely bias results — for example, the word "autism" was consistently rated negatively (-1.44 on a scale from -4 to +4) by participants whose survey responses formed one of the sentiment dictionaries used, so leaving in the word "autism" in the documents would have negatively biased the overall sentiment for those documents.

### ***VADER-Like Dictionary***

For the autism corpus only, an individualized VADER-like sentiment dictionary was built from participant survey responses, in which 535 undergraduate student participants recruited

from the University of Virginia's Psychology Department Participant Pool rated words on a scale of -4 (most negative sentiment) to +4 (most positive sentiment).

With the exception of textbooks in the fields of Gender Studies (GS) and Human Sexuality (HS), each unique word within all of textbooks in the autism corpus was assembled into a set of words. Only words that had "correct spelling" (using the python library "enchant"), words in the English dictionary (defined as the set of lower-cased words in the Python "english-words" library's set "web2lowercase"), or plurals of words in the English dictionary (found using the Python library "inflect") were considered. Nine additional words were added to the set as well, for attention-checking purposes. This resulted in a set of 9,513 unique words from the autism corpus for the participants to rate.

To avoid survey fatigue for participants, this set of words was divided into 48 Qualtrics surveys. Each survey included 198 unique words from the autism corpus, with an additional 10 words at the beginning of the survey included as training for the participants ("raters"), and an additional 6 words interspersed throughout the survey as attention checks. Each word was to be rated according to its valence on a scale from -4 (very negative) to +4 (very positive) — this scale was adopted from Hutto and Gilbert's VADER dictionary creation method, which also inspired the rest of the dictionary-making process (2014). Each survey was considered a random condition, and around 11 raters took each survey.

In order to ensure data validity, attention checks interspersed within the main portion of the survey were employed. These consisted of four explicit attention checks (e.g. "Please choose option +3") and six implicit attention checks in the form of "golden items".

Like in the creation of the VADER lexicon by C. J. Hutto and Eric Gilbert, "golden items" were employed within the surveys (2014). "Golden items" were defined as words with well-known (pre-validated) sentiments (e.g. the word "slavery" is very negative) (Hutto & Gilbert, 2014). Golden items were selected from the VADER lexicon itself — these were words which had very consistent (a standard deviation of 0.4 or less) ratings by participants who helped

create the VADER lexicon. Sixteen words were chosen. Seven of these words were already in the autism corpus set of unique words (“love”, “charming”, “safe”, “prepared”, “reach”, “backed”, “shy”) but were added to each survey which did not already include them; the other nine were added to each survey in addition to the 198 original unique words.

Ten of these words were employed at the very beginning of the survey as training, while the other six were interspersed throughout the survey as implicit attention checks. One golden item was given every 33 words in the main body of the survey. The words, their corresponding ratings (i.e. discrete rating choice closest to the word’s mean rating), and their purpose are shown in Table 1.

**Table 1**

*VADER-Like Dictionary Golden Items*

<b>Word</b>	<b>Corresponding Rating</b>	<b>Survey Purpose</b>
love	3	Attention check
charming	3	Training
proud	2	Training
safe	2	Attention check
resolvable	1	Training
prepared	1	Attention check
reach	0	Attention check
backed	0	Training

shy	-1	Training
stutter	-1	Training
depriving	-2	Training
ungrateful	-2	Training
cruelness	-3	Attention check
repulse	-3	Training
slavery	-4	Training
terrorist	-4	Attention check

Raters who either failed any explicit attention checks or “incorrectly” rated more than 20% of the golden items had all of their survey answers excluded. An “incorrect” rating of a golden item was defined as a rating not within two points of the “correct” answer (e.g. the word “ungrateful” had a corresponding rating of -2. Raters who rated the word “ungrateful” between 0 and -4 were considered “correct” in their ratings.) Raters were not told which words were golden items, so they could not anticipate these checks.

Raters were given the following instructions at the beginning of the survey:

“On the following pages, please rate each word's valence (that is, how positive or negative that word seems according to its definition) on a scale of -4 to +4, where -4 is extremely negative, 0 is completely neutral, and +4 is extremely positive. Please choose the option you think most other people would choose. You may come across words that seem like duplicates. For example, you may encounter a word that is simply the plural version of a word you saw before. Please answer these words as you would normally. If you do not know the definition of a word, please choose 0.”

The raters were then given ten “test words” as training. Following this training section, they were given the same instructions again, and told to rate the next 204 words (198 from the autism corpus unique words set, 6 golden items) as instructed.

After excluding those who did not pass the attention checks (74 participants), the mean number of valid responses per survey was 10.3125. The sentiment dictionary was then created using a very similar method as the VADER dictionary: each word’s ratings were consolidated into one list, and the mean and standard deviation of each word’s ratings were found (Hutto & Gilbert, 2014).

Sentiment analysis using this dictionary was conducted on the autism corpus only. Before analysis, each textbook was “cleaned” — that is, all words were lower-cased, and stop words and punctuation were removed. Afterwards, each autism corpus textbook’s words were found in the built VADER-like sentiment dictionary and their corresponding mean scores were summed. The sum of the words’ scores was then divided by the number of words used for analysis (i.e. the number of words in the document, excluding stop words) to get the mean sentiment per word. This formed the entire textbook’s sentiment score, and these scores were placed into a data frame containing the textbook’s name alongside its score.

Averaging was done in order to avoid biasing results based on Word document length — dividing the “raw sum” by the number of words used for analysis simply gave the mean sentiment per word in the document, while giving a sum without division would unfairly give longer textbooks more extreme scores — for example, a textbook Word document with 500 words total, where 400 of those words were rated +4 and 100 of those words were rated -4 would have a “raw sum” of 1200, but an average total score of 2.4. A textbook Word document with 100 words total, where 80 of those words were rated +4 and 20 of those words were rated -4 would have a “raw sum” of 240, but an average total score of 2.4. If the raw sums were compared, the longer textbook Word document would be rated as much more positive than the

shorter textbook Word document, even though the proportion of positive words to negative words would have been the same in both documents.

### ***SentiWordNet***

The Python library NLTK's module SentiWordNet, which has a large sentiment analysis dictionary built with machine learning, was used to perform sentiment analysis on both the autism and the transgender corpuses (Baccianella et al., 2010). SentiWordNet has 155,287 words in its internal dictionary, as opposed to the VADER-like dictionary, which had 9,513 (Kumar et al., 2018). Before analysis, each textbook was “cleaned” — that is, all words were lower-cased, and stop words and punctuation were removed.

After cleaning, each remaining word in each textbook was given two scores (positive and negative, ranging from -1 to +1) by the SentiWordNet module's dictionary. Each word in the SentiWordNet module's dictionary has a separate positive, negative, and “neutral” score, but for simplicity, only positive and negative scores were used. These scores were added together to produce the word's “true” score. For each textbook, all of the textbook's words' scores were summed. Because SentiWordNet's scores range from -1 to +1, each sum was multiplied by 4 in order to be consistent with the VADER-like sentiment dictionary's scores ranging from -4 to +4. This multiplied sum was then divided by the number of words taken to create that sum (i.e. the number of words in the document analyzed, excluding stop words). This formed the textbook's sentiment score, and these scores were placed into a data frame containing the textbook's name alongside its score.

### ***ChatGPT***

ChatGPT is a powerful chatbot which uses large language models to generate human-like conversation (Hetler, 2023). It also has an API available which can be used in programming software (“Product”, 2024). ChatGPT can follow instructions written in natural language to solve complex problems — including sentiment analysis (“Product”, 2024). When asked how ChatGPT performs sentiment analysis (“When you create these sentiment scores,

are you taking context into account? If not, explain your process.”), the chatbot produced the following response:

The sentiment scores provided are generated based on the text's overall tone and the language used within the specified context. The scores are calculated using natural language processing techniques that analyze the words, phrases, and structure of the text to determine its sentiment. However, the process does not involve a deep understanding of context in the way humans might interpret it. Instead, it relies on patterns and associations identified during the model's training on diverse datasets. (OpenAI, 2024)

The ChatGPT-3.5-Turbo API was employed to perform sentiment analysis for each textbook in both the autism and the transgender corpuses. Unlike with the lexicon-based approaches, stop words were not removed from the documents, although headers were removed. The following prompt was then given to ChatGPT:

For every piece of text I send from now on, give a single sentiment “score” between -4 and 4 for the following piece of text. -4 represents the most negative sentiment, while 4 represents the most positive sentiment. Give the score as a decimal with up to five significant figures. Do not give any other text except for the score.

In general, the textbook documents were too large for ChatGPT to analyze fully. Instead, ChatGPT took each section of text that was separated by newlines as separate paragraphs, and gave a sentiment score to each. Occasionally, ChatGPT would give multiple scores for the same section — only the first one was used if this happened. All the sentiment scores given per paragraph were averaged into one mean score per textbook. This formed the textbook’s sentiment score, and these scores were placed into a data frame containing the textbook’s name alongside its score.

## Results

In the introduction, it was hypothesized that (collapsed across fields) both autism and transgender textbook excerpts may have improved over time, beginning with negative sentiment and progressing into more neutral sentiment. However, the differences between sentiment analysis methods and the differences between different fields of the textbooks examined were also explored. The results of the hypothesis and these explorations are below.

### **SentiWordNet vs. ChatGPT vs. VADER-Like Dictionary**

Each sentiment analysis method used was first compared against the others. With autism, this meant comparing SentiWordNet usage vs. ChatGPT's API vs. the VADER-like dictionary created.

### **Fields**

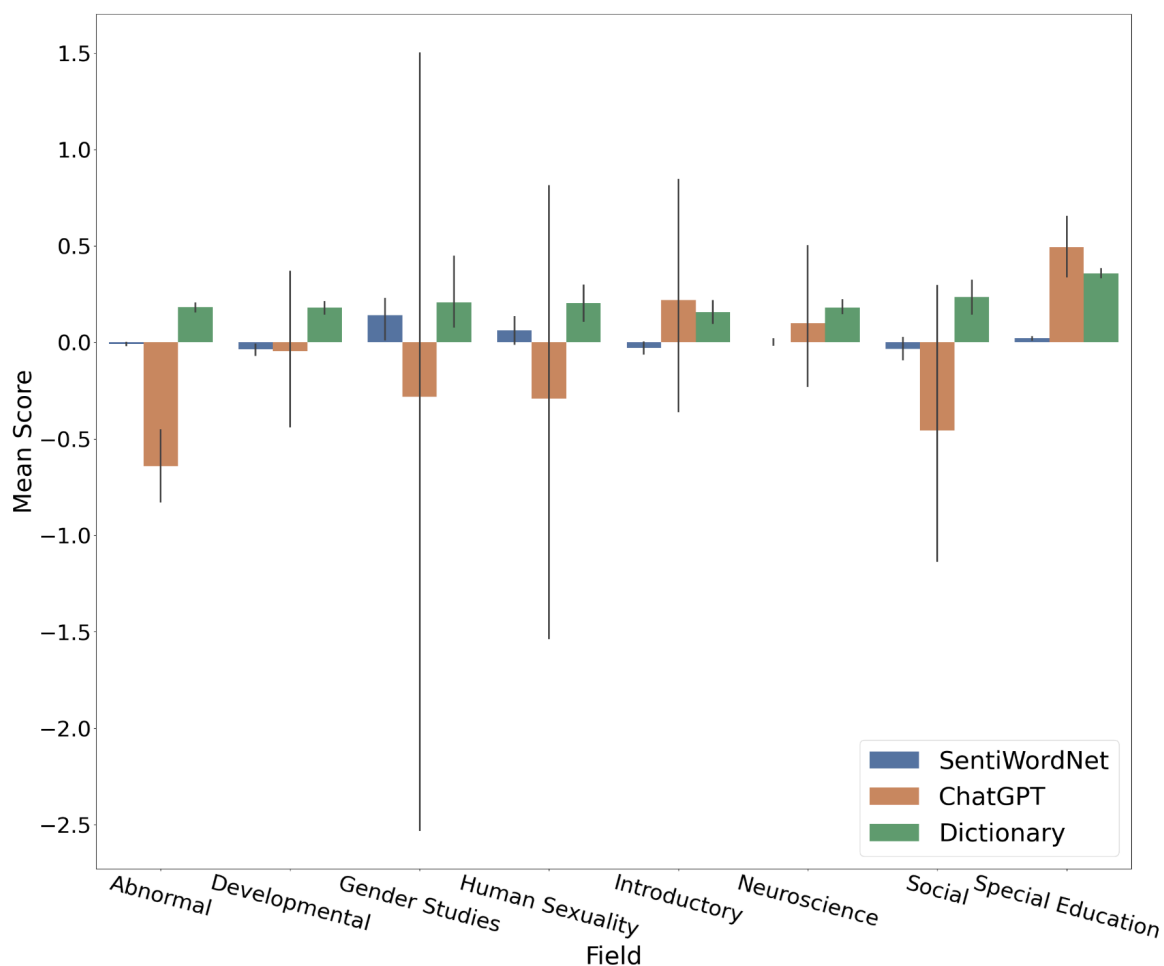
Across all textbooks, the VADER-like dictionary had an average sentiment score of 0.22 ( $SD = 0.14$ ), SentiWordNet had an average sentiment score of -0.005 ( $SD = 0.070$ ), and ChatGPT had an average sentiment score of -0.032 ( $SD = 1.18$ ). A one-way ANOVA test showed that the average sentiment score differed by the type of sentiment analysis used,  $F(2, 217) = 8.94, p < .001$ . Post-hoc Tukey analyses showed that the VADER-like dictionary yielded a higher sentiment score than both ChatGPT ( $p < .001$ ) and SentiWordNet ( $p = .002$ ), but SentiWordNet did not significantly differ from ChatGPT ( $p = .91$ ).

Figure 1 shows the average sentiment score for each field as a function of the type of sentiment analysis conducted. As the figure shows, the VADER-like dictionary yielded scores that were consistently positive across all fields. SentiWordNet yielded scores consistently around 0. ChatGPT yielded scores that were both positive and negative. ChatGPT's scores had enormous variability for most fields (indicated by the large error bars), with the exception of abnormal psychology (which had consistently negative scores) and special education (which had consistently positive scores). Therefore, very different results emerge depending on the method of sentiment analysis used.



**Figure 1**

*Autism SentiWordNet, ChatGPT, and VADER-like Dictionary Average Field Sentiment*



### **Historical Trends**

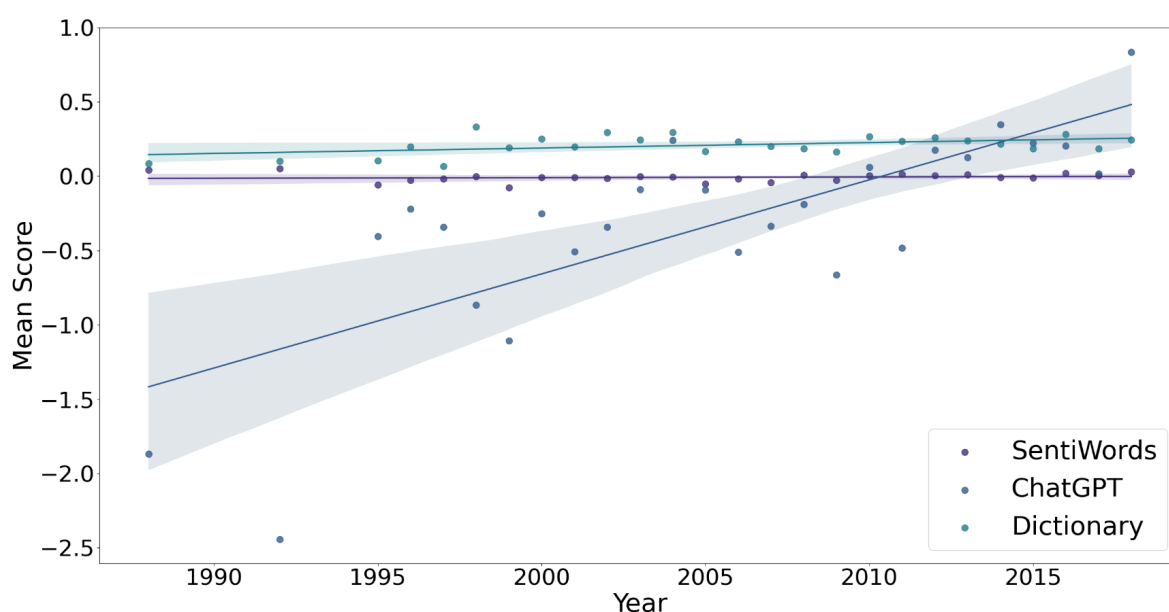
Along with average autism corpus textbook sentiment scores categorized by field, the three sentiment analysis methods also differed with average autism corpus textbook sentiment scores across years. Given increasing awareness of and attention to neurodiversity, it was hypothesized that textbooks might treat autism more positively over time. In fact, as Figure 2 shows, only ChatGPT showed this type of trajectory. A 3 (sentiment method) x 26 (year) ANOVA showed that both main effects were significant. The methods differed from each other,  $F(2, 75) =$

41.73,  $p < .001$ , and the sentiment scores increased over time,  $F(25, 75) = 3.61$ ,  $p = .012$ .

However, these main effects must be considered in light of a significant interaction between method and year,  $F(50, 75) = 3.37$ ,  $p = .013$ . As Figure 2 shows, only ChatGPT showed an increase in sentiment across years.

**Figure 2**

*Autism SentiWords, ChatGPT, and VADER-Like Dictionary Historical Trends*



### ChatGPT API

As described in the introduction, there are many ways to conduct sentiment analysis. One goal of the work here was to investigate how similar or different three methods were on the same data. As shown above, indeed these methods yielded different results. Therefore the question is: which sentiment analysis method is the most appropriate?

Context contributes to sentiment. For example, consider this sentence: “You should read the book.” By a strictly lexicon-based approach, all words in this sentence are inherently neutral. However, suppose this sentence was written as part of a book review. The sentence would then have a positive sentiment. On the other hand, suppose this sentence was written as part of a

film review. In this case, the sentence would hold a negative sentiment. Context is crucial to understanding sentiment.

SentiWordNet and VADER-like dictionaries do not take context into account. However, ChatGPT does (even if its ability is still below humans). When given the above sentence in either context, ChatGPT responds as follows:

Prompt: "Consider a book review. What kind of sentiment does the sentence 'you should read the book' have towards that book? Give a single sentiment 'score' between -4 and 4 for the following piece of text. -4 represents the most negative sentiment, while 4 represents the most positive sentiment. Give the score as a decimal with up to five significant figures."

Answer: "The sentence 'you should read the book' typically carries a positive sentiment towards the book, as it implies a recommendation or endorsement. In terms of sentiment score:

Sentiment score: 3.5." (OpenAI, 2024)

Prompt: "Consider a film review. What kind of sentiment does the sentence 'you should read the book' have towards that film? Give a single sentiment 'score' between -4 and 4 for the following piece of text. -4 represents the most negative sentiment, while 4 represents the most positive sentiment. Give the score as a decimal with up to five significant figures."

Answer: "The sentence 'you should read the book' in a film review context carries a negative sentiment towards the film, as it implies that the book is likely a better experience.

Sentiment score: -3.8." (OpenAI, 2024)

Unlike SentiWordNet and the VADER-like dictionary, which both gave a sentiment score of exactly 0.0 to "you should read the book", ChatGPT is able to recognize that there is sentiment associated with this sentence due to context, and gives contrasting answers as a

result depending on whether the sentence is written in a book review versus a film review.

Therefore, ChatGPT may be the most appropriate method for determining textbook sentiment out of the three sentiment analysis methods used, and all results shown below rely solely on ChatGPT.

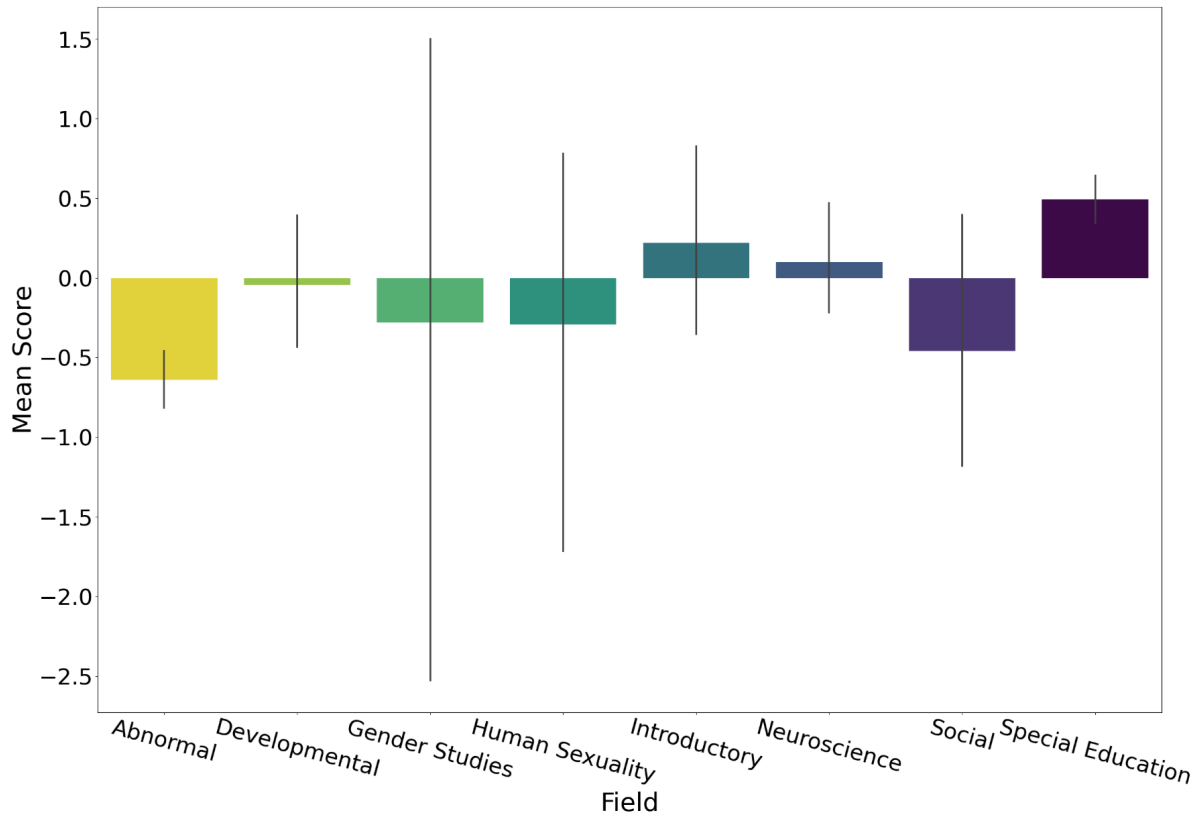
### **Fields**

Using ChatGPT, the first analysis asked whether there were differences across fields in the average sentiment scores (collapsed across all years of publication). As Figure 3 shows, abnormal psychology textbooks received the lowest average sentiment score and special education received the highest. A one-way ANOVA on the data in Figure 3 confirmed that sentiment scores differed by field;  $F(7, 210) = 4.06, p < .001$ . Post-hoc Tukey analyses showed that the average sentiment score for abnormal psychology differed from introductory psychology ( $p = .027$ ) and special education ( $p < .001$ ). None of the other pairwise comparisons were significant.

To investigate whether sentiment scores differed from 0 (i.e., were significantly positive or negative), a series of one-sample t-tests were carried out. Only two fields showed an average sentiment score different from 0: Abnormal psychology was significantly negative;  $t(45) = -6.77, p < .001$ , and special education was significantly positive;  $t(48) = 6.24, p < .001$ .

### **Figure 3**

*Mean Sentiment Analysis Score per Textbook Field Using Autism Search Terms by ChatGPT*



### ***Historical Trends***

After examining field differences, the next investigation focused on differences between years. Figure 2 in the section examining differences between sentiment analysis methods showed that ChatGPT yielded a significant increase in sentiment over time collapsed across all fields. Here, the question was whether fields showed different rates of improvement in sentiment scores over time. Gender studies and human sexuality were not included in this analysis because there were only 3 gender studies textbooks with autism-relevant content (and all came from the same year), and there were only 7 human sexuality textbooks with autism-relevant content.

To investigate whether ChatGPT sentiment scores changed over time and whether there were differences between fields, a series of mixed-effects regression models were compared. The null model predicted sentiment scores from the random effect of textbook author:

Score~(1|Author). The null model was compared to a model that added a fixed effect of publication year: Score~Year+(1|Author). The logLikelihood of the null model with 3 degrees of freedom was -308.57. The logLikelihood of the new model with 4 degrees of freedom was -295.43, which represents a significant increase in goodness of fit,  $\Delta\chi^2(1) = 26.29$ ,  $p < .001$ . The AIC and BIC statistics decreased from the null to the new model, also indicating an increase in goodness of fit (AIC: 623.14 to 598.85; BIC: 633.15 to 612.20). Adjusted  $R^2$  increased from 0.44 to 0.53. Number of observations: 208; Groups: Authors: 53.

The second model was compared to a third model that added a fixed effect of field: Score~Year+Field+(1|Author). The addition of fields resulted in a significant improvement in model fit. The logLikelihood of the model with field and 9 degrees of freedom was -286.67, which represents a significant increase in goodness of fit from the model that just included publication year,  $\Delta\chi^2(5) = 17.51$ ,  $p = .004$ . The AIC statistic decreased from the second to the third model (from 598.85 to 591.35), but the BIC statistic increased (from 612.20 to 621.38). Adjusted  $R^2$  increased from 0.53 to 0.59. Number of observations: 208; Groups: Authors: 53.

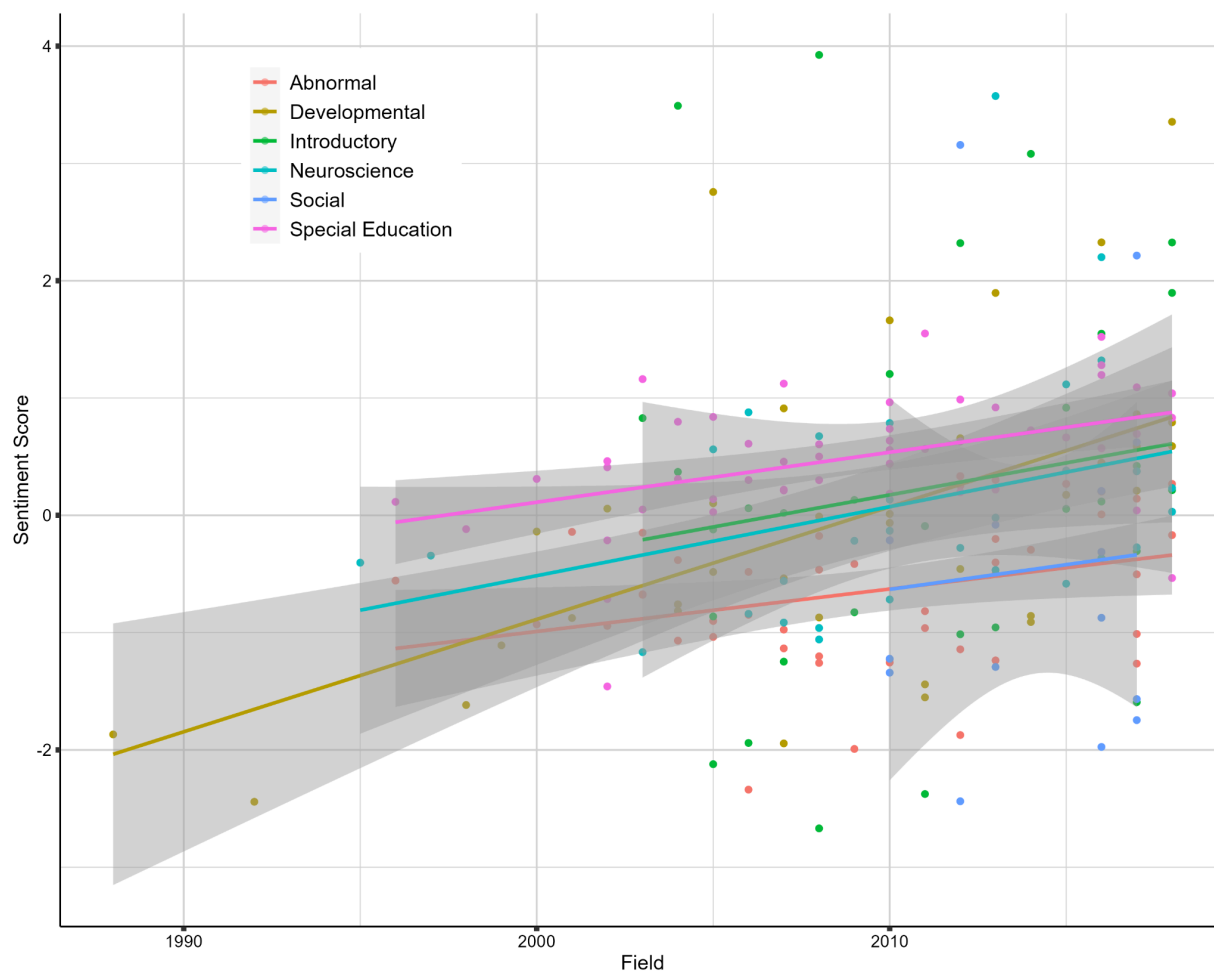
A final model was constructed that included the interaction between year and field to investigate whether some fields improved at a faster rate over time than others: Score~Field\*Year+(1|Author). The model with this interaction did not result in a significant improvement in fit over the model that was additive,  $\Delta\chi^2(5) = 7.55$ ,  $p = .183$ . The fact that the model with the interaction term between field and year was not a significant improvement over the additive model indicates that the increase in sentiment scores over time did not differ by field.

Figure 4 shows predictions from the best-fitting, additive model. As the figure shows, sentiment scores increased over time;  $F(1, 158) = 33.15$ ,  $p < .001$ , and some fields had higher sentiment scores than others;  $F(5, 40) = 3.46$ ,  $p = .01$ . Specifically, collapsing across years, neuroscience and special education had higher sentiment scores over time than abnormal

psychology (neuroscience vs. abnormal psychology:  $t(51) = 3.00, p = .050$ ; special education vs. abnormal psychology:  $t(41) = 2.99, p = .050$ ).

**Figure 4**

*Mean Autism Corpus Textbook Sentiment Score per Year by Field Using ChatGPT*



### ***Autism vs. Transgender***

The ChatGPT sentiment analysis yielded a significant increase in sentiment over time, suggesting that (collapsed across textbook fields), the way autism is written about has become more positive. However, it is possible that more recent textbooks are simply using more positive or neutral language generally to describe everything. To investigate this possibility, textbooks

excerpted with autism search terms were compared with textbooks excerpted with transgender search terms using ChatGPT's API.

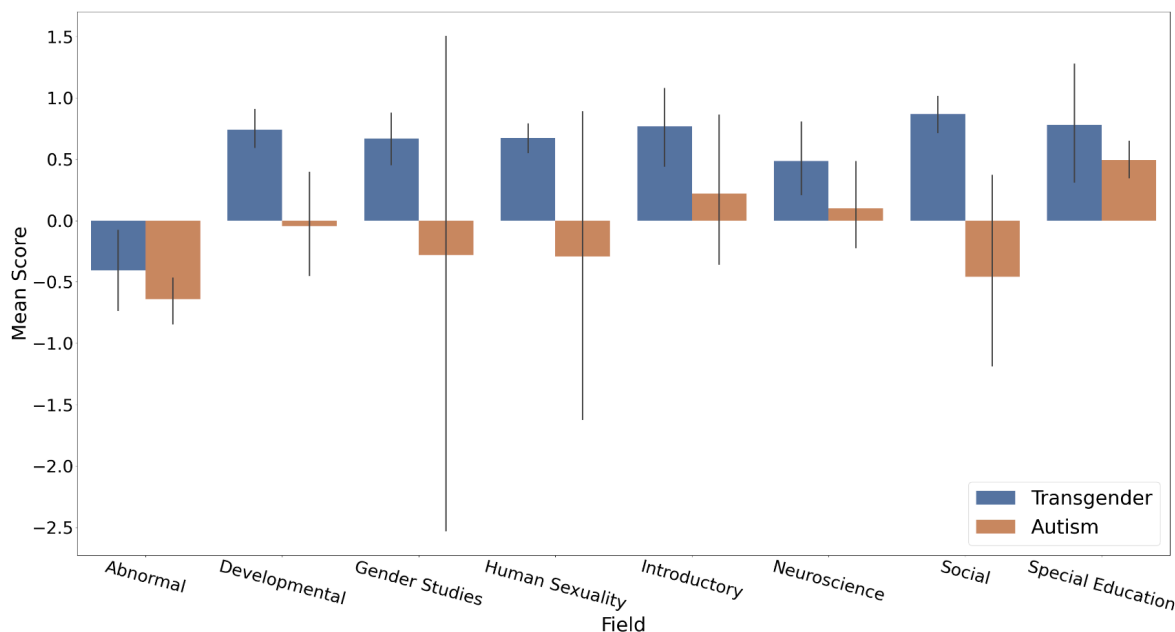
First, excerpts with transgender search terms were compared to excerpts with autism search terms as a function of what field the excerpts came from (collapsed across year). As Figure 5 shows, excerpts with transgender search terms were positive in textbooks from 7 of the 8 fields whereas excerpts with autism search terms were much more variable across the fields. Each field for the transgender corpus of textbooks had an average sentiment score which was significantly different from zero;  $p_s < .006$ .

Next, a 2 (type) x 8 (field) ANOVA was conducted on the data in Figure 5. The ANOVA yielded significant main effects of type (transgender vs. autism);  $F(1, 440) = 39.16, p < .001$ , and of field;  $F(7, 440) = 10.14, p < .001$ . There was no interaction between type and field. The figure shows that transgender excerpts generally had higher sentiment ratings than autism excerpts, and there was some variability across the different fields in terms of the absolute level of rating given.

### **Figure 5**

*Differences Between Autism and Transgender Excerpted Textbooks Analyzed by ChatGPT*

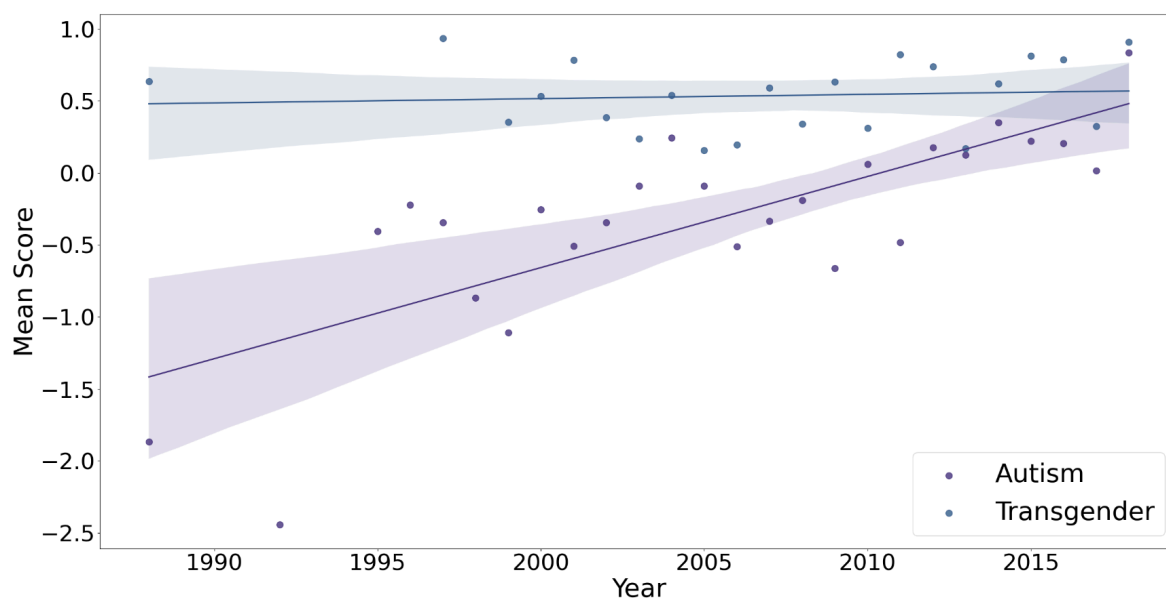




As seen in Figure 6, the transgender corpus sentiment scores did not improve over time whereas the autism sentiment scores did. A 2 (type: transgender vs. autism) x 26 (year) ANOVA showed that both main effects were significant. The types differed from each other,  $F(1, 44) = 62.98, p < .001$ , and the sentiment scores increased over time,  $F(25, 44) = 29.50, p < .001$ . However, these main effects must be considered in light of a significant interaction between type and year,  $F(1, 44) = 17.70, p < .001$ . As Figure 6 shows, only autism showed an increase in sentiment across years, from negative to neutral, while transgender sentiment scores were always neutral.

**Figure 6**

*Historical Trajectory Comparison Between Autism and Transgender-Excerpted Textbooks by ChatGPT*



A summary of the findings is as follows: First, it was found that different sentiment analysis methods yielded different results by subfield and by year, and it was determined that for following analyses, ChatGPT might be the most appropriate. Next, ChatGPT found that for autism, the field of abnormal psychology was found to have consistently negative sentiment scores per textbook collapsed across years, while special education was found to have consistently positive sentiment scores per textbook collapsed across years. ChatGPT also found that overall, collapsed across fields, sentiment scores pertaining to autism significantly improved over time.

Afterwards, autism was compared to transgender subjects. It was found that fields writing about transgender subjects tended to write about transgender subjects positively, with the exception of abnormal psychology, which continued to show negative sentiment scores collapsed across years. Collapsed across fields and looking at years, transgender writing was neutral and stayed neutral, while autism showed a significant difference by improving overall from negative to neutral. The implications of these results will be discussed further in the discussion section.

## Discussion

In the introduction, it was hypothesized that textbook authors may write more positively about autism over time, due to the increase in autism visibility and neurodiversity awareness. Other analyses, like examining different fields and comparing sentiment analysis methods, were purely exploratory in nature and had no initial hypotheses attached. The results section showed the differences between sentiment analysis methods, fields, historical trends, and comparisons between autism and transgender subjects. These results and their implications are discussed below.

### Sentiment Analysis Methods

The method used for investigating sentiment in textbooks with autism content matters. It was shown that there were statistically significant differences between the three methods used (SentiWordNet, VADER-like dictionary, and ChatGPT). SentiWordNet produced the most neutral sentiment scores (that is, it produced, on average, sentiment scores closest to 0), while the VADER-like dictionary produced the most positive sentiment scores. ChatGPT was also neutral, though it had enormous variability in its scores.

The differences between SentiWordNet's neutrality and the VADER-like dictionary's positivity are worth noting. SentiWordNet's sentiment dictionary was created through machine learning, while the VADER-like dictionary was created through human ratings. That is, SentiWordNet's lexicon was created through semi-supervised learning, while the VADER-like lexicon was created by averaging human ratings given for each word in the autism corpus. It is possible that SentiWordNet's sentiment scores appeared more neutral because its dictionary was more neutral overall — after all, SentiWordNet's dictionary was evaluated for accuracy before being released to the public, while the VADER-like dictionary's accuracy was not evaluated at all beyond using attention simple checks, and simply took users' responses at face value without altering them or checking for accuracy (Baccianella et al., 2010). The VADER-like

dictionary's positivity may be due to a positivity offset, where people may have been more likely to rate neutral situations (or in this case, neutral words) as being positive (Norris et al., 2011).

On the other hand, it may also be possible that the human-rated VADER-like dictionary indeed was more accurate, and SentiWordNet's machine-learning-based lexicon was unduly negative. For example, when looking at the VADER-like dictionary and SentiWordNet's base lexicon, seemingly objective words such as "mildly", "drawing", and "facing", while they were rated neutrally with an exact score of 0.0 in the VADER-like dictionary, all had negative scores attached to them in SentiWordNet's base lexicon (-0.04, -0.22, and -0.002, respectively, from a scale of -1 to +1).

Given that ChatGPT's sentiment analysis makes use of context, unlike the VADER-like dictionary and SentiWordNet, it was decided that ChatGPT would be used in the focused analyses about historical trends and differences between fields.

## **Autism**

Out of the eight fields investigated, ChatGPT provided sentiment scores about their autism content for only two: abnormal psychology and special education. Abnormal psychology textbooks, on average, had negative sentiment when describing autism. Special education textbooks, on average, had positive sentiment when describing autism. Why might this be?

"Abnormal", on its own, is a word with negative connotations. It would seem that abnormal psychology, as a field dedicated to pathology and aspects of human behavior deemed problematic by society, involves a negative lexical bias towards conditions like autism. Autism could be described using negative vocabulary, because autism (along with everything else in the DSM) is seen as a disorder and therefore as negative. Even the most positively rated abnormal psychology textbook had a sentiment score of only 0.38, and included passages like this:

[Hans Asperger] described a group of boys who possessed rather good language and cognitive skills, but had marked social problems because they acted like pompous 'little professors' and were physically awkward. ... At [preschool age], when most children

develop social and interactive skills, these children have difficulty reading the social cues of others, taking turns talking, and are unable to interpret language subtleties. Early in their lives, they tend to become preoccupied with a narrow set of interests about which they may talk extensively, not realizing that such one-sided conversations are not socially appropriate. (Whitbourne, 2017, p. 123)

Despite being more positive than the other abnormal psychology textbooks, this textbook uses a charged affective vocabulary such as “pompous” and “awkward”, and explicitly stigmatizes autistic children as having “social problems” and being “not socially appropriate.”

As “abnormal” is an inherently negative word, “special” is inherently positive. Special education, as a field, may then have a positive lexical bias towards autism and other developmental conditions. The “special” in “special education” implies an extraordinariness in these students which is not seen in others – a small implication that autism may be different, but ultimately additive in nature. Even the most negatively rated (sentiment score of -0.71) special education textbook with a significant amount of autism content still had sentences like this: “Not all children with autism exhibit each of the characteristics described by Lovaas and Newsom (1976). Many children with autism are ‘quite loving and caring, thoughtful and creative’ (Greenspan & Weider, 1997, p. 88)” (Heward et al., 2002, p. 486). Despite being more negative than other special education textbooks, this textbook still describes autistic children with positive qualities, even trying to explicitly deny popular claims as made by famous researchers. Because of special education’s positive lexical bias, it would be reasonable to assume that it would have a more positive performance in sentiment analysis.

In terms of historical trends, across all fields, autism textbook sentiment increased from 1988 to 2018, from negative to neutral. This suggests that the traditional stigma surrounding autism is lessening, though not yet to the point where autism is considered a positive attribute, possibly due to a persisting medicalization and bias against mental health conditions like it (Corrigan & Watson, 2002). The positive trend in these textbooks may be representing that

change in popular perception – and working to shape popular perception as well, as described in the introduction.

As noted, sentiment scores in all fields increased at the same rate; there was no interaction between year and field. This is interesting because the starting sentiment scores for some fields (like abnormal psychology and social psychology) were lower than others (like special education and introductory psychology). One explanation could have to do with the age of the textbooks of each field. Developmental psychology, which had the oldest textbooks, started with very low sentiment scores. Introductory psychology, on the other hand, consisted of relatively newer textbooks, and these textbooks started with higher sentiment scores. These scores then may reflect an overall improvement in autism public sentiment and professional discourses generally, rather than improvements in any specific field.

### **Autism vs. Transgender**

As a comparison case, the same kinds of analyses were conducted on another corpus available from many of the same textbooks — namely, a corpus that included text relevant to transgender issues. In the autism corpus, collapsing across years, the only two fields that yielded a sentiment score different from zero were special education (greater than 0) and abnormal psychology (less than 0). In contrast to the transgender corpus, collapsing across years, all fields (except abnormal psychology) yielded sentiment scores that were greater than 0.

To investigate whether the improvement in sentiment scores across time was unique to autism or was shared with other types of people described in the same textbooks, sentiment scores in the autism corpus were compared with sentiment scores in the transgender corpus. The autism sentiment scores uniquely increased over time; the sentiment scores in the transgender corpus stayed the same, hovering around 0.5.

In terms of the difference between types of textbook excerpts, all fields in the transgender corpus on average wrote about transgender subjects positively — except for

abnormal psychology, which on average wrote about transgender subjects negatively. This is in contrast to autism, where only special education and abnormal psychology had significant effects either way — special education also being positive, and abnormal psychology also being negative.

The lack of interaction between field and type may be due to the enormous variability seen in ChatGPT's analysis of autism textbooks. Because there were textbooks across the full spectrum from negative to positive for most fields, it was difficult to determine any significant differences between the transgender corpus and the autism corpus for those fields. For example, gender studies. Gender studies usually had positive sentiment scores in the transgender corpus, but was extremely variable in the autism corpus – for the three gender studies textbooks which had autism content, one had a relatively high sentiment analysis score (1.5), one had a relatively low sentiment analysis score (-2.53), and the last one had a fairly neutral sentiment analysis score (0.18). Because these textbooks were so differently scored, it would have been impossible to predict what other gender studies textbooks would have been rated, and so comparison between the autism gender studies textbooks and the transgender gender studies textbooks would have yielded no significant results.

### **Limitations**

It is important to note several limitations of the current work. There are several important limitations to the use of ChatGPT as a sentiment analysis method. First, ChatGPT was unable to analyze large amounts of text at once. Each set of excerpts for one particular textbook had to be given to ChatGPT paragraph by paragraph, and the sentiment scores given to each of these paragraphs were averaged into one number per textbook. While a paragraph still provides more context than a collection of individual words added together (as was the case with the lexicon-based methods), it would have been ideal for ChatGPT to consider the full context of each excerpt. Unfortunately, token limitations limited this. Secondly, ChatGPT has an infamous tendency to produce artificial hallucinations, as described in the introduction. Additionally,

ChatGPT did not show evidence of how it came to its conclusions, and so how exactly it generated these scores remains a “black box” mystery. It is possible that while analyzing the data, ChatGPT may have fabricated some sentiment scores, giving an inaccurate rating by chance to various pieces of text. Lastly, ChatGPT may also have simply “misunderstood” some paragraphs (and by extension the full textbook excerpts), and given inaccurate scores as a result, as it is an algorithm that does not involve a deep understanding of context.

Another limitation was the number of textbooks available in each field. Some fields (e.g., gender studies) had only 3 textbooks (and all from 2017), while other fields (e.g., human sexuality) had 7 textbooks (spanning 2011 to 2017). These differences meant that no historical trends could be analyzed for the autism content within gender studies, and there was not an adequate amount of content for human sexuality’s autism content to be analyzed in terms of a historical trend either.

The way textbooks with transgender content were excerpted was also less than ideal. Transgender search terms included just the word “gender”, which may have skewed results because of how broad of a word it is. The word “gender” caught many passages that may have related to gender, but did not talk about transgender people specifically. Rather, these excerpts included content such as gender differences in mental disorder presentations and prevalence. It would not be unreasonable to assume that a passage describing gender differences in depression prevalence would be seen as having a neutral sentiment, and indeed, ChatGPT gives such a passage a neutral score, as seen below:

Prompt: “Give a single sentiment ‘score’ between -4 and 4 for the following piece of text. -4 represents the most negative sentiment, while 4 represents the most positive sentiment. Give the score as a decimal with up to five significant figures: ‘The earlier gender imbalance in depression lessens considerably after the age of 65. In early childhood, boys are more likely to be depressed than girls, but an overwhelming surge of depression in adolescent girls produces an imbalance in the sex ratio that is maintained



until old age, when just as many women are depressed but increasing numbers of men are also affected (Piske et al., 2009). From the perspective of the life span, this is the first time since early childhood that the sex ratio for depression is more closely balanced.” (Barlow et al., 2018)

ChatGPT: “The text discusses the shifting patterns of gender imbalance in depression across different life stages, highlighting how it becomes more balanced in older age groups. There's a factual tone without explicitly stating a positive or negative sentiment towards the topic.

Sentiment score: 0.0” (OpenAI, 2024)

This inclusion of the word “gender” in transgender search terms may have significantly skewed transgender corpus data, as shown above with a gender passage being scored as 0. Therefore, while a significant difference between autism and transgender sentiment across years was found, there remains more work to be done to see if these two subjects are actually written about differently.

Discourse concerning transgender people as a comparison case may also itself be problematic. Since transgender people have become more visible in society and social acceptance is growing, their case may be too similar to autism for a comparison to make sense (Verbeek et al., 2020). After all, autism is also more salient and less stigmatized now (Grinker, 2020). Perhaps a group that has just as much stigma as it had in previous years would have made for a more reasonable comparison to autism. For example, stigma against schizophrenia remains high, and even a regression in attitude towards schizophrenia over time has been shown (e.g. more people in 2018 believed that people with schizophrenia are dangerous than people in 1996) (Pescosolido et al., 2021). Using schizophrenia as a comparison, if schizophrenia sentiment within textbooks stayed the same while autism sentiment improved, it could be said with more confidence that the stigma within textbooks concerning autism is lessening. Conversely, if schizophrenia sentiment also improved, and autism sentiment

improved, then it would be less clear if the improvement was due to autism stigma lessening or overall mental health stigma lessening. Therefore, transgender people may not have been an appropriate comparison case.

## Conclusion

This study's results are readily summarized. Textbook content regarding autism is, overall, improving in valence across time. Collapsed across all fields, the trend of textbooks which write about autism has gone from having negative sentiment scores to neutral sentiment scores, supporting the original hypothesis that autism content will improve in sentiment over time. This is in contrast to excerpts about transgender people within those same textbooks, which showed a neutral, static sentiment over time. Autism sentiment is improving in textbooks, but transgender sentiment in textbooks has not seen any change. Moreover, special education tends to write about autism (and transgender subjects) more positively than other fields. Abnormal psychology tends to write about autism (and transgender subjects) more negatively than other fields. Other fields are so variable in their average sentiment scores that no definitive conclusions could be made. The contrast between special education and abnormal psychology may be due to differences in how they might think about neurodivergent people and neurodiversity — “special” students vs. “abnormal” patients.

Sentiment analysis, as a method, benefits from AI inclusion. Large language model-based software like ChatGPT are now capable of evaluating credible sentiment scores from context. Because of this, large-scale sentiment analysis, depending on neither human reading nor less accurate lexicon-based methods, is now possible. No longer is it necessary to manually read and rate large amounts of text. Further studies could take advantage of this new technology and explore other textual sources of sentiment outside of textbooks and education, such as news articles, scientific journals, and social media. Accurate voice recognition and the improvement of speech-to-text may even open up sentiment analysis to the possibility of analyzing speeches, lectures, and other non-textual sources which may make a difference in popular sentiment surrounding different subjects.

Overall, while autism sentiment has improved across textbooks over time, more work needs to be done. Autistic people deserve to be spoken about positively, and to be seen as

valuable individuals outside of case studies and prevalence statistics. Education can make this possible, and that begins with educational materials such as textbooks. Through education, maybe one day autism will cease to be seen as a deficit, and instead will be widely recognized as part of a larger neurodiversity.

## References

- Al-Shabi, M. A. (2020). Evaluating the performance of the most important Lexicons used to Sentiment analysis and opinions Mining. *International Journal of Computer Science and Network Security*, 20(1), 51–57.
- Baccianella, S., Esuli, A., & Sebastiani, F. (2010). SentiWordNet 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining. In *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*. European Language Resources Association (ELRA).
- Barahona-Corrêa, J. B., & Filipe, C. N. (2016). A concise history of Asperger Syndrome: The short reign of a troublesome diagnosis. *Frontiers in Psychology*, 6.  
<https://doi.org/10.3389/fpsyg.2015.02024>
- Barlow, D. H., Durand, V. M., Lalumiere, M. L., & Hofmann, S. G. (2018). *Abnormal psychology: An integrative approach* (8th ed.). Cengage.
- Bhattacharyya, M., Miller, V. M., Bhattacharyya, D., & Miller, L. E. (2023). High rates of fabricated and inaccurate references in ChatGPT-generated medical content. *Cureus*.  
<https://doi.org/10.7759/cureus.39238>
- Brockman, G., Eleti, A., Georges, E., Jang, J., Kilpatrick, L., Lim, R., Miller, L., & Pokrass, M. (2024, April 23). *Introducing ChatGPT and Whisper APIs*. OpenAI.  
<https://openai.com/index/introducing-chatgpt-and-whisper-apis>
- Cambridge University Press. (2009, April 7). *Dropping common terms: stop words*. NLP.  
<https://nlp.stanford.edu/IR-book/html/htmledition/dropping-common-terms-stop-words-1.html>
- Corrigan, P. W., & Watson, A. C. (2002). Understanding the impact of stigma on people with mental illness. *World psychiatry : official journal of the World Psychiatric Association (WPA)*, 1(1), 16–20.

- Cutillas-Fernández, M. A., Jiménez-Barbero, J. A., Herrera-Giménez, M., Forcén-Muñoz, L. A., & Jiménez-Ruiz, I. (2023). Attitudes and beliefs of mental health professionals towards Trans people: A systematic review of the literature. *International Journal of Environmental Research and Public Health*, 20(15), 6495.  
<https://doi.org/10.3390/ijerph20156495>
- de Wynter, A., Wang, X., Sokolov, A., Gu, Q., & Chen, S.-Q. (2023). An evaluation on large language model outputs: Discourse and memorization. *Natural Language Processing Journal*, 4, 100024. <https://doi.org/10.1016/j.nlp.2023.100024>
- Evans, B. (2013). How autism became autism. *History of the Human Sciences*, 26(3), 3–31.  
<https://doi.org/10.1177/0952695113484320>
- Friend, M. P. (2018). *Special education: Contemporary Perspectives for School Professionals* (5th ed.). Pearson Education, Inc.
- Grinker, R. R. (2020). Autism, “stigma,” disability. *Current Anthropology*, 61(S21).  
<https://doi.org/10.1086/705748>
- Hetler, A. (2023, December 15). *What is chatgpt? everything you need to know*. WhatIs.  
<https://www.techtarget.com/whatis/definition/ChatGPT>
- Heward, W. L., Alber, S. R., & Konrad, M. (2002). *Exceptional children: An introduction to special education* (7th ed.). Pearson.
- Hodges, H., Fealko, C., & Soares, N. (2020). Autism spectrum disorder: Definition, epidemiology, causes, and clinical evaluation. *Translational Pediatrics*, 9(S1).  
<https://doi.org/10.21037/tp.2019.09.09>
- Hooley, J. M. (2006). *Abnormal psychology* (13th ed.). Pearson Education, Inc.
- Hutto, C. (2021, March 15). *README*. GitHub.  
<https://github.com/cjhutto/vaderSentiment/blob/master/README.rst>

- Hutto, C., & Gilbert, E. (2014). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the International AAAI Conference on Web and Social Media*, 8(1), 216–225. <https://doi.org/10.1609/icwsm.v8i1.14550>
- Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child*, 2, 217–250.
- Kitchin, J. L., & Karlin, N. J. (2021). Awareness and stigma of autism spectrum disorders in undergraduate students. *Psychological Reports*, 125(4), 2069–2087. <https://doi.org/10.1177/00332941211014144>
- Katrekar, A. (n.d.). *An Introduction to Sentiment Analysis* [PowerPoint slides]. Global Logic. <https://www.globallogic.com/se/wp-content/uploads/2019/12/Introduction-to-Sentiment-Analysis.pdf>
- Kumar, J., Rout, J. K., Katiyar, A., & Jena, S. K. (2018). Sentiment analysis using weight model based on sentiwordnet 3.0. *Advances in Intelligent Systems and Computing*, 131–139. [https://doi.org/10.1007/978-981-10-8633-5\\_14](https://doi.org/10.1007/978-981-10-8633-5_14)
- LeVay, S., Baldwin, J. I., & Baldwin, J. D. (2015). *Discovering human sexuality* (3rd ed.). Sinauer Associates : Oxford University Press.
- Lewin, N., & Akhtar, N. (2020). Neurodiversity and deficit perspectives in *The Washington Post's* coverage of autism. *Disability & Society*, 36(5), 812–833. <https://doi.org/10.1080/09687599.2020.1751073>
- Lucy, L., Demszky, D., Bromley, P., & Jurafsky, D. (2020). Content analysis of textbooks via natural language processing: Findings on gender, race, and ethnicity in Texas U.S. history textbooks. *AERA Open*, 6(3), 233285842094031. <https://doi.org/10.1177/2332858420940312>
- Mcgee, M. (2012). Neurodiversity. *Contexts*, 11(3), 12–13. <https://doi.org/10.1177/1536504212456175>
- Medicine: frosted children. (1948, April 26). *Time*, LI(17). Retrieved from <https://content.time.com/time/subscriber/article/0,33009,798484,00.html>.

- Mittmann, G., Schrank, B., & Steiner-Hofbauer, V. (2023). Portrayal of autism in mainstream media – a scoping review about representation, stigmatisation and effects on consumers in non-fiction and fiction media. *Current Psychology*, 43(9), 8008–8017.  
<https://doi.org/10.1007/s12144-023-04959-6>
- Mocarski, R., King, R., Butler, S., Holt, N. R., Huit, T. Z., Hope, D. A., Meyer, H. M., & Woodruff, N. (2019). The rise of transgender and gender diverse representation in the media: Impacts on the population. *Communication, Culture and Critique*, 12(3), 416–433.  
<https://doi.org/10.1093/ccc/tcz031>
- Norris, C. J., Larsen, J. T., Crawford, L. E., & Cacioppo, J. T. (2011). Better (or worse) for some than others: Individual differences in the positivity offset and negativity bias. *Journal of Research in Personality*, 45(1), 100–111. <https://doi.org/10.1016/j.jrp.2010.12.001>
- OpenAI. (2024). *ChatGPT* [Large language model]. <https://chat.openai.com/chat>
- OpenAI. (n.d.). *GPT-4 Turbo and GPT-4*. OpenAI.  
<https://platform.openai.com/docs/models/gpt-4-turbo-and-gpt-4>
- OpenAI. (2024). *Pricing*. OpenAI. <https://openai.com/pricing>
- OpenAI. (2024). *Product*. OpenAI. <https://openai.com/product>
- OpenAI. (n.d.). *What are tokens and how to count them?* | *OpenAI help center*. OpenAI.  
<https://help.openai.com/en/articles/4936856-what-are-tokens-and-how-to-count-them>
- Pescosolido, B. A., Halpern-Manners, A., Luo, L., & Perry, B. (2021). Trends in public stigma of mental illness in the US, 1996-2018. *JAMA Network Open*, 4(12).  
<https://doi.org/10.1001/jamanetworkopen.2021.40202>
- Redfield, S. E. (2017). *Enhancing justice: Reducing bias*. American Bar Association, Judicial Division.
- Rosen, N. E., Lord, C., & Volkmar, F. R. (2021). The diagnosis of autism: From Kanner to DSM-III to DSM-5 and beyond. *Journal of Autism and Developmental Disorders*, 51(12), 4253–4270. <https://doi.org/10.1007/s10803-021-04904-1>



*Sample usage for sentiwordnet.* NLTK. (2023, January 2).

<https://www.nltk.org/howto/sentiwordnet.html>

Schmidt, E. (2022, June 2). *Required reading: How textbook adoption in 3 states influences the nation's K-12 population.* APM Research Lab.

<https://www.apmresearchlab.org/10x-textbook-adoption>

Silberman, S. (2015). *Neurotribes: The legacy of autism and the future of neurodiversity.* Avery Publishing.

Srivastava, R., Bharti, P. K., & Verma, P. (2022). Comparative analysis of lexicon and machine learning approach for sentiment analysis. *International Journal of Advanced Computer Science and Applications*, 13(3), 71–77. <https://doi.org/10.14569/ijacsa.2022.0130312>

*Stop words list.* CountWordsFree. (n.d.). <https://countwordsfree.com/stopwords>

Taboada, M., Brooke, J., Tofiloski, M., Voll, K., & Stede, M. (2011). Lexicon-based methods for sentiment analysis. *Computational Linguistics*, 37(2), 267–307.

[https://doi.org/10.1162/coli\\_a\\_00049](https://doi.org/10.1162/coli_a_00049)

*Transgenderism: Definition, meaning, and origin in anti-LGBTQ hate.* GLAAD. (2024, March 28).

<https://glaad.org/transgenderism-definition-meaning-anti-lgbt-online-hate/>

Truică, C.-O., Apostol, E.-S., Șerban, M.-L., & Paschke, A. (2021). Topic-based document-level sentiment analysis using contextual cues. *Mathematics*, 9(21), 2722.

<https://doi.org/10.3390/math9212722>

Turnock, A., Langley, K., & Jones, C. R. G. (2022). Understanding stigma in autism: A narrative review and theoretical model. *Autism in Adulthood*, 4(1), 76–91.

<https://doi.org/10.1089/aut.2021.0005>

Verbeek, M. J., Hommes, M. A., Stutterheim, S. E., van Lankveld, J. J., & Bos, A. E. (2020).

Experiences with stigmatization among transgender individuals after transition: A qualitative study in the Netherlands. *International Journal of Transgender Health*, 21(2), 220–233. <https://doi.org/10.1080/26895269.2020.1750529>

Wankhade, M., Rao, A. C., & Kulkarni, C. (2022). A survey on sentiment analysis methods, applications, and challenges. *Artificial Intelligence Review*, 55(7), 5731–5780.

<https://doi.org/10.1007/s10462-022-10144-1>

Whitbourne, S. K. (2017). *Abnormal psychology: Clinical perspectives on psychological disorders* (8th ed.). McGraw-Hill Education.

Zastrow, C., Kirst-Ashman, K. K., & Hessenauer, S. L. (2010). *Understanding human behavior and the social environment* (8th ed.). Cengage.

Zeidan, J., Fombonne, E., Scolah, J., Ibrahim, A., Durkin, M. S., Saxena, S., Yusuf, A., Shih, A., & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research*, 15(5), 778–790. <https://doi.org/10.1002/aur.2696>

## Appendix

### Textbooks Employed for Sentiment Analysis

**Table A1**

*Textbooks Used*

Most Recent Name	Field	First Author	Editions	Year Range	Autism Search Terms	Transgende r Search Terms
Abnormal Psychology: An Integrative Approach	Abnormal	Barlow	4-8	2004- 2018	Yes	Yes
Casebook in Abnormal Psychology	Abnormal	Brown	1-5	1996- 2017	Yes	Yes: 5; No: 1-4
Abnormal Psychology	Abnormal	Comer	5-9	2003- 2015	Yes	Yes
Abnormal Psychology	Abnormal	Hooley	13-17	2006- 2017	Yes	Yes
Casebook in Child Behavioral Disorders	Abnormal	Kearney	6	2017	Yes	Yes

Abnormal Child Psychology	Abnormal	Mash	2-6	2001-2016	Yes	Yes: 2, 5-6; No: 3-4
Abnormal Psychology in a Changing World	Abnormal	Nevid	6-10	2005-2018	Yes	Yes: 8-10; No: 6-7
Abnormal Psychology	Abnormal	Nolen-Hoeksema	3-7	2003-2017	Yes	Yes
Understanding Abnormal Behavior	Abnormal	Sue	7-11	2002-2016	Yes	Yes
Abnormal Psychology: Clinician Perspectives on Psychological Disorders	Abnormal	Whitbourne	4-8	2003-2017	Yes	Yes
The Developing Person Through the Life Span	Developmental	Berger	5-10	2000-2017	Yes	Yes
Infants, Children, and Adolescents	Developmental	Berk	3-8	1998-2016	Yes	Yes: 8; No: 3-7
Developmental Psychology	Developmental	Bornstein	2-7	1988-	Yes	Yes: 2, 4-5,

Science: An	al			2016		7; No: 3, 6
Advanced						
Textbook						
Development	Development	Feldman	3-8	2002-	Yes:	Yes
Across the Life	al			2017	4-8;	
Span					No: 3	
Infants,	Development	Gonzalez-M	10-11	2014-	No	Yes
Toddlers, and	al	ena		2018		
Caregivers: A						
Curriculum of						
Respectful,						
Responsive,						
Relationship-Bas						
ed Care and						
Education						
Human	Development	Kail	2-7	1999-	Yes: 7;	Yes: 2, 5-7;
Development: A	al			2016	No: 2-6	No: 3-4
Life-Span View						
Theories of	Development	Miller	6	2018	Yes	Yes
Developmental	al					
Psychology						
Development	Development	Newman	8-13	2001-	Yes:	Yes

Through Life: A	al			2018	13; No:	
Psychosocial					8-12	
Approach						
Life-Span	Development	Santrock	11-16	2007-	Yes:	Yes
Development	al			2017	12-16;	
					No: 11	
Life-Span	Development	Sigelman	4-9	2002-	Yes	Yes: 5, 7-9;
Human	al			2018		No: 4, 6
Development						
Race, Class,	Gender	Andersen	7-9	2009-	No	Yes
and Gender: An	Studies			2015		
Anthology						
Gender:	Gender	Brannon	5-7	2009-	No	Yes
Psychological	Studies			2017		
Perspectives						
Diversity and	Gender	Healey	3-5	2012-	No	Yes
Society: Race,	Studies			2017		
Ethnicity, and						
Gender						
Psychology of	Gender	Helgeson	3-5	2008-	No	Yes
Gender	Studies			2017		
The Gendered	Gender	Kimmel	4-6	2009-	No	Yes

Society	Studies			2017		
Identities & Inequalities: Exploring the Intersection of Race, Class, Gender, and Sexuality	Gender Studies	Newman	1-3	2005- 2017	Yes: 3; No: 1-2	Yes
Introducing Gender and Women's Studies	Gender Studies	Robinson	2-4	1997- 2015	No	Yes
Race, Class, and Gender in the United States: An Integrated Study	Gender Studies	Rothenberg	8-10	2009- 2017	Yes: 10; No: 8-9	Yes
Gendered Lives	Gender Studies	Wood	10-12	2013- 2017	Yes: 12; No: 10-11	Yes
Sexuality Now: Embracing	Human Sexuality	Carroll	3-5	2009- 2016	Yes: 5; No: 3-4	Yes

Diversity						
Our Sexuality	Human Sexuality	Crooks	11-13	2011-2017	Yes: 11, 13; No: 12	Yes
Exploring the Dimensions of Human Sexuality	Human Sexuality	Greenberg	4-6	2010-2017	No	Yes
Human Sexuality	Human Sexuality	Hock	2-4	2009-2016	No	Yes
Understanding Human Sexuality	Human Sexuality	Hyde	11-13	2010-2017	Yes: 13; No: 11-12	Yes
Discovering Human Sexuality	Human Sexuality	LeVay	1-3	2009-2015	Yes: 2-3; No: 1	Yes
Human Sexuality: Diversity in Contemporary America	Human Sexuality	Yarber	7-9	2009-2016	Yes: 9; No: 7-8	Yes
Psychology: Foundations and	Introductory	Bernstein	6-10	2003-2016	Yes	Yes



Frontiers						
Psychology: Modules for Active Learning	Introductory	Coon	10-14	2005-2018	Yes: 10-11, 13-14; No: 12	Yes
Psychology: A Concise Introduction	Introductory	Griggs	1-5	2005-2018	Yes: 3-5; No: 1-2	Yes: 1, 4-5; No: 2-3
Introduction to Psychology	Introductory	Kalat	8-11	2007-2017	Yes: 9-11; No: 8	Yes: 8, 11; No: 9-10
Understanding Psychology	Introductory	Morris	7-11	2005-2016	Yes	Yes: 7, 9-11; No: 8
Psychology	Introductory	Myers	7-11	2004-2015	Yes	Yes
Essentials of Psychology: Concepts and Applications	Introductory	Nevid	1-5	2005-2018	No	Yes
PSYCH	Introductory	Rathus	1-5	2008-2018	Yes: 5; No: 1-4	Yes

Psychology	Introductory	Wade	8-12	2005- 2017	Yes: 8, 12; No: 9-11	Yes
Psychology: Themes and Variations	Introductory	Weiten	6-10	2004- 2017	Yes: 6, 8, 10; No: 7, 9	Yes: 7-10; No: 6
Neuroscience: Exploring the Brain	Neuroscience	Bear	3-4	2006- 2017	Yes	Yes
Behavioral Neuroscience	Neuroscience	Breedlove	5-8	2007- 2017	Yes	Yes
Physiology of Behavior	Neuroscience	Carlson	9-12	2006- 2017	Yes	Yes
Brain & Behavior: An Introduction to Biological Psychology	Neuroscience	Garett	2-5	2008- 2018	Yes	Yes
Fundamental Neuroscience for Basic and	Neuroscience	Haines	5	2017	No	Yes

Clinical Applications						
Developmental Cognitive Neuroscience	Neuroscience	Johnson	1-4	1997- 2015	Yes	Yes: 4; No: 1-3
Biological Psychology	Neuroscience	Kalat	10-12	2008- 2016	Yes	Yes: 12; No: 10-11
Fundamentals of Human Neuropsycholog y	Neuroscience	Kolb	4-7	1995- 2015	Yes	Yes: 7; No: 4-6
Biopsychology	Neuroscience	Pinel	7-10	2007- 2018	Yes	Yes
Cognition: Exploring the Science of the Mind	Neuroscience	Reisberg	6	2016	Yes	Yes
Social Psychology	Social	Aronson	8-9	2012- 2016	Yes	Yes
Social Psychology and Human Nature	Social	Baumeister	2-4	2010- 2017	Yes: 4; No: 2-3	Yes: 2, 4; No: 3

Social Psychology	Social	Branscombe	12-14	2008- 2017	No	Yes
Social Psychology	Social	Franzoi	5-7	2008- 2016	No	Yes
Social Psychology	Social	Gilovich	2-4	2010- 2016	Yes	Yes
Social Psychology	Social	Gruman	2-3	2011- 2017	Yes: 3; No: 2	Yes
[Unknown]	Social	Kassin	9-10	[Unkn own]	No	Yes
[Unknown]	Social	Myers	10-12	[Unkn own]	Yes: 12; No: 10-11	Yes
[Unknown]	Social	Rogers	2-4	[Unkn own]	Yes	Yes
[Unknown]	Social	Zastrow	8-10	[Unkn own]	Yes	Yes
Special Education: Contemporary Perspectives for	Special Education	Friend	1-5	2007- 2018	Yes	Yes: 2-5; No: 1

School						
Professionals						
Special Education in Contemporary Society: An Introduction to Exceptionality	Special Education	Gargiulo	3-6	2005- 2017	Yes	Yes: 6; No: 3-5
Human Exceptionality: School, Community, and Family	Special Education	Hardman	8-12	2004- 2017	Yes	Yes: 10, 12; No: 8-9, 11
Exceptional Children: An Introduction to Special Education	Special Education	Heward	7-11	2002- 2017	Yes	Yes: 10-11; No: 7-9
Teaching Students with Language and Communication Disabilities	Special Education	Kuder	1-5	1996- 2018	Yes	Yes: 5; No: 1-4

Teaching Special Students in General Education Classrooms	Special Education	Lewis	5-9	1998- 2016	Yes	Yes: 9; No: 5-8
Assessing Learners with Special Needs: An Applied Approach	Special Education	Overton	4-8	2002- 2016	Yes: 4, 6-8; No: 5	Yes: 8; No: 4-7
Teaching Students with Special Needs in Inclusive Settings	Special Education	Smith	3-7	2000- 2015	Yes	Yes: 4-5, 7; No: 3, 6
Exceptional Lives: Special Education in Today's Schools	Special Education	Turnbull	4-8	2003- 2016	Yes	Yes
Teaching Students Who are Exceptional, Diverse, and At	Special Education	Vaughn	3-7	2002- 2018	Yes	Yes: 7; No: 3-6

---

Risk in the

General

Education

Classroom