Cohesion and Coherence in Essays Generated by ChatGPT: A Comparative Analysis to University Students' Writing

By
Hanaa Youssef Shaarawy Ismail
Associate Professor of Linguistics, School of Linguistics and Translation,
Badr University in Cairo (BUC)

Abstract

ChatGPT is a natural language processing tool that can assist in many language-related tasks (ChatGPT, personal communication, May 20, 2023). It can also produce human-like writing. Cohesion and coherence are major requirements of any piece of writing, and it is assumed that ChatGPT can generate successful essays regarding these language requisites. Adopting Halliday and Hasan's (1976) theory of cohesion and van Dijk and Kintsch's (1983) cognitive situation model to achieve coherence, the researcher investigates cohesion and coherence in essays produced by ChatGPT compared to essays produced by students with respect to the same language features. The current study revolves around two major research questions: how are cohesion and coherence manifested in essays produced by ChatGPT, and how do these same language features resemble or differ from essays written by students? To answer these questions, students' essays from first and fourth years in the English Department at one of the Egyptian private universities and essays produced by ChatGPT on the same prompts were assessed by the two courses' instructors, qualitatively analyzed by ChatGPT and quantitatively analyzed using Coh-Metrix software version 3.0, and then compared. Findings revealed that ChatGPT can produce human-like essays not only in terms of cohesion and coherence, but also regarding lexical choices and advanced grammatical structures, which poses a great challenge of how to avoid having students plagiarizing from ChatGPT. Accordingly, the article concludes with some recommendations that utilize ChatGPT to enhance learners' linguistic proficiency without sacrificing the integrity, honesty and reliability of the learning process.
Keywords: ChatGPT, cohesion, coherence, Coh-Metrix 3.0

التماسك والترابط فى المقالات الصادرة عنChatGPT
دراسة تحليلية مقارنة مع كتابات الطلاب الجامعيين

هناء يوسف شعراوى اسماعيل
قسم اللغة الانجليزية
كلية اللغات والترجمة – جامعة بدر بالقاهرة

مستخلص

يعد ChatGPT أداة إلكترونية لمعالجة اللغة الطبيعية يمكنها أن تصدر مقالات تشبه تلك التى يكتبها الإنسان. ويعد التماسك والترابط بين الأفكار والعبارات من أهم السمات التى تتسم بها المقالات التى يكتبها الإنسان. فتهدف هذه الدراسة إلى استخدام نظرية التماسك لهاليداى وحسن (1976) ونظرية الترابط لفان ديك وكينتش (1983) لتحليل المقالات الصادرة عن ChatGPT والمقالات التى يكتبها الإنسان متمثلة فى مقالات طلاب الفرقة الأولى والفرقة الرابعة بقسم اللغة الانجليزية فى إحدى الجامعات الخاصة بالقاهرة. تطرح الدراسة الحالية تساؤلين هما: كيف يظهر التماسك والترابط فى المقالات الصادرة عن ChatGPT وكيف تختلف أو تتوافق هاتين السماتين مع المقالات التى يكتبها الطلاب؟ وقد خضعت جميع المقالات موضوع الدراسة إلى تقييم القائمين بالتدريس كما خضعت للتحليل النوعى باستخدام ChatGPT ثم التحليل الكمى باستخدام Coh-Metrix Software version 3.0. وأظهرت النتائج قدرة ChatGPT على إصدار مقالات تشبه تلك التى يصدرها الطلاب ليس فقط من حيث التماسك والترابط ولكن أيضا فيما يتعلق باختيار المفردات والتراكيب النحوية المتقدمة مما يشكل تحديا كبيرا لكيفية تجنب الانتحال العلمى نتجية لوجود ChatGPT ، وبالتالى تختتم الدراسة ببعض التوصيات لاستخدام ChatGPT للنهوض بالعملية التعليمية دون المساس بالنزاهة أو المصداقية.

الكلمات المفتاحية: ChatGPT، التماسك، الترابط، Coh-Metrix Software 3.0

Cohesion and Coherence in Essays Generated by ChatGPT: A Comparative Analysis to University Students' Writing

By
Hanaa Youssef Shaarawy Ismail
Associate Professor of Linguistics, School of Linguistics and Translation,
Badr University in Cairo (BUC)

Introduction

Writing is the most daunting task for university students and educators. Students are to submit weekly written assignments to show their understanding of a point, and tutors are to mark these assignments and give feedback. Artificial Intelligence (AI) has recently provided Large Language Models (LLMs) that can be of great benefit and help to students and teachers like Bing, Google Bard and ChatGPT. This article is concerned with ChatGPT which is

> an AI language model developed by OpenAI. It is based on the GPT-3.5 architecture, which stands for "Generative Pre-trained Transformer 3.5." GPT-3.5 is designed to understand and generate human-like text based on the input it receives. It has been trained on a large corpus of diverse text data, allowing it to generate responses, provide information, engage in conversations, and assist with various language-related tasks. ChatGPT can be used for a wide range of applications, including chatbots, virtual assistants, content generation, language translation, and more. (ChatGPT, personal communication, May 20, 2023)

ChatGPT can assist the teacher in tailoring reading passages, creating grammar and vocabulary exercises, providing feedback on students' work and aiding in lesson planning (Cook, 2023). Another key capability of ChatGPT is generating essays. To test the reliability of essays produced by ChatGPT, the present article aims at examining Halliday and Hasan's (1976) theory of cohesion and van Dijk and Kintsch's (1983) cognitive situation model to achieve coherence as two fundamental linguistic components for any successful communication in ChatGPT essays and compare them to human essays with respect to the same language constituents. The article ends up with some recommendations for using ChatGPT as a tool that helps students in the learning process.

Theoretical Framework

Halliday and Hasan (1976) maintained that cohesion is a semantic concept that has to do with the relationships or ties within a text. They added that cohesion could be "grammatical" and could be "lexical" (p. 6).

Grammatical cohesion falls within four types: "reference", "substitution", "ellipsis" and "conjunction". Lexical cohesion, further, has to do with the choice of words. This includes "reiteration" and "collocation". "Reiteration" could appear as "repetition", "synonymy/near-synonymy", "antonymy" or "superordination". "Collocation" is that part of lexical cohesion where combination of words co-occurs together.

van Dijk and Kintsch's (1983), on the other hand, presented a cognitive semantic situation model for coherence as a strategy for discourse comprehension. The situation model is a mental model which is "a cognitive representation of the events, actions, persons, and in general the situation, a text is about" (pp. 11-12). The situation model activates previous and general knowledge from the "semantic memory" about similar situations (p. 12). The components of the model include: "propositional strategies", "local coherence strategies", "macrostrategies" and "schematic strategies". "Propositional strategies" means the creation of propositions. Propositions are realized from the meaning of words "activated from semantic memory and syntactic structures of clauses" (p. 14). "Local coherence strategies" emerge from the connection between propositions. The order of the clauses, cohesive devices together with "knowledge from long-term memory" help in establishing the connection among propositions (p. 15). "Macrostrategies" or "global coherence strategies" are inferred macropropositions when local propositions are organized to form the macrostructure (topic) of a text (p. 15). "Schematic strategies" form the "superstructure" of the text which is the overall organization of the macropropositions (p. 16). van Dijk (1985) added that the "context" and the "control system" are constraints that specify the relevant propositions that are to be retrieved from similar situations in "episodic memory" for the communicative goal of the text (p. 67).

Reinhart (1980) has earlier set the same conditions for a text to be coherent as those discussed by van Dijk and Kintsch (1983) but in a different wording. The first condition is "cohesion" which is the connectedness between sentences. The second is "consistency" which is a semantic condition relating the meaning of propositions to one another. The third is "relevance" that limits the relations of the meaning of the sentences to the topic and the context of the text (p. 164).

van Dijk and Kinstch's (1983) situation model and coherence also triggered latter scholars to work in the same arena. Tapiero (2007) highlighted the relevance of the retrieved model to the associated given text based on the addressee's prior knowledge. Kehler (2022) adopted a relational concept of coherence that differentiates discourse from a series

of unrelated utterances and that explains how "pragmatic enrichments" arise from the inference of unstated propositions. These relations are: "Resemblance, Contiguity in Time or Place, and Cause-Effect". Resemblance "requires that commonalities and contrasts among corresponding sets of parallel relations and entities be organized" (p. 126) and he presented "Parallel relations" and "Exemplification" as two sub-sets of "Resemblance" and alluded to other subsets as "Contrast, Generalization, Exception, and Elaboration" (p. 127). "Contiguity" is explained through "Occasion relation" (p. 127), and "Cause-Effect" includes the relations of "Result, Explanation, Violated Expectation, and Denial of Preventer" (p. 126). (For a detailed discussion of coherence relations with examples, see Kehler, 2022, pp. 124-128).

Literature Review

Halliday and Hasan's (1976) theory of cohesion lays the foundation of many studies. Bahaziq (2016) examined the different cohesive devices found in one student's writing. Grammatical cohesion was found to be more frequent than lexical cohesion, and it was recommended to give the student more practice on cohesion to improve essay writing.

Coherence has also been the subject of recent research. For example, Siekmann et al. (2022) designed rubrics to assess students' "narrative" and "argumentative" essays regarding the "structure and coherence" of these essays to aid learners and teachers in developing coherent writing. Burke et al. (2023) used computational tools to examine "semantic coherence" in the speech of people with "Alzheimer's disease" and healthy people. Results indicated no significant differences in local coherence between the two groups, but people with Alzheimer showed less global coherence in their speech, which implies that computational tools can help in the clinical diagnosis of people with Alzheimer. Jiang and Niu (2022) examined discourse coherence in "neural machine translation" and human translation of Chinese literary texts into English by applying the "theme-rheme structure" in the systemic functional linguistics framework. Findings revealed that the product of "neural machine translation" is incoherent since it misses links to preceding propositions.

Since its emergence in November 2022, the applications of ChatGPT to different fields have been the focus of a plethora of research. For instance, Hill-Yardin et al. (2023) examined the ability of ChatGPT to write scientific research. The result was that it provided fake references with "shallow", "generic" "robotic" language (p. 152). Lund (2023) explores the technology behind ChatGPT. Eke (2023) discussed

how ChatGPT could be a threat to academic integrity and called for an unexpensive tool that can detect plagiarism and differentiate between AI work and human work. Alser and Waisberg (2023) called for not using ChatGPT in academic medical research without acknowledging it paying attention to plagiarism, since when detecting plagiarism in some of ChatGPT writing, it was found to be copied from the internet and also some of the sources written by ChatGPT are fake references that do not exist. On the other hand, Javaid et al. (2023) discussed the importance of using ChatGPT in education since it can help teachers and students and save their time and effort; however, they did not tackle the issue of plagiarism or suggested an altering method of using ChatGPT that enriches academic integrity. Thus, no research to date has examined any linguistic aspect in essays generated by ChatGPT or compared these essays to human ones linguistically, the focus of the current study.

Research Questions

The current study poses two research questions:

1. How are cohesion and coherence manifested in essays produced by ChatGPT?
2. How do these same language features resemble or differ from essays written by students?

To answer these questions, essays written by students, and essays generated by ChatGPT on the same topics are analyzed in the following section.

Methodology
Data and Procedures

The data used in this study constitute essays from fourth year Comparative Literature course representing the most advanced level, among all courses, in essay writing in the English Department at one of the Egyptian private universities, and essays from first year Writing course designating the beginning level of essay writing in the same Department. Essay writing, in these courses, is an essential component and a compulsory question in the final exam.

At a very initial step, ChatGPT was asked to generate a research consent form which was modified as follows and sent to students of the fourth and the first years through their teachers to grant the researcher permission to use their essays for research purposes. All students responded positively to the consent form.

Research Consent Form

I hereby grant permission to [Researcher's Name] to use my essay titled [title of the essay] for research purposes. I understand that my essay will be used solely for academic research and analysis in its entirety or in part, and it may be quoted, paraphrased, or referenced in academic publications or presentations. I acknowledge that my participation is voluntary. I agree that the research findings may be published or presented, but my personal information will remain confidential and will not be disclosed. (ChatGPT, personal communication, June 8, 2023)

A prompt from the fourth-year course was provided to the researcher by the teacher of the course. The topic was "A Comparative Reading between Hamida in Mahfouz's *Midaq Alley* and Hester in Hawthorne's *The Scarlet Letter*". Another prompt from the first-year course was also given to the researcher by the course teacher. The prompt was "Write a compare and contrast essay on Apples and Oranges". These prompts were the last prompts given to students at the end of the semester just before their final exams after they have been trained throughout the semester on how to write essays successfully.

ChatGPT was asked to write essays on the same prompts of the two University courses. ChatGPT generated one advanced essay on each prompt. Nevertheless, ChatGPT was asked to simplify the "Apples and Oranges" essay to comply with the beginning level of the students in the Writing course. It was also asked to write two body paragraphs: one on similarities and one on differences to correspond to the same schematic structure of essays written by students in the Writing course. Yet, no instructions were given to ChatGPT on the Literature prompt regarding in text citation, indentation or capitalization rules. It was posted on ChatGPT with the same wording provided by the Literature tutor.

After teachers of the two courses graded students' essays according to the assessment criteria they usually follow, the highest scored essay in each of the two courses was taken to represent the human sample of the study. It is worth noting that the highest score essay is the one that is written by the first rank student who always gets the highest score in essays written across the semester. Thus, one essay from that student was assumed to be sufficient as a preliminary step in comparing a human essay to a ChatGPT essay especially that no research to date has investigated this matter to the best of the researcher's knowledge.

Teachers of the two courses were then asked to assess the essays generated by ChatGPT according to the same marking criteria without knowing that these essays were a ChatGPT product to avoid any

prejudice for or against. The marking criteria followed in the literature course assigns 60% of the grade to content and 40% to language, and the marking criteria in the Writing course adopts The International English Language Testing System (IELTS) band descriptors that covers "task response, coherence and cohesion, lexical resources and grammatical range and accuracy" ("IELTS", 2023) giving each of these four parameters 25% of the grade. It was found out that in the literature course, the essay generated by ChatGPT got the same grade, which was 9 out of 10, exactly as the highest grade given to the best student in the course, and the teacher asserted that the ChatGPT essay did not get 10 because it lacks intext citation, while the human essay lacks some supporting ideas. However, the grade given to the essay generated by ChatGPT in the "Apples and Oranges", beginning level, prompt was 8 out of 10 which surpassed the highest grade given to the best student in the Writing course; namely, 7 out of 10, since, according to the grader, both essays lack some supporting materials, but this deficiency is more in the human essay than in the ChatGPT essay. Thus, ChatGPT can produce advanced human like essays in an advanced literature course, and its simplified beginning level essay outdid that of the highest score essay in the Writing course.

In an attempt to utilize ChatGPT in the qualitative analysis of the essays, the researcher uploaded the human essay and the ChatGPT essay of each course as a completely new topic on ChatGPT and asked ChatGPT to analyze the essays in terms of cohesion (grammatical and lexical) and coherence (link between propositions, schematic structure and global coherence) without defining any of the terms. ChatGPT was able to analyze the essays regarding these linguistic features, which means that as a model it is trained to produce cohesive and coherent texts. Here is a sample of ChatGPT analysis of Essay 1 (see Appendix) in terms of grammatical and lexical cohesion:

Pronouns: The use of pronouns such as "they," "both characters," and "each" helps maintain coherence by referring back to previously mentioned subjects. This allows for smoother transitions between ideas and keeps the focus on the characters and their actions.
Connectives: The essay employs connectives such as "on the other hand," "though," "in both works," and "despite" to establish relationships between contrasting ideas or introduce new points. These connectives help signal shifts in focus or highlight similarities and differences between the characters and their actions.
Conjunctions: The use of conjunctions like "and," "but," and "moreover" helps connect related ideas and create a logical flow of information. They assist in presenting supporting evidence or introducing additional details that reinforce the main arguments.
Demonstrative pronouns: The essay utilizes demonstrative pronouns like "this," "these," and "that" to refer back to specific concepts or ideas mentioned earlier. They help tie different sections of the essay together and provide clarity by pointing to specific elements being discussed.

(ChatGPT, personal communication, June 19, 2023)

Synonyms: The essay employs synonyms to avoid repetitive language and introduce variety. For example, "representation" is used as a synonym for "portrayal," "predicament" is used instead of "situation," and "retaliation" is used in place of "revenge." The use of synonyms helps maintain reader interest and adds depth to the essay.

Repetition: The essay repeats certain keywords and phrases to reinforce ideas and create a sense of cohesion. Examples include the repetition of terms like "social and religious norms," "predicament," "consequences," and "defied." The repetition of these terms strengthens the essay's main points and ties the different sections together.

Parallelism: The essay employs parallel structures in sentences or phrases to establish connections and enhance clarity. For instance, in the sentence "She exuded majesty and grace that portrayed a strong self-confidence," the parallel structure of "majesty and grace" reinforces the idea of self-confidence. This technique helps the reader follow the essay's arguments more effectively.

(ChatGPT, personal communication, June 19, 2023)

An example of ChatGPT analysis of Essay 5 (see Appendix) regarding the "link between propositions", ChatGPT asserted that the link between propositions is clear, since each

paragraph focuses on a specific aspect: the introduction introduces the topic, the body paragraph discusses similarities, and the following body paragraph discusses differences. The use of transitional words and phrases, such as "both," "however," and "in conclusion," helps to connect and guide the reader through the progression of ideas. This creates a cohesive structure, where each paragraph builds upon the previous one, leading to a logical conclusion.

(ChatGPT, personal communication, June 19, 2023)

Concerning the schematic strategies and global coherence, the essay has

a structured format with an introduction, body paragraphs on similarities and differences, and a conclusion. This organization helps readers understand the overall framework of the essay and anticipate the information that will be presented.

The essay maintains global coherence by consistently focusing on the topic of comparing apples and oranges. The introduction introduces the fruits and sets the context, while the body paragraphs delve into the similarities and differences, respectively. The conclusion summarizes the main points and reinforces the overall idea of appreciating the unique qualities of both fruits. This consistent focus on the topic throughout the essay enhances global coherence.

(ChatGPT, personal communication, June 19, 2023)

ChatGPT was, then, asked to compare the human essay to the one it produced in each course and to state which one is better with respect to cohesion and coherence. ChatGPT did the task and showed that Essays 2 and 5 (the ones it produced, see Appendix) are much stronger than the human essays in terms of cohesion and coherence. It maintained that Essay 2, for example, "demonstrates stronger organization and coherence. Its clear introduction, systematic approach, and well-developed sections

contribute to a more cohesive and coherent presentation of ideas. The essay maintains a logical progression throughout, allowing the reader to follow the author's argument easily", but Essay 1 needs "stronger transitions, clearer topic sentences, and a more explicit roadmap of the overall argument to improve its global coherence" (GhatGPT, personal communication, June 19, 2023). Accordingly, the researcher asked ChatGPT to improve the human essay of the Writing Course and the improved version (Essay 6, see Appendix) got 9.5 out of 10 by the human course grader who deducted half a mark because the conclusion is a bit long. ChatGPT was also asked to improve the human essay of the Literature course and add intext citation and Works Cited (MLA Style) to it, but the improved version (Essay 3, see Appendix) got 8 out of 10, since, according to the grader, ideas are less developed than those found in Essay 2 (see Appendix). All the six essays, three essays from each course; one human essay getting the highest score in the course (Essays 1 and 4) and one ChatGPT essay on the same prompt (Essays 2 and 5) together with the modified improved version of the human essay processed by ChatGPT (Essays 3 and 6), represented the data of the current study.

To reinforce ChatGPT qualitative analysis and validate the human assessment of the six essays in terms of cohesion and coherence, the Comparative Literature essays and the Writing essays were quantitatively analyzed using Coh-Metrix 3.0. Coh-Metrix is a Web-based computational linguistic tool developed by the University of Memphis, and it analyses cohesion and coherence in discourse and "over 200 [other] measures of language, text, and readability" through a set of indices (Graesser et al., 2004, p. 194). The indices used in the current study are "Referential Cohesion", "Latent Semantic Analysis", "Lexical Diversity", "Connectives", "Situation Model", "Syntactic Complexity", "Syntactic Pattern Density", "Word Information" and "Readability" (Quispesaravia et al., 2016).

"Referential cohesion" measures "noun overlap", "argument overlap" and "stem overlap" between nearby sentences that create "local cohesion" and the same overlaps among all the sentences in a text creating "global cohesion" (Graesser et al., 2004, p. 199-200). "Latent semantic analysis (LSA) … also provides measures of semantic overlap between sentences or between paragraphs. LSA considers meaning overlap between explicit words and also words that are implicitly similar or related in meaning" (McNamara et al., 2011, p. 3). High cohesion is achieved when the referential cohesion and LSA are high, and their

"measures vary from 0 (low cohesion) to 1 (high cohesion)" ("Coh-Metrix version 3.0 indices", n.d.).

"Lexical diversity" has an inverse relation with cohesion. When the "lexical diversity" is high in a text, the text is to be suffering from "low cohesion". "Lexical diversity" refers to the "type-token" relationship. "Types" are the different words in a text, and "tokens" are the number of words in a text. When a word is used many times in a text, "lexical diversity" decreases and "cohesion" increases (McNamara, et al. 2011, p. 3).

This inverse relation is also at play between "the situation model" and "cohesion". The "situation model", as presented earlier, has to do with "coherence". It is the mental processing of a text that goes beyond the "explicit" wording to draw "inferences" that relate ideas inside the text to each other and to "prior knowledge" generating a "coherent" text (McNamara, et al. 2011, p. 4). Zwaan and Radvansky (1998) suggested five dimensions of the situation model to be applied to "deep comprehension: causation, intentionality (goals), time, space, and people" (as cited in Dowell, et al., 2016, p. 78). A break in coherence occurs when there is a discontinuity on one or more of these dimensions. Accordingly, the text uses cohesive devices called "particles" to signal the presence of the "discontinuity" and redeem deep cohesion. These "particles" could be "causal (e.g., *because*, *enable*), intentional (*in order to*, *so that*), and temporal (*before*, *later*)". Coh-metrix provides indices of these particles that "stitch together the actions, goals, events, and states conveyed in the text" (Graesser et al., 2011, p. 227). Consequently, when the cohesion particles in Coh-Metrix are high, it gives the meaning that the coherent situation model or deep cohesion is discontinued, and these particles are just compensating for this discontinuity.

"Connectives" are also related to the "situation model" coherence or "deep cohesion". They are a type of "particles" that link sentences and clauses to each other whenever there is a "discontinuity" in the dimensions of coherence. So, "connectives" are positively related to "cohesion", but negatively related to "coherence". Moreover, if the topic under discussion is a "less familiar" topic, cohesive devices can compensate for the difficulty (Graesser et al., 2011, p. 230).

"Syntactic complexity" has to do with the length of the sentence and whether it has embedded clauses or not. "The syntax in a text tends to be easier to process when there are shorter sentences, few words before the main verb of the main clause, and few words per noun-phrase" ("Coh-Metrix version 3.0 indices", n.d., n.p.). Closely related to this index is "syntactic pattern density". If a text "has a higher noun and verb phrase

incidence, it is … informationally dense with complex syntax" ("Coh-Metrix version 3.0 indices", n.d., n.p.). So, the more of these two indices, the more difficult the text is to be understood. Although "syntactic complexity" and "syntactic density" are closely related, they are measured differently. Coh-Metrix (n.d.) measures the "mean number of words before verbs and nouns" when dealing with "syntactic complexity", but it measures the "incidence score" when dealing with "syntactic density" (n.p.). An incidence score "is the number of word classes or constituents per 1,000 words" (Graesser et al., 2004, p. 197).

The most important incidence in the "word information" index is "word concreteness". Texts that include concrete rather than abstract words are easier to process in memory (Dowell, et al., 2016, p. 78). So, the higher the score, the more concrete words are in a text. "Readability" index has to do with the length of the sentence. The shorter the sentence, the higher the "readability" score. Readability is "a number from 0 to 100, with a higher score indicating easier readability" ("Coh-Metrix version 3.0 indices", n.d., n.p.). In what follows, the results of Coh-Metrix analysis of these indices are presented and discussed.

Results and Discussion

Coh-Metrix version 3.0 focuses mostly on coherence. Eight of its ten indices consider coherence in a text and only two indices, "local cohesion" and "connectives" are devoted to cohesion. Table 1 shows the score of each index in the Literature essays in relation to cohesion and coherence, and Table 2 presents the same indices in the Writing essays. Some indices are directly related to cohesion/coherence; others are inversely related to cohesion/coherence. The best score of indices that are directly/positively related to cohesion/coherence are highlighted in yellow, and the best score of indices that are indirectly/inversely related to cohesion/coherence are highlighted in green. As discussed earlier, Essay 1 and Essay 4 are the human essays, Essays 2 and 5 are the ChatGPT essays, and Essays 3 and 6 are the modified human essays by ChatGPT.

Table 1
Coh-Metrix Analysis of Literature Essays

Essay #	Local Cohesion	Global Cohesion	LSA	Lexical Diversity	Situation Model	Connectives	Syntactic Complexity	Syntactic Density	Concreteness	Readability
Essay 1	0.293333	0.208667	0.1921	0.745	0.344	103.809998	3.3955	176.1903	363.3970032	50.20199966
Essay 2	0.296	0.298667	0.252	0.628[b]	0.083	97.1989975	4.1865	180.8075	406.8739929	21.24399948
Essay 3	0.359667[a]	0.2645	0.2263	0.733	0.286	109.123001	4.059	177.0663	393.3529968	28.72900009

Note. The best score of each index in the three Literature essays is highlighted.
[a] Yellow cells reflect the direct relationship between Coh-Metrix indices and cohesion/coherence.
[b] Green cells reflect the inverse relationship between Coh-Metrix indices and cohesion/coherence.

Table 1 indicates that Essay 2 (ChatGPT essay) surpasses in six indices (60%) of the total number of indices whether these indices are directly or indirectly related to cohesion/coherence. All six indices are related to "coherence" which has to do with the cognitive abilities of the producer to organize a text and relate the meaning of its propositions to one another. This discloses the fact that ChatGPT manages to produce essays that excel human essays regarding not only coherence ("global coherence, situation model and connectives", but also lexical choices ("LSA, lexical diversity and concreteness").

Essay 1 (the human essay) exceeded in three indices: "syntactic complexity, syntactic density and readability". However, all three indices are related to producing simple, short sentences that facilitate understanding, which does not reflect highly proficient. Consequently, when reexamining the results of "syntactic complexity, syntactic density and readability" presented in Table 1, it became obvious that Essay 2 (the ChatGPT essay) is characterized by high "syntactic complexity", high "syntactic density" and low "readability", denoting the high writing proficiency level of essays produced by ChatGPT, which complies with McNamara et al. (2010) assertion that complex structures "may result in text that is difficult to process, [but] it … may be reflective of more sophisticated, skilled language production" (p. 63).

Only one index, "local cohesion", surpassed in Essay 3 (ChatGPT modified human essay). On a deeper look at all the indices in Table 1, it would be evident that in all the indices, Essay 3 occupies the second rank except for "local cohesion" and "connectives" where it occupies first in "local cohesion" in a direct relationship with cohesion, and third in "connectives" in an indirect relationship with coherence. Essay 3 is a modification of Essay 1. ChatGPT fixed all the drawbacks of Essay 1 to come out with a more proficient essay from which students can learn and avoid similar mistakes in the future. It works on Essay 1 and refined it. Hence, it is in between Essay 2 and Essay 1. It is not up to the level of Essay 2, and it is better than Essay 1. Regarding the indices of "local

cohesion" and "connectives", ChatGPT pinpointed a break in coherence in Essay 1, which was clear when it was asked to compare Essays 1 and 2 (see above). Accordingly, it inserted words overlaps as "local cohesion" and "particles" as "connectives" to compensate for the break. This is in accordance with ChatGPT's capability in producing "personalized learning experiences … [that] can help students improve their academic performance …" (Ray, 2023, p. 137).

All that has been presented indicates that the literature essay produced by ChatGPT is the best essay of the three. The human grader gave this essay 9 out of 10, the same grade given to Essay 1. However, it is believed that when humans evaluate students' work, they are influenced by what they present in class, and part of their evaluation measures students following their instructions in achieving the task. Accordingly, since ChatGPT has not received any instructions; yet it produced an essay that meets the criteria of the human grader, it can be approved to be the best essay of the three. Furthermore, Essay 3 received the least human grade, which agrees with the break in coherence demonstrated by Coh-Metrix analysis.

Table 2 presents the results of the Writing essays. It is worth mentioning that Essay 5 is a simplified version of the original essay produced by ChatGPT, as aforementioned. Nevertheless, it excelled in 6 indices (60%) of Coh-Metrix as shown in Table 2, which suggests that albeit its simplicity and shortness (see Appendix), Essay 5 managed to exceed Essay 4 (the human essay) in most of the indices.

Table 2
Coh-Metrix Analysis of Writing Essays

Essay #	Local Cohesion	Global Cohesion	LSA	Lexical Diversity	Situation Model	Connectives	Syntactic Complexity	Syntactic Density	Concreteness	Readability
Essay 4	0.22233	0.191	0.2067	0.635	0.375	123.077003	1.467	188.462	426.2669983	71.98500061
Essay 5	0.39583[c]	0.332	0.3639	0.489[d]	0.143	150.942993	1.518	167.453	418.5050049	63.15800095
Essay 6	0.24583	0.26333	0.2546	0.613	0.412	154.255005	2.2025	171.542	440.7669983	43.04299927

Note. The best score of each index in the three Writing essays is highlighted.
[c] Yellow cells reflect the direct relationship between Coh-Metrix index and cohesion/coherence.
[d] Green cells reflect the inverse relationship between Coh-Metrix index and cohesion/coherence.

Table 2 shows that the scores of Essay 4 are the best in "connectives", "syntactic complexity" and "readability". Essay 4 is a short essay that contains mostly short, simple sentences, a few complex clauses and some pronouns to link ideas together (see Appendix), which "facilitates reading speed and comprehension" (Graesser et al., 2011, p. 226), but this does not reflect "high proficient writing" which is more in Essay 5 and most in Essay 6. Having these three indices the highest in Essay 6 together with "concreteness" implies that no matter to what

extent a student's writing exhibits beginning or elementary writing level, ChatGPT can upgrade it by adding "connectives" to redeem any flaw in coherence, by using more complex embedded clauses to elevate the level of syntax, which affects "readability", since there would be more load on the working memory (McNamara et al., 2010, p. 64). At the same time, ChatGPT uses more concrete words in Essay 6 than those produced in Essays 4 and 5, which complies with the fact that beginning level students tend to use more abstract than proficient students (Graesser et al., 2011, p. 231). Consequently, Essay 6 has the best scores in four of the indices: "Connectives", "syntactic complexity", "concreteness" and "readability". The remaining six indices are best, as mentioned earlier, in Essay 5. So, whether the essay is a ChatGPT product (Essay 5) or a modification of a student essay (Essay 6), it came up to be better than the human product. This is in accordance with the human grading that allotted the least grade to the human essay (7 out of 10). It is worth noting that Essay 6 is written in a more academic, formal style than Essay 5 because it has Essay 4, where students are taught academic writing, as an input; thus, ChatGPT processed the same style and came up with Essay 6, which again proves the ability of ChatGPT to help in improving student's level through "personalized learning" (Ray, 2023, p. 137), and for this reason, the human grader gave Essay 6 the highest score, 9.5 out of 10.

All that has been presented responds to the research questions upon which this study was based. It shows how cohesion and coherence are manifested in the essays produced by ChatGPT in two different courses with two different levels of students, and whether these essays are similar or different from human essays.

Conclusion

The purpose of this study was to investigate cohesion and coherence in essays produced by ChatGPT using Halliday and Hasan's (1976) theory of cohesion and van Dijk and Kintsch's (1983) cognitive situation model. It raised two research questions: how cohesion and coherence are deployed in essays generated by ChatGPT and how these same language features are exploited in human essays on same Literature and Writing courses topics. Data was assessed by the courses' instructors, qualitatively analyzed using ChatGPT and quantitively analyzed using Coh-Metrix version 3.0. Results indicated that the essay produced by ChatGPT in the Literature course got the highest score, and the essay modified by ChatGPT in the Writing course got the highest score. This result might challenge the integrity of any task assigned to students, since students might ask ChatGPT to do the task for them and submit it as their work. Therefore, some recommendations are proposed
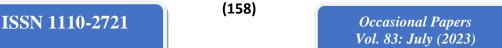
here to make use of ChatGPT without scarifying the honesty and reliability of the learning process.

In Writing courses, students might be asked, instead of writing an essay at home, to ask ChatGPT to provide them with at least three essays on a certain topic on different levels: beginning, intermediate and advanced. The teacher just tells them the topic and that they have to bring three essays on three language levels. Students will get different essays depending on the language level they use while chatting with ChatGPT. They study these essays thoroughly according to a criterion explained to them in class. This criterion might be the rubrics for assessment, for example. They could memorize syntactic structures or vocabulary from ChatGPT essays, since it presents authentic passages. After that, they bring all these essays to class, discuss them with their peers under the teacher's monitoring, and submit all the work they got, and then sit for a quiz to write an essay in class on the same topic.

In Literature courses, students might be asked to bring contemporary versions of classic novels or plays from ChatGPT and study the new version thoroughly at home. Then, in class, they can compare this version to the original version in an essay form as an open book quiz. In any content course, students ask ChatGPT about a specific concept in a course and come to class to present it. They might use other AI tools like "Midjourney" or "slidego" to help them in preparing a presentation on the concept they were assigned to. In translation courses, ChatGPT can provide students with the translation, and they can compare what they get, evaluate it and decide on the best translation.

Consequently, many skills will be developed and enhanced. Students will learn how to be focused by learning how to give ChatGPT specific instructions to provide them with what they actually need. Students will depend on themselves for getting the required information; thence, autonomous learning will develop. Students' language competence and performance will develop since they are exposed to authentic language while chatting with ChatGPT.

These are but a few suggested applications of ChatGPT that could enhance the learning process without scarifying honesty and integrity and fearing plagiarism. ChatGPT has just emerged, and much linguistic research is still needed to explore it. Different cognitive abilities could be empowered with the use of AI, and cognitive linguistic research is one of the fields that could unleash this power. The present study is just a prior work in the field, and it is hoped that more studies could address the same topic.

References

Alser, M., & Weisberg, E. (2023). Concerns with the Usage of ChatGPT in Academia and Medicine: A Viewpoint. *American Journal of Medicine Open, 9*, 1-2. https://doi.org/10.1016/j.ajmo.2023.100036

Bahaziq, A. (2016). Cohesive devices in written discourse: A discourse analysis of a student's essay writing. *English Language Teaching, 9*(7), 112-119. http://dx.doi.org//10.5539/elt.v9n7p112

Burke, E., Gunstad, J., & Hamrick, P. (2023). Comparing global and local semantic coherence of spontaneous speech in persons with Alzheimer's disease and healthy controls. *Applied Corpus Linguistics, 3*, 1-6. https://doi.org/10.1016/j.acorp.2023.100064

Coh-Metrix version 3.0 indices. (n.d.). Coh-Metrix Memphis. Retrieved May 19, 2023, from http://cohmetrix.memphis.edu/cohmetrixhome/documentation_indices.html#Referential

Cook, B. (2023, March 7). *Revolutionize Your English Language Teaching: 5 Things You Can Do Differently Using ChatGPT* [Webinar]. British Council EnglishScore. https://5907605.fs1.hubspotusercontent-na1.net/hubfs/5907605/ChatGPT%20webinar.mp4?_hsmi=249318084&_hsenc=p2ANqtz--Ec_AB1DWAxl_ctmab9LfJ7ZsfmJOCHVybRQgnfPsd0wy2oXDLC-QUSJwxNlOZSh5KuQzJNIU2-pojl3elS1mcAq59tg

Dowell, N. M., Graesser, A.C., & Cai, Z. (2016). Language and discourse analysis with Coh-Metrix: Applications from Educational material to learning environments at scale. *Journal of Learning Analytics, 3*(3), 72–95. http://dx.doi.org/10.18608/jla.2016.33.5

Eke, D. O. (2023). ChatGPT and the rise of generative AI: Threat to academic integrity? *Journal of Responsible Technology, 13*, 1-4. https://doi.org/10.1016/j.jrt.2023.100060

Graesser, A.C., McNamara, D. S., & Kulikowich, J. M. (2011). Coh-Metrix. *Educational Researcher, 4*(5), 223-234. http://doi:10.3102/0013189X11413260

Graesser, A.C., McNamara, D. S., Louwerse, M.M., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36*(2), 193-202. Retrieved June 20, 2023, from https://link.springer.com/article/10.3758/bf03195564

Halliday, M.A.K., & Hasan, R. (1976). *Cohesion in English*. Longman.

Hill-Yardin, E., Hutchinson, M. R., Laycock, R., & Spencer, S. J. (2023). A Chat(GPT) about the future of scientific publishing. *Brain, Behavior, and Immunity, 110*, 152-154. https://doi.org/10.1016/j.bbi.2023.02.022

IELTS. (2023). IDP IELTS. Retrieved June 20, 2023, from https://ielts.idp.com/results/scores/writing.

Javaid, M., Haleem, A., Singh, R.P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through ChatGPT Tool towards

ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations 3*, 1-12. https://doi.org/10.1016/j.tbench.2023.100115

Jiang, Y., & Niu, J. (2022). How are neural machine-translated Chinese-to-English short stories constructed and cohered? An exploratory study based on theme-rheme structure. *Lingua, 273*, 1-20. https://doi.org/10.1016/j.lingua.2022.103318

Kehler, A. (2022). Coherence establishment as a source of explanation in linguistic theory. *Annual Review of Linguistics, 8*, 123-142. https://doi/10.1146/annurev-linguistics-011619-030357

Lund, B. D. (2023). A brief review of ChatGPT: Its value and the underlying GPT technology. https://doi/10.13140/RG.2.2.28474.06087/1

McNamara, D. S., Crossley, S. A., & McCarthy, P. M. (2010). Linguistic features of writing quality. *Written Communication, 27*(1), 57-86. https://doi/10.1177/0741088309351547

McNamara, D., Graesser, A.C., Cai, Z., & Kulikowich, J. M. (2011). Coh-Metrix Easability Components: Aligning Text Difficulty with Theories of Text Comprehension. *AERA*, 1-12. Retrieved June 20, 2023, from https://www.researchgate.net/publication/228455723.

Quispesaravia, A., Perez, W., Cabezudo, M., & Alva-Manchego, F. (2016). Coh-Metrix-Esp: A Complexity Analysis Tool for Documents Written in Spanish. In N. Calzolari, K. Choukri, T. Declerck, S. Goggi, M. Grobelnik, B. Maegaard, J. Mariani, H. Mazo, A. Moreno, J. Odijk, & S. Piperidis (Eds.), *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* (pp. 4694-4698). European Language Resources Association (ELRA). https://aclanthology.org/volumes/L16-1/

Ray, P. P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems, 3*, 121-154. https://doi.org/10.1016/j.iotcps.2023.04.003

Reinhart, T. (1980). Conditions for text coherence. *Poetics Today, 1*(4), 161-180. https://doi.org/10.2307/1771893

Siekmann, L., Parr, J. M., & Busse, V. (2022). Structure and coherence as challenges in composition: A study of assessing less proficient EFL writers' text quality. *Assessing Writing, 54*, 1-13. https://doi.org/10.1016/j.asw.2022.100672

Tapiero, I. (2007). *Situation Models and Levels of Coherence: Toward a Definition of Comprehension* (1st ed.). Routledge. https://doi.org/10.4324/9780203810040

van Dijk, T.A. (1985). Cognitive situation models in discourse production: The expression of ethnic situations in prejudiced discourse. In J. P. Forgas

(Ed.). *Language and social situations. Springer series in social psychology* (pp. 61-79). Springer. https://doi.org/10.1007/978-1-4612-5074-6_4

van Dijk, T. A., & Kintsch, W. (1983). *Strategies of discourse comprehension*. Academic Press. Retrieved May 19, from https://discourses.org/wp-content/uploads/2022/06/Teun-A-van-Dijk-Walter-Kintsch-1983-Strategies-Of-Discourse-Comprehension.pdf

Appendix
Comparative Literature Essays

Essay 1 (Human Essay)

A Comparative Reading between Hamida in Mahfouz's Midaq Alley
and Hester in Hawthorne's The Scarlet Letter

Midaq Alley is a novel, written by the Egyptian Nobel Laureate Naguib Mahfouz. It serves as a representation of Egypt with its multitude of social classes and cultural backgrounds during the era of the British occupation. On the other hand, The Scarlet Letter is written by the American writer Nathaniel Hawthorne. It gives insightful glimpses on several issues that were rooted in America such as immigration, puritanism, and hypocrisy. Midaq Alley's Hamida is the leading female figure, whom the plot is centered around. Being an orphan, she grew up with her foster mother "Umm Hamida", but she lacks proper parental guidance. On the other side of the spectrum, Hester is an adulteress who immigrated from Holland to America, to end up settling down in Boston amongst the puritans. Though both characters shamelessly defied the religious and social norms of their societies and did not feel guilty about their acts, they reacted differently to the outcomes of the risky state they got themselves into.

In both works, both characters went against the social and religious rules of their society and showed confidence while doing so. In The Scarlet Letter, Hester carried the illegitimate baby in her hands and walked like a queen and did not care about how people looked at her. She was first described as having a "natural dignity and power of character," with particular attention paid to the haughty smile and commanding gaze that show no self-consciousness of her predicament. Her behavior was described as "haughty', but she may be in anguish, as if her heart had been cast into the street for them all to spurn and tread upon, but her face did not reflect that. She exuded majesty and grace that portrayed a strong self-confidence. In Miadq Alley, Hamida was already autonomous because she went for solo walks at a period when Muslim women were not permitted to do so, and she wanted to break the cycle of sacrificing her own happiness. No one was allowed to go alone during World War II, but Hamida wanted to disobey the expectation that women should be fearful and composed. In Midaq alley, When Umm Hamida asks "Have you forgotten that you are engaged? And that I confirmed it by reading the Qur'an with Abbas?" Hamida "shouted in full, angry scorn, "that barber!". Hamida violated the Qur'an by offering herself to another man while being engaged, but she did not care or feel guilty. Furthermore, Hamida insisted on her decision to get a better life by marrying a wealthy man, and she became a prostitute to get a wealthy life without a slight sense of shame.

Even though both characters acted against the religious beliefs of their societies and did not follow the prevalent norms, they grappled with their fallouts dissimilarly. Having no parents or opportunities for nepotism and amidst the darkness of the alley, Hamida saw in Ibrahim Faraj a sanguine hope that she could take advantage of, to have a better future and unbounded freedom. Little did she know that she was the one who is taken advantage of, manipulated, and exploited in the most degrading way possible. She found it unfathomable and could not wrap her head around the fact that she eluded a prison, the alley, only to get trapped in a bigger cage, and when she finally did hold an in-depth insight of her situation, there was no coming back. Sorrow sank in her soul as she realized that she was at the mercy of Faraj and that she had been chasing a mirage rather than a dream, "Hamida followed him inside. She now knew that retreat was impossible and that the past was completely erased. she was resigned to her fate; nevertheless, she wondered where happiness lay" (Mahfouz 150). Her feelings of humiliation and imprisonment rose as time went by, indicating the wrath of the storm to come, "Hamida still felt strangely restless and dissatisfied. Not entirely ruled by her sexual instincts, she longed for emotional power." (Mahfouz 175). Following the stage of realization, Hamida's rebellious nature started emerging on the surface; she was determined on getting all that she wanted without compromises or binding shackles. She embarked on plotting her revenge and decided to kill two birds, Abbas and Ibrahim Faraj, with one stone. Taking a closer look over her situation, she was at war with herself, steady searching for freedom and constantly running into ruckus that ruin her life more than it already was. On the contrary, Hester, when faced with the consequences of her predicament, chose to remain silent and passive. She totally submitted to her fate, shouldering the burden all alone. Instead of running away with her illegitimate daughter, she stayed and bore the humiliating punishment with complete surrender. On the scaffold, she refused to give away the name of her partner in sin. Moreover, she did not even try to convince the clergymen of mitigating her penalty as she became a widow and single mother. More than that, the child that brought her sin to light, was not perceived by her as a curse but rather as a blessing that transformed her life to the better. She did not act her anger out on the child, but fought for her custody instead, which suggests that feelings of anger and rebellion did not exist in Hester's heart, "God gave her into my keeping. I will not give her up!" (Hawthorne 98). A closer look at Hester will reveal that she kept up with severe injustice and suffering, with no intentions of retaliation in mind, only because she lacked the wild and aggressive temper that can make a way for her to revolt.

In conclusion, both characters in Midaq Alley and The Scarlet Letter did not conform to the norms of their societies and committed sinful transgressions in an almost overconfident manner. However, when the consequences of their deeds started to unfold and threaten their dignity and self-perception, they acted differently, each depending on her inner feelings and long-standing aims. Despite the cultural gap between the two works, Mahfouz and Hawthorne succeed in voicing out the human conscience with its bad and good, creating universal-themed literary works.

Essay 2 (ChatGPT Essay)

Midaq Alley, written by Naguib Mahfouz, and The Scarlet Letter, penned by Nathaniel Hawthorne, are two novels that explore the consequences of sin and the societal reactions towards the sinners. While Midaq Alley delves into the lives of individuals residing in a poor Egyptian alley during the British occupation, The Scarlet Letter focuses on the puritanical society of seventeenth-century America. This essay aims to compare the characters of Hamida in Midaq Alley and Hester in The Scarlet Letter, their commitment to sin, and the divergent responses of their respective societies.

Both Hamida and Hester find themselves entangled in sinful acts, driven by their personalities and circumstances. In Midaq Alley, Hamida, an ambitious and attractive young woman, yearns for a life of luxury and despises the poverty-stricken existence in the alley. Her pursuit of material wealth intensifies under the deteriorating economic conditions of the British occupation. When Faraj, a man who promises her a better life, deceives her, Hamida succumbs to becoming a prostitute, an unforgivable sin in both traditional and religious contexts. Similarly, in The Scarlet Letter, Hester, a beautiful and young woman, finds herself abandoned by her husband in the puritanical society of Boston. Loneliness and a lack of affection drive her into the arms of another man, resulting in her committing adultery, a grave transgression in the strict puritan community.

Mahfouz and Hawthorne present their sinners, Hamida and Hester, in contrasting manners. In Midaq Alley, Hamida is depicted as a sinner whose actions are perceived as heinous throughout the novel. Even Abbas, the only one who truly loves her, denounces her as a "shameless criminal." The alley's society vehemently condemns Hamida for eloping with a man, illustrating the unrelenting judgment and rejection she faces. This is further exemplified in the tragic ending, where Hamida is physically attacked, leading to Abbas's demise. Conversely, in The Scarlet Letter, Hester's portrayal evolves from initial condemnation to gradual acceptance. Despite the initial harsh judgment, including the public humiliation of wearing a scarlet letter "A" for adultery, the puritan society begins to empathize with Hester's plight over time. Even her estranged husband, Chillingworth, forgives her and recognizes their shared responsibility for the transgression. The society, represented by Chillingworth, views Hester as a victim rather than a criminal, leading to her eventual glorification.

The societal reactions to the sinners in the two novels differ significantly. In Midaq Alley, the Egyptian society ostracizes and condemns Hamida after her return as a prostitute, considering her a criminal who has violated societal norms. This rejection culminates in a violent incident, highlighting the society's refusal to forgive her. Conversely, in The Scarlet Letter, the puritan society initially condemns Hester for her adultery, acknowledging the severity of her sin. However, over time, they display a surprising capacity for empathy and forgiveness. The society gradually embraces Hester, recognizing her resilience and personal growth. The meteor with the scarlet letter "A" in the sky further reinforces her transformation, ultimately leading to her elevated status as a charitable and sacred woman.

In conclusion, the characters of Hamida in Midaq Alley and Hester in The Scarlet Letter both commit sins that are deemed unforgivable by their societies. However, the authors' portrayals and the societal reactions differ significantly. Mahfouz depicts Hamida as a sinner who faces constant condemnation, while Hawthorne presents Hester's journey from condemnation to acceptance. The contrasting reactions of the Egyptian and puritan societies emphasize the societal attitudes towards sin and their capacity for forgiveness. Through the divergent portrayals of sinners and societal responses, both novels provide insights into the complexities of human nature and the impact of societal judgment on individuals. (ChatGPT, personal communication, May 28, 2023)

Essay 3
Improvement of the human essay (Essay 1) by ChatGPT with documentation MLA style

A Comparative Reading between Hamida in Mahfouz's Midaq Alley and
Hester in Hawthorne's The Scarlet Letter

Midaq Alley, written by the Egyptian Nobel Laureate Naguib Mahfouz, provides a vivid representation of Egypt's social classes and cultural diversity during the era of British occupation (Smith 45). Similarly, The Scarlet Letter, penned by the American writer Nathaniel Hawthorne, delves into the issues of immigration, puritanism, and hypocrisy in seventeenth-century America (Jones 23). This essay aims to compare the characters of Hamida in Midaq Alley and Hester in The Scarlet Letter, focusing on their defiance of societal norms and contrasting responses to the consequences they face.

Both Hamida and Hester exhibit a resolute confidence as they challenge the social and religious rules of their respective societies. In The Scarlet Letter, Hester fearlessly carries her illegitimate baby, displaying a dignified and commanding presence (Hawthorne 87). Despite the shame and judgment cast upon her, she exudes grace and self-assurance, as if her heart remains untouched by the public's disdain. Likewise, in Midaq Alley, Hamida demonstrates her autonomy by defying societal expectations that restrict the freedom of Muslim women. She walks alone during World War II, refusing to be confined by fear and conformism. When confronted by Umm Hamida about her engagement, Hamida boldly dismisses her suitor, further challenging societal conventions without remorse (Mahfouz 56). Her pursuit of a wealthy life leads her to become a prostitute, displaying an unyielding determination to escape her impoverished circumstances.

While both characters transgress the religious beliefs and social norms of their societies, their responses to the consequences of their actions differ significantly. Hamida, orphaned and desperate for a better future, sees Ibrahim Faraj as her savior, offering her a chance at unbounded freedom (Mahfouz 123). Unbeknownst to her, she falls victim to manipulation and exploitation, finding herself trapped in a different form of confinement. As she realizes the futility of her pursuit, sorrow engulfs her, and she yearns for emotional fulfillment amidst the humiliation and imprisonment she experiences (Mahfouz 145). Her rebellious nature emerges, leading her to plot revenge and seek a path to freedom, ultimately descending into a cycle of self-destruction (Mahfouz 168).

In contrast, Hester, when faced with the consequences of her sin, chooses a path of silent acceptance and endurance. She willingly bears the burden of her adultery, refusing to reveal the name of her partner in sin on the scaffold (Hawthorne 105). Despite the public humiliation and scorn, she embraces her daughter as a blessing rather than a curse. Hester's resilience and lack of anger and rebellion are evident in her unwavering commitment to her child's well-being and her determination to protect her. Her journey is one of personal growth and transformation, as she gains the empathy and recognition of the puritan society, eventually being regarded as a sacred and charitable woman (Hawthorne 130).

In conclusion, both Hamida in "Midaq Alley" and Hester in "The Scarlet Letter" defy the norms of their societies and engage in sinful acts with confidence. However, their responses to the consequences of their actions diverge significantly, with Hamida's rebellion leading to self-destruction and Hester's acceptance leading to personal growth and societal recognition. Through these contrasting portrayals, "Midaq Alley" and "The Scarlet Letter" shed light on the complexities of human nature and the varying societal attitudes towards sin (Smith 78; Jones 42).

Works Cited

Hawthorne, Nathaniel. The Scarlet Letter. Penguin Classics, 2016.

Jones, Sarah. "Exploring Puritanism in 'The Scarlet Letter'." Literary Analysis Journal, vol. 25, no. 3, 2014, pp. 21-45.

Mahfouz, Naguib. Midaq Alley. Anchor, 1992.

Smith, John. "Class and Identity in Mahfouz's 'Midaq Alley'." Journal of Middle Eastern Literature, vol. 10, no. 2, 2005, pp. 35-52.
(ChatGPT, personal communication, June 19, 2023)

Apples and Oranges Essays

Essay 4 (Human Essay)
Apples and Oranges

Fruits are the most delicious-healthy desserts anyone can have. There are many kinds of fruits. The most well known in every country are apples and oranges. There and some similarities and differences between them.

Apples and oranges share some features together. Both fruits are healthy and nutritious. They contain a lot of minerals that every human needs. Apples are similar to oranges in growing on trees. They come after white flowers. Apples and oranges have a sweet taste and also can be made into fresh juices.

However, they are different from each other in many things. Oranges have one color which is orange, while apples can be red, green or yellow. The type of skin also differs. The skin of oranges is rough and cannot be eaten, whereas the skin of apples is smooth and can be eaten. Oranges are grown in warm climates. On the other hand, apples can only be grown in cool climates.

Finally, apples and oranges are the most popular fruits around the world. They are different but both are essential to human's health. Therefore, they should be found in every home due to their benefit.

Essay 5 (Chat GPT Essay)
Apples and Oranges

Apples and oranges are two fruits that have similarities and differences. Let's explore how they are alike and how they are different.

Apples and oranges are both fruits that are good for us. They have vitamins and minerals that help keep our bodies healthy. They also have dietary fiber, which helps our digestion and makes us feel full. Another similarity is that both fruits have a lot of water in them, which helps to keep us hydrated.

Apples and oranges look and taste different. Apples come in different colors like red, green, and yellow, while oranges are always orange. Apples can be round or slightly oval-shaped, while oranges are round. Apples have a crunchy texture and different flavors, like sweet or a little sour. Oranges have a soft and pulpy texture and taste sweet and tangy. Oranges are also juicier than apples.

In conclusion, apples and oranges have both similarities and differences. They are both healthy fruits with vitamins, minerals, and fiber. However, they look and taste different. Apples come in different colors and have a crunchier texture, while oranges are orange, have a soft texture, and are juicier. Whether we prefer the crispness of apples or the juiciness of oranges, both fruits offer unique qualities to enjoy. (ChatGPT, personal communication, June 19, 2023)

Essay 6
Improvement of the human essay (Essay 4) by ChatGPT

Comparing Apples and Oranges

Fruits come in various types, and two of the most popular ones are apples and oranges. While they share some similarities, there are also notable differences between them. This essay will explore the commonalities and contrasts between apples and oranges, highlighting their distinct characteristics and the reasons why they are beloved by many.

Apples and oranges have several similarities. First and foremost, both fruits are packed with essential nutrients, making them healthy choices. They provide vitamins, minerals, and fiber that are beneficial for our bodies. Additionally, both apples and oranges are commonly enjoyed as fresh snacks or used to create refreshing juices. They offer a delightful balance of sweetness and tanginess, satisfying our taste buds. Overall, these shared qualities make apples and oranges appealing options for those seeking nutritious and delicious fruits.

Despite their similarities, apples and oranges also possess distinctive characteristics. One noticeable difference lies in their appearance. Oranges have a vibrant orange color, while apples can be found in various shades such as red, green, or yellow. Furthermore, the texture of their skins sets them apart. Oranges have a rough, pebbly skin that is unsuitable for consumption, while apples feature a smooth and edible skin. Another significant contrast is their preferred climate for growth. Oranges thrive in warm climates, whereas apples flourish in cooler regions. These differences in color, skin texture, and preferred growing conditions make each fruit unique in its own way.

In conclusion, apples and oranges, though both fruits, possess both similarities and differences. They are both healthy and nutritious, providing valuable vitamins and minerals. Both fruits are enjoyed for their taste and versatility, whether eaten as a snack or transformed into refreshing juices. However, they differ in terms of appearance, with oranges displaying a vibrant orange color and apples exhibiting a range of colors. The texture of their skins also sets them apart, as oranges have rough skin while apples have smooth, edible skin. Additionally, their preferred growing conditions differ, with oranges thriving in warm climates and apples requiring cooler environments. These distinctions contribute to the individual appeal of each fruit. Regardless of their differences, both apples and oranges bring unique flavors and health benefits to the table, making them valuable additions to any home.

(ChatGPT, personal communication, June19, 2023).