Validating Psychological Insight from Chat: An LLM-Driven Approach

Per Stark

Mon 10, June, 2024

Abstract

This study investigates the feasibility of leveraging Large Language Models (LLMs) for conducting nuanced psychological analysis of chat messages, focusing on work/educational-related stress. We analyze anonymized data from a Discord chat group of students in an educational program, collected longitudinally from fall 2022 to summer 2023. Employing the GPT-4.0 model, our analysis examines the interplay of experienced demands, support, degree of control, and their relationship to general sentiment. Our findings reveal a strong positive correlation between perceived control and sentiment and a significant positive correlation between perceived support and sentiment. Interestingly, while demands initially showed a negative correlation with sentiment, this relationship weakened upon accounting for control and support, suggesting a mitigating effect.

These results align with the established psychological demand-control-support model, indicating the validity of GPT-4.0's interpretations in this domain. This study highlights the potential of LLMs, particularly GPT-4.0, to provide valuable insights into the complex dynamics of stress, learning, and well-being in educational settings. Future research should focus on validating these findings against established psychological measures and exploring the generalizability of these results across diverse educational contexts.

Contents

1 Introduction

The field of natural language processing (NLP) has made significant strides in recent years, particularly with the advent of Large Language Models (LLMs). These models have demonstrated remarkable capabilities in various language-related tasks, such as sentiment analysis, text generation, and language translation (Vaswani *et al.*, 2017). However, the application of LLMs in psychological analysis of text data remains relatively unexplored, some preliminary studies has been made, for example using LLMs to identify suicidality risk in online posts (Chim *et al.*, 2024). Traditional approaches to analyzing psychological aspects of language have primarily relied on manual coding or rule-based systems, which can be time-consuming and labor-intensive (Tausczik and Pennebaker, 2010). This thesis aims to investigate the feasibility of using LLMs to conduct psychological analysis on chat messages, specifically focusing on work/educational-related stress.

1.1 Background

There has been significant advancements in recent years regarding NLP and classifying psychological conditions. One application of it that's perhaps reached the widest adoption is sentiment analysis. Sentiment analysis involves determining the emotional tone or attitude expressed in a piece of text, typically categorizing it as positive, negative, or neutral[1]. While sentiment analysis provides valuable insights into the overall emotional state of the writer, it may not capture the nuances and complexities of psychological constructs such as stress, control, and support. In recent years the efforts to use NLP in regards to mental health has been primarily focused on classifying documents, such as questionaires or users (Nguyen *et al.*, 2022), although new research is being done.

1.2 Demand-Control-Support Model

Stress, particularly in the context of work or education, has been extensively studied in psychology. The demand-control-support model (Karasek and Theorell, 1990) suggests that the interaction between job demands, job control, and social support plays a crucial role in determining an individual's level of stress. High demands, coupled with low control and lack of support, are associated with increased stress levels and poor psychological well being (Kristensen, 1995). Conversely, a sense of control and adequate support can buffer the negative effects of high demands which results in a less impacted, if at all, psychological well being. The current focus of study, a chat group wherein the students exchange information, support and frustration could be an interesting area to perhaps glean information about these psychological factors.

[1] https://en.wikipedia.org/wiki/Sentiment_analysis

1.3 Stress and learning

One theory that provides some insight on the relationship on stress and learning is Vygotsky's concept of Zone of Proxmial Development (ZPD) (Vygotsky and Cole, 1978). The ZPD refers to the gap between what a learner can do independently and what they can achieve with guidance and support from another, for example a teacher.

According to Vygotsky, learning occurs most effectively when students are challenged within their ZPD. This means that the tasks and material presented should be slightly beyond the learner's current abilities, creating a moderate level of stress or cognitive dissonance. This optimal level of stress is often referred to as "positive stress" or "eustress" (Sarada and Ramkumar, 2015). When learners are pushed to work within their ZPD, they are motivated to stretch their capabilities and engage in problem-solving, leading to growth and development.

It's however worth to note that too high stress levels might impact learning negatively, as well as impact psychological well being negatively. (Schwabe and Wolf, 2010) showed that learning under stress impairs memory. In summary it's important to manage the levels of stress and demands, too high and learning performance and psychological well being is impacted negatively, but too low and students will not be challenged enough and learning performance is negatively impacted.

1.4 Large Language Models

Currently the type of models that's most in focus of research regarding NLP (Chim *et al.*, 2024) are Large Language Models (LLMs) (Touvron *et al.*, 2023). LLMs are a type of deep learning model that has been trained on vast amounts of data, enabling them to *understand*[2] and generate natural language. Models such as BERT (Devlin *et al.*, 2019) or GPT (Radford *et al.*, 2019) has shown state of the art performance on many natural language processing tasks, such as text classifications, question answering and more.

Given enough diverse training data, these models perform well across many areas of interest without any additional supervision (Radford *et al.*, 2019) which offers a great tool for researchers and users that do not have to train their own model to their specific dataset.

1.5 Purpose

The primary aim of this study is to explore the feasibility of using a Large Language Model (LLM) to analyze chat messages and interpret various psychological phenomena, focusing on work/educational-related stress. By leveraging the capabilities of LLMs, this research seeks to investigate a approach to psychological

[2]There is an intriguing conversation to be had regarding understanding, to what extent does the model understand language and the world. But that is beyond the scope of the current paper.

analysis that could save time and resources compared to traditional methods, potentially enabling new avenues of investigation.

The study will specifically examine the chat messages exchanged among students in a Discord chat group, where they share experiences, questions, and concerns related to their educational program. The analysis will focus on variables such as experienced demands, support, and degree of control, and their relationship to general sentiment.

While LLMs offer promising opportunities for psychological analysis, it is important to acknowledge the potential limitations and challenges associated with this approach. These may include biases in the training data, the need for model interpretability, and the importance of validating the results against established psychological measures. As an exploratory study, this research aims to assess the feasibility of using LLMs in this context and identify areas for future refinement and validation.

In addition to the primary focus on work/educational-related stress, this study will also investigate the temporal dynamics of stress and related variables over the course of the academic year, leveraging the longitudinal nature of the data. This analysis may provide insights into how stress levels and coping mechanisms evolve over time in response to academic demands and support.

To make the findings accessible and actionable, the study will present the research results through a user-friendly visualization tool.

2 Method

This study employed a mixed-methods approach, combining natural language processing (NLP) techniques with psychological analysis. The data consisted of anonymized chat messages from a Discord chat group of students in an educational program, collected longitudinally from fall 2022 to summer 2023. The GPT-4o Large Language Model (LLM) was used to analyze the chat messages and extract variables related to experienced demands, support, and degree of control, as well as sentiment. This section will meticulously detail the methodology employed in this study, encompassing a comprehensive description of the data collection process, data cleaning and preparation steps, and the specific LLM utilized.To examine the relationships between the variables, we employed multiple regression analysis, which allowed us to control for the effects of multiple independent variables on the dependent variable, sentiment

2.1 Collection

The data for this study was collected from a Discord chat group used by students of the EC Fullstack .NET webbutvecklare program. The DiscordChatExporter, an open-source tool, was employed to extract message histories from various Discord channels, limiting the extraction to the specific channels dedicated to different courses (e.g., “html-css”, “javascript-frontend”, “javascript-frontend”, “c-sharp”, “datalagring”, “asp-net”).

2.2 Parsing and Cleaning

After exporting the message histories in HTML format, a custom Python script[3] was developed to parse and clean the data. This program utilizes the BeautifulSoup library for HTML parsing and organizes the relevant information, such as timestamps and message content, by week number. The parsed data is stored in a dictionary where the keys represent the week numbers and the values are lists of message data for each corresponding week.

During the parsing process, the script keeps track of the total number of messages and the number of messages with valid timestamps. Messages without a valid timestamp are skipped and not included in the analysis. After cleaning 67290 individual messages remained for analysis.

2.3 Analysis

The sentiment analysis of the collected and parsed Discord data was conducted using the same Python script. This script connects to the OpenAI API and employs the GPT-4o model to perform the analysis on the input text data.

For each week, the script concatenates the message content into a single text string. If the text exceeds the size limit of the model, it is split into smaller chunks. Each chunk is then sent to the OpenAI API with a specific prompt, instructing the model to assign scores (0-100) for ‘Demand’, ‘Control’, ‘Support’, ‘Sentiment’, and ‘Engagement’ based on the Demand Control Support Model and sentiment analysis. The “temperature” parameter in the API request is set to 0 to minimize the degree of randomness in the model’s responses, ensuring more consistent and deterministic output.

2.4 Consolidation

After analyzing all the weeks, the script consolidates the data by combining the results for each unique week number. The consolidated data includes the total number of messages and the average scores for each category (Demand, Control, Support, Sentiment, Engagement) across all the chunks within a week.

2.5 Ethical Considerations

To ensure the anonymity and privacy of all participants, several measures were taken during the data collection and analysis process. Firstly, all personal identifying information, such as names and user IDs, was removed from the data, leaving only the message content for analysis. This anonymization process not

[3] https://github.com/perstarkse/ec-student-sentiment/blob/master/tools/interpreter-v2.py

only protects the participants' identities but also reduces the risk of bias in the analysis performed by the model, as it eliminates any potential influence from knowing the identity of the message authors.

Informed consent was not explicitly gathered from the participants, as the data was collected from a public chat group. However, considering the anonymization process and the absence of manual analysis, the ethical risks were deemed acceptable. The researchers carefully considered the balance between the potential benefits of the study and the protection of participants' privacy.

To further mitigate any potential ethical concerns, the collected data was securely stored and accessible only to the researchers directly involved in the study. The data will be retained only for the duration necessary to complete the research and will be securely deleted upon the conclusion of the study.

3 Results

The multiple regression analysis revealed a more nuanced understanding of the relationships between the variables. The model explained a significant proportion of the variance in sentiment, indicating a strong correlation between perceived control and sentiment. Interestingly, the relationship between demands and sentiment weakened when accounting for control and support, suggesting a mitigating effect of these factors. This finding aligns with the demand-control-support model, which posits that social support and control can buffer the negative effects of high demands.

3.1 Messages over Time

This table shows the results of the message volume analysis conducted on the chat messages. The message volume analysis was performed by counting the number of messages sent in each course.

As can be seen in the table, the average number of messages sent per week for all courses is 1725.38. This indicates that there is a high level of engagement in the chat groups. However, there is some variation in message volume across the different courses. For example, JavaScript Frontend has the highest average number of messages sent per week (3826.67), while Datalagring has the lowest (511.83).

Course	Total Messages	Max Messages per Week	Min Messages per Week	Average Messages per Week
html_css	11999	4374	568	1999.83
js_frontend	22960	5511	2205	3826.67
js_backend	8746	3591	144	1457.67
c#	7521	3051	70	1074.43
datalagring	3071	1356	136	511.83

Course	Total Messages	Max Messages per Week	Min Messages per Week	Average Messages per Week
asp.net	13063	3204	122	1451.44
Overall	67360	5511	70	1684.00

3.2 Sentiment over Time

This table below shows the results of the sentiment analysis conducted on the chat messages. The sentiment analysis was performed using the GPT-3.5-turbo model from OpenAI. The model was instructed to assign a sentiment score (0-100) to each message, based on the overall tone and sentiment expressed in the message.

The result showed some variety in sentiment over time based on the current course. Average sentiment for the courses ranged between 46.33 to 60.5 with some variety in standard deviation.

In fig. 1 you can see the changes over time for each course.

Course	Max Sentiment	Min Sentiment	Sentiment Standard Deviation	Average Sentiment
html_css	62	50	4.49	53.83
js_frontend	61	41	6.95	49.33
js_backend	60	46	6.01	51.17
c#	70	46	8.17	57.86
datalagring	65	45	7.50	55.33
asp.net	62	38	7.33	51.78
Overall	70	38	7.07	53.23

3.3 Correlational Analysis

Pearson's r correlation coefficients were calculated to examine the relationships between the Demand-Control-Support model variables (Control, Demand, Support) and sentiment scores. Pearson's r is a measure of the linear correlation between two variables, ranging from -1 to +1. A value of +1 indicates a perfect positive correlation, meaning that as one variable increases, the other variable also increases. Conversely, a value of -1 indicates a perfect negative correlation, where as one variable increases, the other variable decreases. A value of 0 indicates no linear correlation between the variables.

The results showed a strong positive correlation between Control and Sentiment ($r = 0.823586$, $p = 0.0$), suggesting that higher levels of perceived control were significantly associated with more positive sentiment in the chat messages.

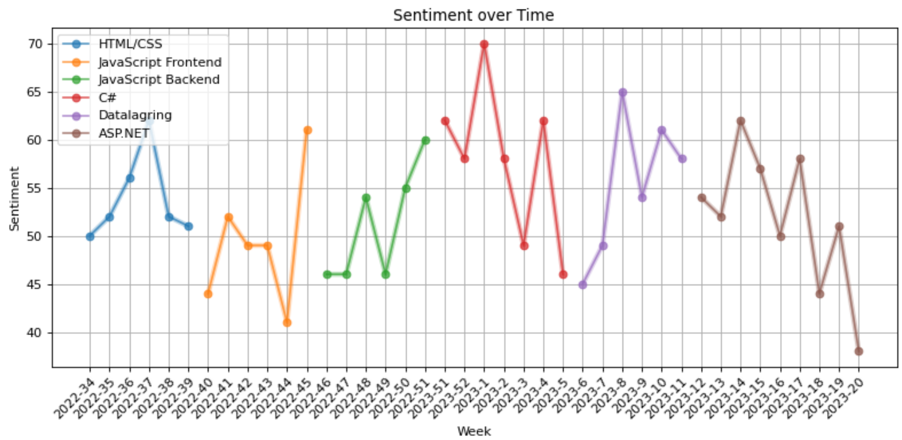

Figure 1: Displays the Sentiment level for each week. The line is split up in accordance to the different courses.

In contrast, there was a moderate negative correlation between Demand and Sentiment (r = -0.549477, p = 0.00024), indicating that higher levels of perceived demands were significantly associated with more negative sentiment.

The correlation between Support and Sentiment was strongly positive (r = 0.688576, p = 0.000001), suggesting that higher levels of perceived support were significantly associated with more positive sentiment.

3.4 Multiple Linear Regression Analysis

A multiple linear regression analysis was conducted to examine the relationship between sentiment and three independent variables: control, demand, and support. The goal was to determine the impact of these variables on sentiment and assess the overall fit of the regression model.

Prior to the analysis, the independent variables were centered by subtracting their respective means. Centering the variables helps mitigate multicollinearity, which occurs when independent variables are highly correlated with each other. This approach reduces the correlation between the independent variables and the constant term in the regression model, improving the stability and interpretability of the coefficients.

The regression results indicate that the model explains approximately 78.8% of the variation in sentiment (R-squared = 0.788). The adjusted R-squared value of 0.770 suggests that the model's explanatory power remains high even after accounting for the number of predictors. The F-statistic of 44.48 and the corresponding p-value of 3.40e-12 indicate that the overall model is statistically significant, meaning that the independent variables collectively have a significant impact on sentiment.

3.4.1 Coefficient Analysis

- Control (coefficient = 0.7533, p-value = 0.000): This variable has a positive and statistically significant relationship with sentiment. As control increases relative to its mean, sentiment tends to increase as well.
- Demand (coefficient = -0.1568, p-value = 0.140): This variable has a negative coefficient but is not statistically significant. This suggests that changes in demand relative to its mean do not have a significant impact on sentiment in this model.
- Support (coefficient = 0.3943, p-value = 0.000): This variable has a positive and statistically significant relationship with sentiment. As support increases relative to its mean, sentiment also tends to increase. The constant term (coefficient = 53.2250) represents the expected value of sentiment when all the centered independent variables are zero, i.e., when control, demand, and support are at their respective means.

3.4.2 Diagnostic Tests

- Normality of Residuals: The Omnibus test (p-value = 0.977) and the Jarque-Bera test (p-value = 0.893) both suggest that the residuals are normally distributed.
- Autocorrelation: The Durbin-Watson statistic (2.574) indicates no significant autocorrelation in the residuals.
- Multicollinearity: The condition number (8.28) suggests that multicollinearity is not a concern in this model.

The plot of the coefficients and the actual vs. predicted sentiment plot provide visual representation of the regression results. The actual vs. predicted sentiment plot demonstrates the model's ability to predict sentiment based on the centered independent variables, with points closer to the diagonal line indicating better predictions.

Overall, the model shows a strong fit, with control and support having significant positive effects on sentiment, while demand does not have a significant impact.

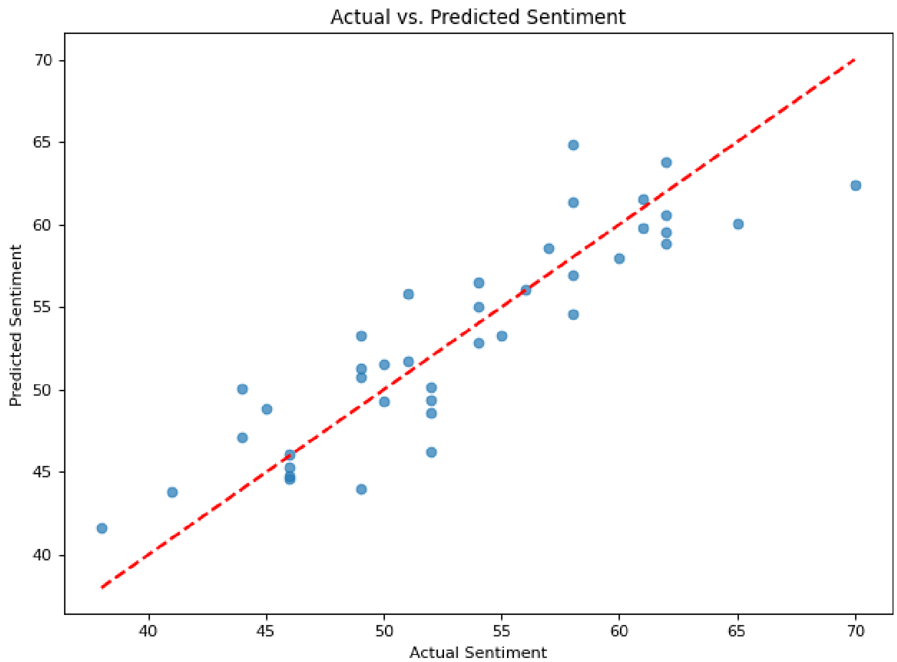

Figure 2: Multiple Linear Regression Analysis

3.5 Demand-Control-Support variables over time

As seen in fig. 3 there is significant change over time in the variables. There is a clear difference in the averages for the measures.

Variable	Max	Min	Standard Deviation	Average
Control	60	41	5.07	50.30
Demand	84	50	6.59	75.00
Support	78	45	6.82	63.00

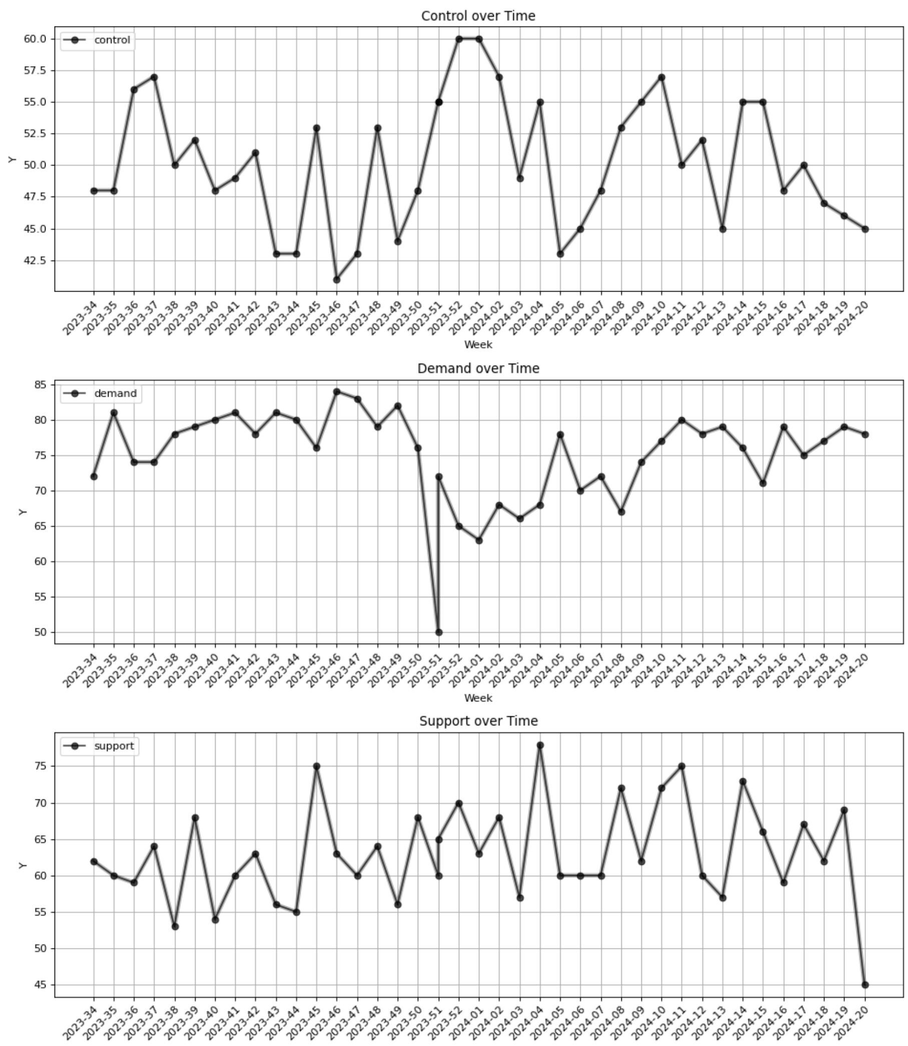

Figure 3: Displays the experienced Demand, Control and Support over time.

4 Discussion

This study explored the feasibility of using LLMs to analyze student chat messages, focusing on the relationship between demands, control, support, and sentiment within the framework of the demand-control-support model. Our findings, particularly the strong alignment with established psychological theory, offer compelling evidence for the validity of this approach.

4.1 Convergent Validity with the Demand-Control-Support Model

A key contribution of this study is the demonstration of convergent validity between LLM-derived insights and the well-established demand-control-support model. The multiple regression analysis revealed a strong positive correlation between perceived control and positive sentiment, aligning with the model's prediction that a sense of control can buffer against stress and contribute to well-being. This finding is further strengthened by the observation that the negative impact of demands on sentiment was significantly mitigated when accounting for control and support, highlighting the crucial role of these factors in promoting positive student experiences.

The consistency between our findings and the demand-control-support model provides strong preliminary evidence for the validity of using LLMs to extract psychologically meaningful insights from textual communication data. This suggests that LLMs can effectively capture and interpret the nuanced ways in which students express their perceptions of demands, control, support, and their emotional responses within a social learning environment.

4.2 Implications for Learning and Well-being

The alignment of our findings with the demand-control-support model, as revealed through LLM analysis, carries important implications for educational practices. The strong link between perceived control and positive sentiment underscores the need for educators to foster student autonomy and agency in learning, a concept central to Vygotsky's Zone of Proximal Development (ZPD) (Vygotsky and Cole, 1978). Providing choices in assignments, encouraging student-led projects, and creating opportunities for self-directed learning can empower students to operate within their ZPD, fostering both skill development and positive emotional experiences.

Furthermore, recognizing the mitigating effect of control and support on the negative impact of demands highlights the importance of providing adequate resources and scaffolding for challenging tasks. This aligns with research demonstrating the buffering effect of social support on stress (Karasek and Theorell, 1990), where a supportive learning environment – characterized by clear guidelines, timely feedback, readily available assistance, and peer collaboration – can help students navigate demanding situations while maintaining positive well-being. While a certain level of challenge, often referred to as “positive stress” or “eustress” (Sarada and Ramkumar, 2015), is essential for growth, LLM-driven insights can help educators strike a balance, ensuring that demands are appropriately calibrated to individual student needs and support systems.

By providing this nuanced understanding of the interplay between demands, control, support, and sentiment, LLMs can empower educators to create learning environments that are both challenging and supportive, fostering both academic growth and student well-being.

4.3 Limitations and Future Directions

While this study provides valuable insights into the feasibility of using LLMs for psychological analysis of chat messages, it is important to acknowledge its limitations. The study relied on a single LLM (GPT-4o) and a specific set of prompts for the analysis. Future research could explore the use of different LLMs and prompts to assess the robustness and generalizability of the findings.

Additionally, the study focused on a specific educational context (students in a Discord chat group) and a particular time frame (fall 2022 to summer 2023). Further research could investigate the applicability of this approach to other educational settings, such as traditional classrooms or online learning platforms, and examine longitudinal data over longer periods to gain a more comprehensive understanding of the dynamics of stress, learning, and well-being.

Another limitation of the study is the reliance on automated analysis without direct validation against established psychological measures or expert human judgment. Future studies could incorporate traditional psychological assessments and compare the results obtained from LLMs with those from validated measures to establish the convergent validity of the approach.

4.4 Conclusion

This exploratory study demonstrates the exciting potential of Large Language Models (LLMs) as powerful tools for investigating psychological phenomena within educational settings. By analyzing student chat messages, we uncovered nuanced relationships between perceived demands, control, support, and sentiment, aligning with established psychological theories like the demand-control-support model. This convergence of findings with existing theory provides compelling evidence that LLMs can effectively capture and interpret psychologically meaningful patterns within textual data.

The alignment of our results with established theory is promising. The strong correlation between perceived control and positive sentiment, along with the mitigating effect of control and support on the impact of demands, underscores the importance of fostering student agency and providing a supportive learning environment – key factors known to promote student well-being.

While further validation through comparisons with traditional psychological assessments and expert human judgment is crucial, this study provides a strong foundation for future research in this burgeoning field. Exploring the capabilities of diverse LLMs, expanding to various educational contexts, and incorporating longitudinal data analysis will further illuminate the dynamic interplay between stress, learning, and well-being.

This study marks an important step towards leveraging the power of LLMs to gain deeper insights into the psychological experiences of students. As research in this area progresses, LLMs hold immense promise for informing the development of targeted interventions and fostering supportive learning environments that empower students to thrive.

References

Chim, J. *et al.* (2024) 'Overview of the CLPsych 2024 shared task: Leveraging large language models to identify evidence of suicidality risk in online posts'. Edited by A. Yates et al., pp. 177–190. Available at: https://aclanthology.org/2024.clpsych-1.15.

Devlin, J. *et al.* (2019) 'BERT: Pre-training of deep bidirectional transformers for language understanding'. Available at: https://arxiv.org/abs/1810.04805.

Karasek, R.A. and Theorell, T. (1990) *Healthy work*. Basic Books. Available at: https://books.google.se/books?id=MrXuAAAAMAAJ.

Kristensen, T. (1995) 'The demand-control-support model: Methodological challenges for future research', *Stress Medicine*, 11, pp. 17–26. Available at: https://doi.org/10.1002/smi.2460110104.

Nguyen, T. *et al.* (2022) 'Improving the generalizability of depression detection by leveraging clinical questionnaires'. Edited by S. Muresan, P. Nakov, and A. Villavicencio, pp. 8446–8459. Available at: https://doi.org/10.18653/v1/2022.acl-long.578.

Radford, A. *et al.* (2019) 'Language models are unsupervised multitask learners', *OpenAI blog*, 1(8), p. 9.

Sarada, P.A. and Ramkumar, B. (2015) 'Positive stress and its impact on performance', *Research Journal of Pharmaceutical, Biological and Chemical Sciences*, 6, pp. 1519–1522.

Schwabe, L. and Wolf, O.T. (2010) 'Learning under stress impairs memory formation', *Neurobiology of Learning and Memory*, 93(2), pp. 183–188. Available at: https://doi.org/https://doi.org/10.1016/j.nlm.2009.09.009.

Tausczik, Y.R. and Pennebaker, J.W. (2010) 'The psychological meaning of words: LIWC and computerized text analysis methods', *Journal of Language and Social Psychology*, 29(1), pp. 24–54. Available at: https://doi.org/10.1177/0261927X09351676.

Touvron, H. *et al.* (2023) 'Llama 2: Open foundation and fine-tuned chat models'. Available at: https://arxiv.org/abs/2307.09288.

Vaswani, A. *et al.* (2017) 'Attention is all you need'. Edited by I. Guyon et al., 30. Available at: https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf.

Vygotsky, L.S. and Cole, M. (1978) *Mind in society: Development of higher* *psychological processes*. Harvard University Press. Available at: https://books.google.se/books?id=RxjjUefze_oC.