Growth Mind-set through Resilient Intelligent Technologies

Work package 5

Handbook

D.1; D.2; D.3; D.4; D.5; D.6

Project Number 2022-1-IT02-KA220-HED-000085944

For information and contacts
University of Camerino
andrea.marconi@unicam.it

Table of contents

Preamble

As part of the ongoing European project GRIT – Growth Mind-set through Resilient Intelligent Technologies, a pilot session was conducted to test a newly developed interactive game designed to explore and promote non-cognitive skills, such as perseverance, resilience, and tolerance, within academic and research environments. The pilot served as an important moment of validation for both the game mechanics and the broader goals of the project. The initiative was carried out across two distinct educational settings: undergraduate and master's programs in Computer Science and Applied Mathematics, as well as within a Doctoral School involving early-career researchers.

The main purpose of this pilot was to introduce the game as a reflective tool, encouraging participants to actively engage with the concept of non-cognitive skills in a way that was both personal and thought-provoking. Rather than simply delivering theoretical content, the game immerses users in simulated scenarios that invite emotional and strategic responses—prompting players to reflect on how they react in challenging or uncertain situations.
Elements of gamification and narrative design created the GRIT interactive experience that becomes more than a tool or a framework for sentiment analyses: it was a space for critical self-exploration. At the end of the session, players are asked to complete an open-ended questionnaire to share their impressions and connect the game's content with their own educational or career journeys.

The game was tested in live classroom settings and at distance, with the objective of creating a dynamic and interactive virtual environment where students could engage enthusiastically with the platform, often sparking spontaneous answer to the related questionnaire about their users experience and the correlation with their academic careers and non-cognitive skills. The feedback collected through the integrated questionnaire was being analysed by the interdisciplinary research team. Initial impressions suggest that participants found the experience meaningful, particularly in how it helped them recognize the often-overlooked role of soft skills in their academic development. The pilot also reinforced the value of using interactive formats to approach themes like emotional regulation, failure, and perseverance topics that are deeply relevant but rarely addressed in traditional academic settings.

Following the success of this initial trial, the GRIT team focused on refining the game based on participant feedback. The next phase was exploring the integration of sentiment analysis technologies to further enhance the evaluation of non-cognitive skills through digital interaction. Based on the sentiment analyses of the questionaries of the users' experiences, they find relevant evidence addressing:

- Defining similarity in achieving academic objectives and in lifestyle
- Clustering correlation among data to identify strong patterns for academic success
- Identify a set of patterns specifically related to civic commitment and social awareness
- Applying Python Function on data frame as computing technique
- Testing unsupervised and supervised machine learning models
- Finalizing the model

Now the researchers are moving on publication and scientific articles.

1. Foundational frameworks

1.1 General description of the task and context of the research

This project aims to explore the relationship between GRIT (passion and perseverance towards long-term goals), emotional experience and narrative responses of students after a gamified activity. To do so, a four-phase protocol has been designed, applied to students from different countries and areas of study.

In the first phase, participants are administered the GRIT questionnaire, developed by Angela Duckworth. This tool assesses two key dimensions: sustained passion for long-term goals and perseverance in effort, even in the face of difficulties or failures.

The questionnaire is administered at the beginning and its results serve as a starting point for further analysis. Students from different countries and academic disciplines participate, which allows for an intercultural and interdisciplinary approach to the study.

Participants then access the interactive game developed specifically for this study[1]. This game has been designed as a playful simulation of decision-making and perseverance, which puts into practice skills linked to GRIT. During the game, the aim is to generate a meaningful and reflective emotional experience that connects with the dimensions assessed in the previous phase.

At the end of the game, participants complete an open-ended questionnaire, in which they are invited to reflect freely on:

- Their decisions during the game.
- Their predominant emotions.
- Perceived difficulty.
- The sense of achievement or frustration.
- Their perception of the connection between the game and their real life.

This qualitative questionnaire allows for the collection of emotional narratives and in-depth personal assessments.

In the final phase, a sentiment analysis of the open-ended responses obtained is carried out. This computational technique allows us to identify the emotional valence (positive, negative, neutral) and, in some cases, the predominant emotion (such as joy, frustration, hope, etc.). These emotional scores are then linked to the GRIT questionnaire scores, with the aim of exploring whether students with higher GRIT experience different or more regulated emotions during the game.

This mixed approach, combining quantitative instruments (GRIT questionnaire), an interactive experience designed ad hoc and a qualitative-computational analysis, allows for a richer and more contextualised understanding of GRIT and its emotional dimension.

1.2 Architecture of the open-ended questionnaire

Let's start by defining what an open question is. An open-ended question is one that invites someone to respond with more than one word. For example, instead of asking, "Are you feeling cheerful?" (which can be answered with a simple "yes" or "no"), we can ask, "How do you feel?" This second

[1] https://www.gritproject.eu/the-game/

approach encourages reflection and can reveal key information about the topic of study that a yes/no answer might not offer.
In surveys, open-ended questions allow respondents to express themselves in their own words, without being limited to pre-defined options. Unlike closed-ended questions, open-ended questions encourage respondents to provide detailed feedback. This type of response is valuable when trying to analyse preferences, behaviours or experiences at a deeper level.
An open-ended question goes beyond a simple "yes" or "no" answer. It invites people to share more about their thoughts, feelings or perspectives. These questions help us understand not only what the other person thinks, but why they think the way they do. By encouraging deeper responses, we can uncover motivations, preferences or a certain purpose that may go unnoticed when using a closed question.
This deeper level of understanding allows us to conduct a more thorough analysis of the topic of study. These questions are fundamental to gathering richer and more meaningful information, making them a valuable tool in research.
One of the main advantages of using open-ended questions in our project is that they create a space for participants to express genuine emotions and opinions, which is really valuable for our work. A question such as "How did you feel when you failed to open the locker on the first try?" allows respondents to share what they really felt, offering a more authentic view of their experience.

1.3 Data collection process and participant profiles

The success of the GRIT pilot hinged on a carefully structured and purpose-driven data collection strategy. From the outset, the research team adopted a collaborative, iterative approach to define clear goals that would align the design of the interactive game with the overarching aim of evaluating non-cognitive skills such as grit, perseverance, resilience, and tolerance within academic and research settings.
Consequently, the GRIT project began by addressing a fundamental question: How can we design a digital game that both engages academic audiences and elicits meaningful reflection on non-cognitive competencies? With this in mind, stakeholders from educational, psychological, and technological domains worked together to identify learning objectives and research priorities. The process involved structured co-design workshops, consultations with educators, and iterative feedback loops that helped crystallize the core goal: to gather data that could illuminate the emotional, motivational, and attitudinal dimensions of academic life.
The project intended outcome was twofold:

1. To collect qualitative and emotional reflections from diverse user groups (students and early-career researchers) through their interaction with the GRIT game.
2. To analyse these reflections using sentiment analysis and machine learning to inform the development of a predictive model for non-cognitive skills in academic success.

The sampling was guided by the need for diversity and representation across academic disciplines and career stages. Participants were recruited through a combination of institutional networks, course integrations, and targeted communications, including:

- Undergraduate and Master students from different Universities and courses i.e. Computer Science and Applied Mathematics, Education Science, Pedagogy, Psychology etc..

- PhD candidates from the International Doctoral School (via workshops and official calls for participation).
- Young researchers (postdoctoral researchers, tenure-track researchers, lecturers).

This multi-tiered approach ensured broad demographic coverage while also minimizing sample bias. Particular attention was paid to gender balance, international representation, and accessibility (e.g., providing both online and offline modes of participation).
The GRIT study relied predominantly on primary data, user-generated responses to an open-ended questionnaire administered at the end of the game. This questionnaire was embedded directly into the game platform for a seamless transition from gameplay to reflection. Additionally, a standalone Google Form was made available to accommodate different user preferences and technical setups. The target group counted more than 380 players and 179 open questionnaires analysed.
The questionnaire focused on:

- Emotional responses to in-game scenarios.
- Self-assessed alignment with portrayed non-cognitive challenges.
- Reflections on academic and personal development.

To complement primary data, secondary data (e.g., metadata on game interaction times, sequence of choices, and response latencies) was collected to support behavioural analysis and enrich sentiment evaluation. In some cases, synthetic data was generated to model hypothetical player responses and train early-stage machine learning models.
The chosen method, a reflective, open-ended questionnaire, allowed for deep qualitative insights. This method was well-suited to GRIT's exploratory goals, enabling participants to articulate complex emotional states and behavioural tendencies beyond what could be captured through structured multiple-choice questions.
To ensure reliability and maximize response quality, the instruments were:

- Pilot-tested across small groups for clarity and emotional accessibility.
- Adapted linguistically to match the language of instruction or origin of the participants.
- Designed for completion within 10–15 minutes, respecting the time constraints of busy academic professionals.

Despite these efforts, the research team encountered a notable challenge: a high drop-off rate in questionnaire completion. While players were highly engaged with the game itself, many opted not to complete the final reflective stage. This highlighted a common tension in gamified research tools, where interaction is high, but follow-through on analytical components requires additional motivation or scaffolding.
Moreover, the ethical considerations were central throughout. All participants provided informed consent and were briefed on data privacy, anonymity, and the academic purpose of the study. The research design received clearance from the research project teams ensuring alignment with GDPR and institutional review standards.
After collection, the data underwent rigorous cleaning and preprocessing:

- Removal of duplicates and empty responses.
- Standardization of text entries for sentiment analysis.
- Structuring of narrative responses for machine learning input.

This phase ensured that the final dataset was not only reliable but also optimized for downstream tasks such as clustering, emotion tagging, and predictive modelling.

1.4 Theoretical framework of Grit on the five-factor model and correlation

We will begin this section by attempting to provide an answer to the term "Personality". Defining this construct is somewhat complex, so we will begin by describing the origin of the word. "Personality" derives from the Latin persona and refers to the masks worn by actors in ancient times. From this point of view, we understand that our personality would be the part that is visible and exposed to others. However, when we turn to any dictionary, we find that this concept is defined as "Individual difference that constitutes each person and distinguishes them from others". This definition makes it clear that, although we may have similarities with other people, we all have individual attributes that differentiate us from others. Based on this conception, we can define personality as a set of qualities that makes each of us different from others and the same as ourselves over time. From this point of view, we are assuming a sense of continuity or consistency of the person over time and in different situations that will allow us to predict and understand their behaviour. We now understand personality to be a deeply embedded pattern of overt cognitive, affective and behavioural traits that persist over long periods of time. This definition places the focus on two fundamental concepts: the existence of individual differences and intrapersonal functioning. We assume that each subject is unique and different from others. Intrapersonal functioning refers to a person's internal processes that guide their behaviour. It is these internal processes that provide a sense of continuity.

At this point, the question arises as to whether personality is biologically determined or whether it is the product of our experiences and interactions with the environment. Numerous studies have shown the influence of biological aspects on our personality (Caspi, 2000; Plomin & Caspi, 1999; Rowe, 1999). Temperament is the part of our personality that is biological in origin. This component manifests itself from early childhood (Strelau, 1998). However, the environment also plays a crucial role in the development of our personality. Family, culture and social class are some of the environmental aspects that have the greatest impact on the shaping of our personality. This component linked to the subject's experiences throughout his or her life cycle is what we call character.

Gordon Allport (1897-1967) is one of the most relevant psychologists in the study of personality. He defined personality as "the internal dynamic organisation of the individual's psychophysical systems which determine his characteristic behaviour and thinking" (Allport, 1961, p. 28). He understood personality as constantly changing and unique.

For Allport, we are the result of heredity and environment. This author considers that our personality can be described on the basis of the presence of a series of traits. For Allport, a trait is a dynamic behavioural tendency that is the product of the integration of numerous habits and expresses a characteristic way in which an individual reacts to his or her environment.

Subsequently, several authors such as Norman (1975), Goldberg (1993) and McCrae-Costa (1985) postulated that an individual's personality was made up of 5 major factors. The latter succeeded in integrating the previous work into a single theoretical framework in their Big Five Factor Model: Neuroticism, Extroversion, Openness, Agreeableness and Conscientiousness. According to this model, the neurotic person would be characterised as insecure and nervous while the extrovert person would be sociable, fun-loving and affectionate. Another major factor that makes up this theory would be Openness, which would be present in the independent, creative and daring person.

Affability can be found in kind and trusting individuals. Finally, Scrupulousness would be found in individuals who are careful, hardworking and organised.
At present, one of the definitions with the greatest consensus and which largely brings together the above is that of Larsen and Buss (2005), who consider that personality is "the set of psychological traits and mechanisms internal to the individual, organised and relatively durable, which influence the interactions and adaptation of the person with his or her intrapsychic, physical and social environment".
That is to say, although part of our personality is biologically determined, we have a wide margin that is subject to our motivations, aspirations and dreams. That is why the way we are is not determined from the start, and we can therefore have a great deal of influence over how we want to be.
Along the same lines is Anglea Duckworth's theory, described in her book "Grit: The Power of Passion and Perseverance" (2016). Angela's studies focus on the analysis of a personality trait that she calls "Grit". For this author, Grit is defined as the combination of:

1. **Passion**: deep and sustained commitment to long-term goals.
2. **Perseverance**: ability to maintain effort in the face of difficulty, boredom or failure.

Duckworth argues that success depends not so much on innate talent, but rather on the ability to sustain effort and interest over long periods, even when progress is slow or uncertain. His work is characterised by the presence of:

1. **Hierarchical goals**: People with high GRIT organise their goals in hierarchies. Higher goals (more abstract and long-term) guide choice and persistence in subordinate goals (more specific and short-term).
2. **Sustained intrinsic motivation**: Unlike momentary motivations, Grit implies a stable interest in a topic or purpose over time.
3. **Resilience in the face of failure**: Grit allows you to keep trying despite obstacles or failures, seeing these as part of the process towards ultimate achievement.
4. **Growth mindse**t: Although not the same, Duckworth builds on Carol Dweck's idea that skills and intelligence can be developed with effort, learning and perseverance.

Another crucial aspect of the issue at hand is the link between personality and emotion. Generally speaking, personality influences how we experience, express and regulate emotions. One of the theoretical models of personality where emotions have been linked is the Big 5 model of personality. This model identifies five broad personality traits that are consistently associated with different emotional profiles:

Feature	**Main emotional link**
Neuroticism	Tendency to experience negative emotions such as anxiety, sadness, anger. It is the strongest predictor of chronic emotional distress.
Extraversion	Increased frequency and intensity of positive emotions (joy, enthusiasm); increased sociability and positive reactivity.

Openness to experience	Aesthetic emotions, emotional curiosity, greater emotional complexity.
Conscientiousness	Better emotional regulation, greater emotional self-control, less impulsivity.
Kindness	Prosocial emotions such as empathy, compassion and less tendency to emotional conflict.

The relationship between personality and emotions is bidirectional and dynamic. On the one hand, personality predisposes to certain emotional patterns. And, in turn, repeated emotional experience can consolidate or shape personality traits (especially during childhood and adolescence).
At this point, one of the objectives of our project is to test whether people who are characterised by high levels of Grit are associated with a particular emotional profile. Another objective is to test whether the sentiment analysis technique is effective in analysing the feelings of these participants. To do this, we generated a game with a high degree of difficulty in which the person had to use his or her determination to achieve a goal. Afterwards, they had to answer open-ended questions. Finally, these answers were analysed using the Sentiment Analysis technique.

2. Data analysis and AI techniques

This section presents the full analytical pipeline developed to process and transform the qualitative textual data collected during the GRIT pilot sessions. The University of Camerino and University of Maribor (UM) led this work package component, focusing on data preparation, multi-layered sentiment extraction, thematic analysis relevant to civic and ethical reasoning, and construction of structured emotional features for predictive modelling.

The following research questions were addressed in this Work package n.5 :

RQ1: Can grit be accurately predicted using sentiment analysis of user-generated texts instead of traditional psychometric surveys?

RQ2: What emotions identified through sentiment analysis most significantly influence grit predictions in students?

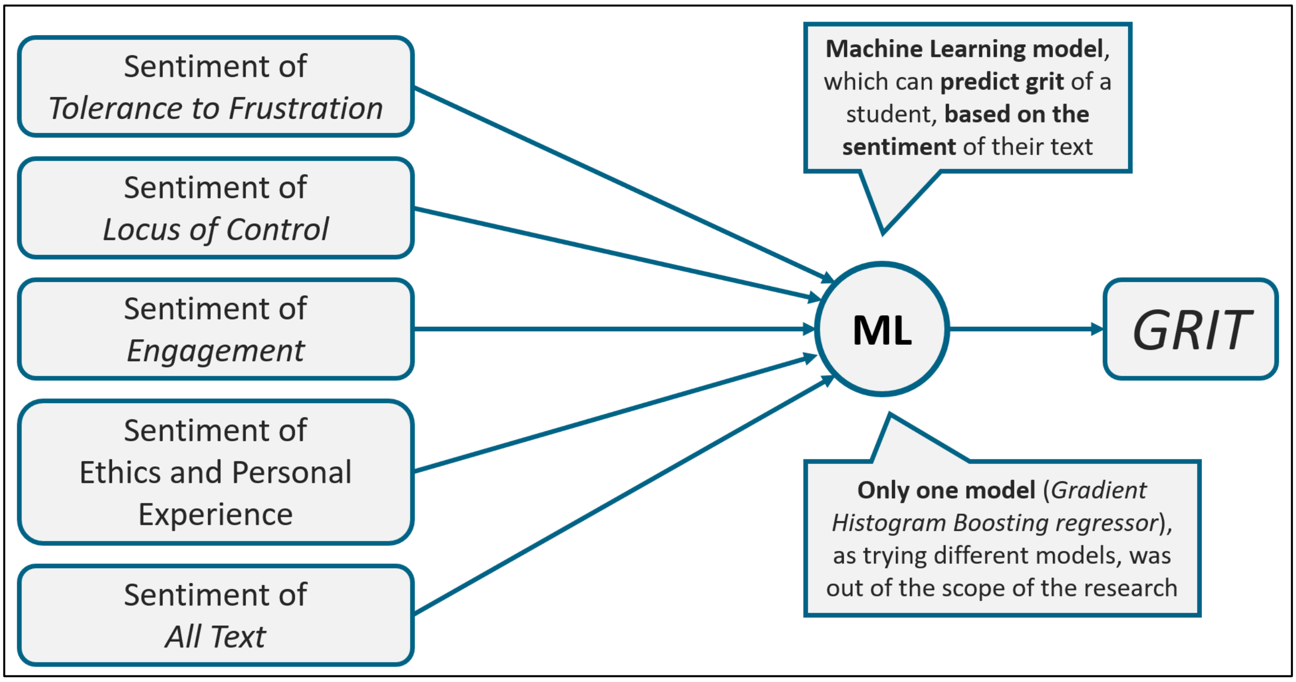
RQ3: How can explainable XAI contribute to interpreting the relationship between emotional sentiments and grit?

The activities investigates these questions through sentiment and emotional analyses of multilingual open-ended responses from students who participated in a serious game designed to provoke challenging, perseverance-testing scenarios. The use of sentiment analysis combined with XAI interpretation not only aims to improve the measurement accuracy of grit but also offers deeper insights into the emotional mechanisms underlying perseverance and passion.

2.1 Qualitative coding and preprocessing pipeline

The data collected through the GRIT game and subsequent questionnaire involved both structured and unstructured inputs. Open-ended questions focused on four psychologically and socially meaningful domains: Tolerance to Frustration, Locus of Control, Engagement, and Ethics and Personal Experience. These responses were provided in natural language and in multiple languages depending on the participant's country, adding complexity to preprocessing. To address this, the UM team developed a set of preprocessing steps including text normalization (lowercasing, punctuation handling), handling of missing values, language detection where appropriate, and formatting clean-up for numerical compatibility.

Each participant's responses from these four domains were preserved both in their original form and as a concatenated feature, referred to as "*All Text*." This combined field served to centralize participants' expressive data and provided a more contextually rich feature for sentiment extraction and modeling. Additionally, each participant responded to a standard 8-item grit questionnaire scored on a Likert scale from 1 to 5. The average of these responses was used to calculate a continuous grit score, which became the primary target variable for predictive modeling.



Demographic information (gender, age, country, academic field, and study level) was also retained to allow subgroup analysis and to enrich interpretive insights from the model. These metadata fields supported cohort comparisons, such as gender-based SHAP analyses discussed in later stages.

2.2 Sentiment Analysis Approach

To convert qualitative reflections into machine-readable features, we implemented a multi-model sentiment extraction pipeline. This pipeline was designed to be modular and comparative in nature, allowing multiple methods to be applied across the same textual input for robustness.

Lexicon-based sentiment scoring was carried out using VADER (Valence Aware Dictionary for Sentiment Reasoning), which generated compound polarity scores and individual probabilities for negative, neutral, and positive tones. VADER was selected for its ability to analyze short and informal texts, such as those commonly submitted in the questionnaire.

Transformer-based models were used to derive categorical emotion scores. These included Hugging Face pipelines that assigned probabilistic scores to basic emotion categories such as joy, sadness, anger, trust, anticipation, fear, surprise, and disgust. In parallel, SpaCy-based models were used to extract psychological VAD dimensions: valence (positivity), arousal (intensity), and dominance (control).

We further integrated a generative approach using the OpenAI ChatGPT API. With structured prompting, the same responses were processed to return either scalar sentiment scores (ranging from −1 to 1), ternary sentiment probabilities (positive, neutral, negative), or full emotion profiles. Additionally, we prompted ChatGPT to return VAD values, providing a generative comparison to SpaCy estimates.

Each set of features was appended to the main data frame with standardized prefixes indicating the method of extraction and the thematic source of the text. This naming schema enabled precise control over feature selection in downstream modelling.

2.3 Clustering and Pattern Discovery

Predicting grit directly from sentiment-derived features alone can be challenging due to its subjective nature. To address this complexity, an alternative approach involves clustering participants into groups based on their responses to structured questionnaire items combined with sentiment analysis features. Clustering enables a deeper understanding of **personality traits** by grouping participants with similar psychological profiles, thereby creating distinctions regarding their grit levels.

We applied the K-Means clustering algorithm (using Euclidean distance), aiming to minimize intra-cluster variation and ensuring that participants within each cluster exhibit similar attitudes, emotional responses, and behavioural patterns. The clustering analysis utilized both structured questionnaire responses and sentiment analysis metrics derived from open-ended textual responses.

To effectively visualize and interpret cluster results, the data were projected onto the first two principal components (PCs) via Principal Component Analysis (PCA). Through correlation analysis, patterns within cluster structures can emerge, providing deeper insights into the distinct groups.

Subsequently, the derived clusters are utilized for training group-specific machine learning regressors to estimate grit scores. The outcomes from these regressors are then compared against the same model trained on the overall dataset, enabling an examination of the relevant features and potentially enhancing prediction accuracy.

2.4 Python-Based Analytical Techniques

All processing and modelling scripts were implemented in Python using standard scientific libraries. The backbone of the feature extraction routine was a modular batch-processing function. This function applied each sentiment extraction method across the relevant columns and merged the results into the main dataset using a systematic prefixing strategy.

This approach made it possible to annotate each participant's response with more than 300 numerical features, capturing both emotional polarity and categorical emotion likelihoods across different extraction models. Dimensionality was reduced in later stages through filtering based on feature variance, correlation analysis, and interpretability.

3. Machine learning model development

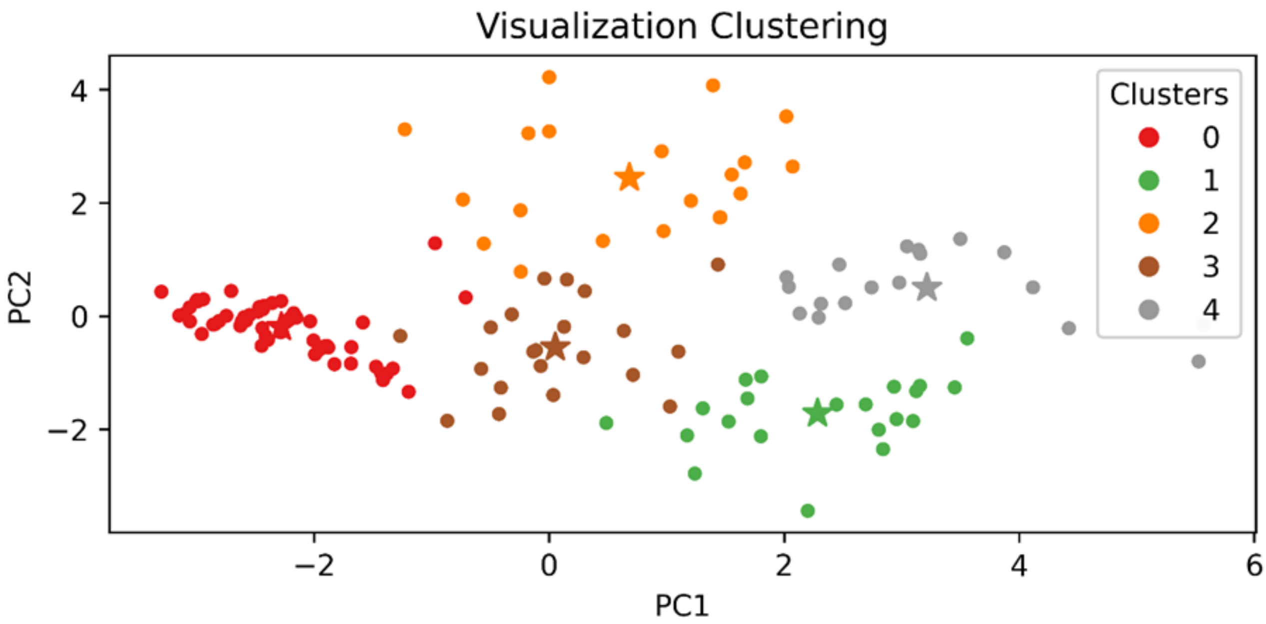
The objective of modeling psychological grit based on the structured sentiment features extracted from participants' free-text responses. This section details the modeling rationale, experimental setup, interpretability techniques, and final outcomes of the grit prediction model.

3.1 Modeling Objectives and Feature Design

Our main goal was to determine whether a participant's expressed emotions, attitudes, and affective tone in written responses could reliably predict their overall grit score, calculated from the structured questionnaire. We approached this as a regression problem, given the continuous nature of the grit outcome variable. The feature set consisted of over 300 variables derived from sentiment analysis models. These included VADER sentiment polarity scores, categorical emotion scores from transformer-based models, VAD (valence-arousal-dominance) dimensions from SpaCy and ChatGPT, and emotion probabilities such as sadness, joy, anticipation, and trust.
Initial experiments showed that using features extracted from the "All Text" composite (combining all open-text fields) yielded similar or even better results than using each thematic section separately. Therefore, the All Text-based sentiment features were prioritized as the main inputs into the modeling pipeline. Notably, we found that VADER sentiment scores, combined with specific emotion categories such as anticipation and trust, produced the most stable and interpretable results. Other methods, such as ChatGPT emotion probabilities and SpaCy VAD features, provided useful nuance, but tended to contribute less predictively on their own.

3.2 Unsupervised Learning Approaches

The unsupervised learning methods employed include K-Means clustering and PCA, both implemented via the scikit-learn library. Given that grit was initially measured on a Likert scale, we selected K = 5 clusters corresponding conceptually to discrete grit categories: "Very High," "High," "Moderate," "Low," and "Very Low". The resulting visualization, presented below, highlights clearly distinguishable profiles with each star indicating a cluster centroid. We also concentrate our attention on the first two (most important) components of the PCA.

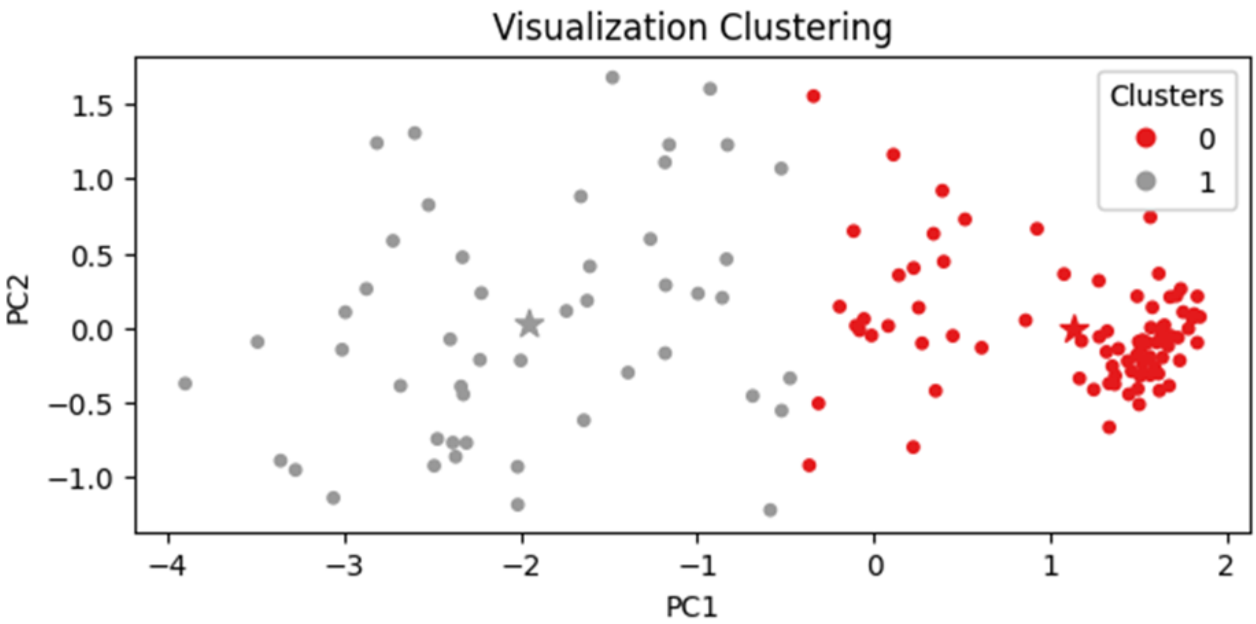


To interpret these clusters more effectively, we computed the mean grit score for each identified group:

Cluster	Mean Grit
C0	2.623
C1	4.179
C2	2.900
C3	3.333
C4	3.972

From these findings, clusters exhibiting higher values on PC1 demonstrate, on average, higher grit scores. Further insights were obtained through correlation analysis between PCs and grit measurements, revealing that PC1 has a strong positive correlation with grit (0.906) while PC2 has a weak negative correlation (-0.40). Consequently, PC1 emerges as a more robust indicator of grit compared to simple grit averages, as it consolidates multiple questionnaire responses into a cohesive psychological dimension. Thus, PC1 offers a valuable tool for accurate classification and interpretation of grit-related clusters.

Additionally, clusters were computed solely using sentiment analysis-derived features and the K-Means algorithm with K = 2 was applied (see figure below), categorizing data into "High grit" and "Low grit" groups. This choice ensured sufficient data points per cluster for effective training of machine learning regressors. Using the established supervised learning pipeline, grit scores for each cluster were subsequently predicted.



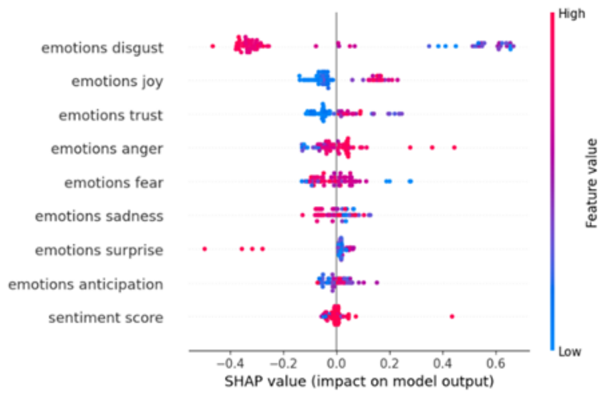
However, the obtained results did not show any improvement in the evaluation metrics for either cluster compared to a model trained on the entire dataset:

Cluster	MSE	MAPE
C0	0.3629	0.1623
C1	0.4663	0.1617

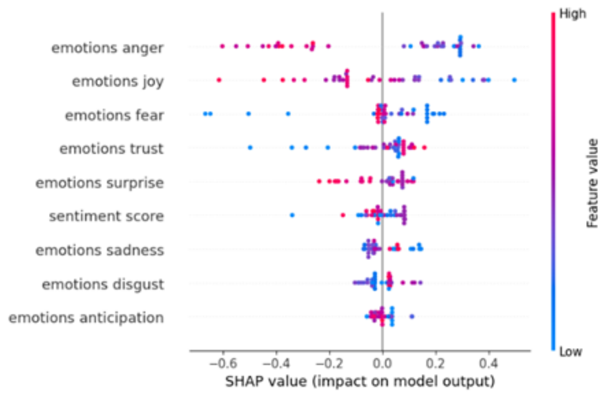
Moreover, predictions for cluster C1 resulted in constant values, limiting the interpretability of feature importance through SHAP[2]. For this reason, the general approach outlined in the subsequent sections remains preferable.

The SHAP dot plots for the two clusters are shown below to highlight the limited interpretability of the results. These plots visualize the impact of each feature on the model's output for the instances within each cluster. Each dot represents a single prediction, with its position on the x-axis indicating the SHAP value (i.e., the feature's impact on the model's output), and its color representing the original feature value. On the y-axis, the features are ordered from the most (the top one) to the least influential (the bottom one).

[2] SHAP, SHapley Additive exPlanations is a method for explaining the predictions of machine learning models, based on game theory. Essentially, SHAP assigns a value to each feature of an input, indicating its contribution to the model's prediction for that specific instance.



SHAP dot plot for C1



SHAP dot plot for C2

For most features, the SHAP values are tightly distributed around zero, indicating minimal impact on the model's output. This suggests that, across both clusters, the majority of features do not meaningfully contribute to the prediction task, resulting in weak explanatory power and limited insights into the model's decision logic.

Due to the overall low contribution of most features, only a small subgroup of variables shows a clear impact. The feature 'joy' emerges as the most impactful in common for both clusters. In cluster C1, higher values of the 'joy' feature are associated with positive SHAP values, whereas in cluster C2 they correspond to negative SHAP values. This contrast suggests that 'joy' plays a discriminative role in differentiating the two clusters. Additionally, 'disgust' and 'anger' are the dominants in C1 and C2, respectively. Nevertheless, their lack of relevance in the opposite cluster, unlike 'joy', prevents them from deriving more detailed or comparative insights.

3.3 Supervised Learning Approaches

For supervised prediction, we used a histogram-based gradient boosting regressor from the scikit-learn library. This choice was motivated by its robustness to overfitting, native handling of missing values, support for non-linear relationships, and strong performance in medium-sized datasets without extensive hyperparameter tuning. The model was trained to predict the continuous grit score based on sentiment-derived features.

Multiple versions of the model were trained using different combinations of input features to evaluate the comparative predictive value of each sentiment extraction method. The best performance was achieved using all extracted features from the All Text field. The model achieved a mean squared error (MSE) of 0.221 and a mean absolute percentage error (MAPE) of 0.105, which we interpret as strong predictive accuracy for this type of subjective psychological construct.

3.4 Model Validation and Performance Metrics

To evaluate model performance rigorously, we adopted a standard train-test split (80/20) using a fixed random seed to ensure reproducibility. Performance was evaluated on the hold-out test set using both MSE and MAPE. The low values of these metrics across multiple runs indicate the stability of the model and its capacity to generalize beyond the training data.

In addition to numerical evaluation, we performed feature-level interpretation using SHAP (SHapley Additive exPlanations). This method assigns an impact score to each input feature, revealing how much it contributed to the prediction. SHAP provides a breakdown of the contribution of each feature toward pushing the prediction higher or lower than the average predicted grit. This helps us identify which aspects of emotional expression drive the model's output.

The SHAP summary plot visually presents this breakdown, where each dot represents a single participant, and the color of the dot corresponds to the raw value of the feature (e.g., red for high and blue for low). Features such as All Text_vader_sentiment_score and All Text_chatgpt_emotions_trust were particularly influential. For instance, when the VADER sentiment score was high (shown in red), the SHAP values were concentrated to the right, indicating that this pushed the grit prediction upward. Conversely, high sadness or disgust values generally pulled the prediction downward.

To complement SHAP, we used partial dependence plots (PDPs), which show how the model's prediction for grit changes as we vary one feature while holding others constant. These plots help us understand whether the relationship between a sentiment feature and grit is linear, monotonic, or exhibits threshold behavior.

The PDP for All Text VADER sentiment, for instance, shows a strong positive relationship: as sentiment becomes more positive, predicted grit steadily increases. In contrast, PDPs for emotions like dominance or anticipation display threshold effects—little change occurs until the emotion score exceeds a certain value, after which grit prediction sharply rises.

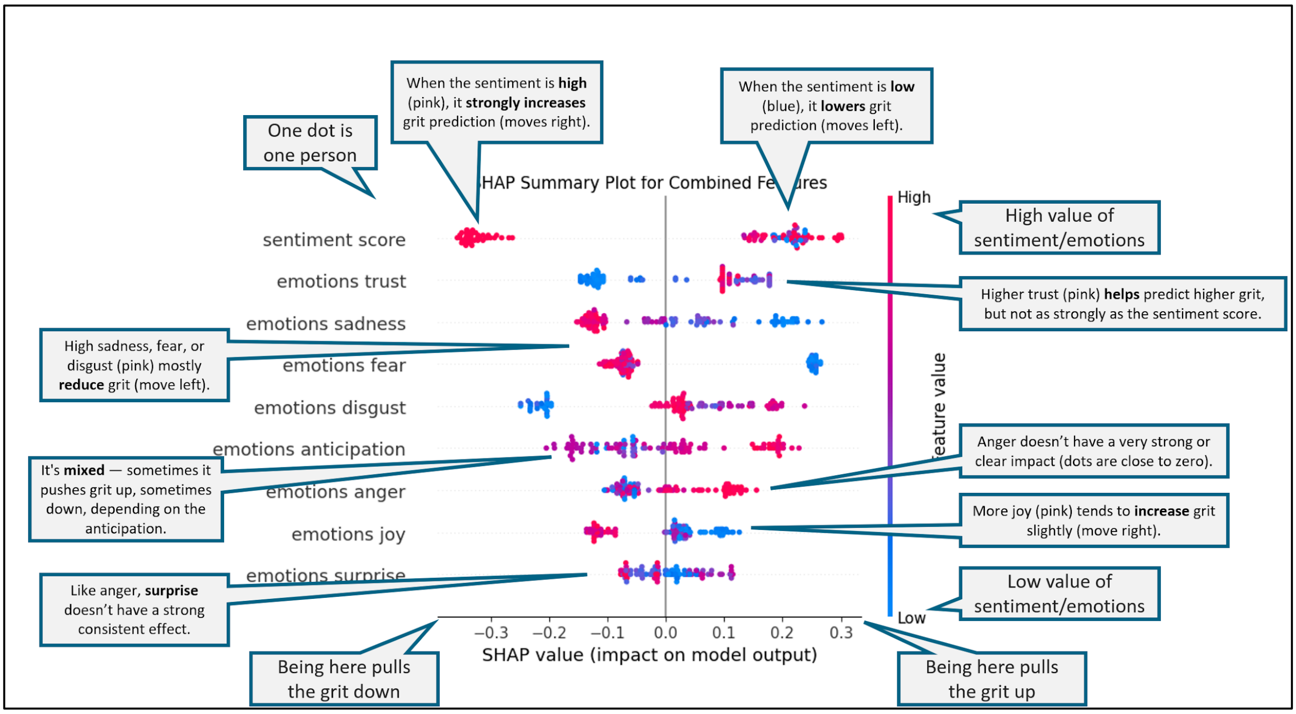
In the original analysis presented during the project dissemination (see WP5 results presentation), the SHAP dot plots were used to highlight directional effects of individual features. For example, trust generally increased predicted grit, while sadness and fear decreased it. Emotions such as anticipation were context-dependent—in some cases, high anticipation led to increased grit predictions, but the relationship was not strictly linear. Emotions like anger and surprise had smaller effects and tended to cluster around zero, indicating minimal or ambiguous impact on the final prediction.

Partial dependence plots extended this analysis by showing aggregate behavior across the full population. These plots indicated that while certain features such as arousal or dominance might not show high global importance in SHAP, their impact on grit was sometimes substantial at specific intervals. This type of interpretability provides key insights into the affective structure of participant responses.

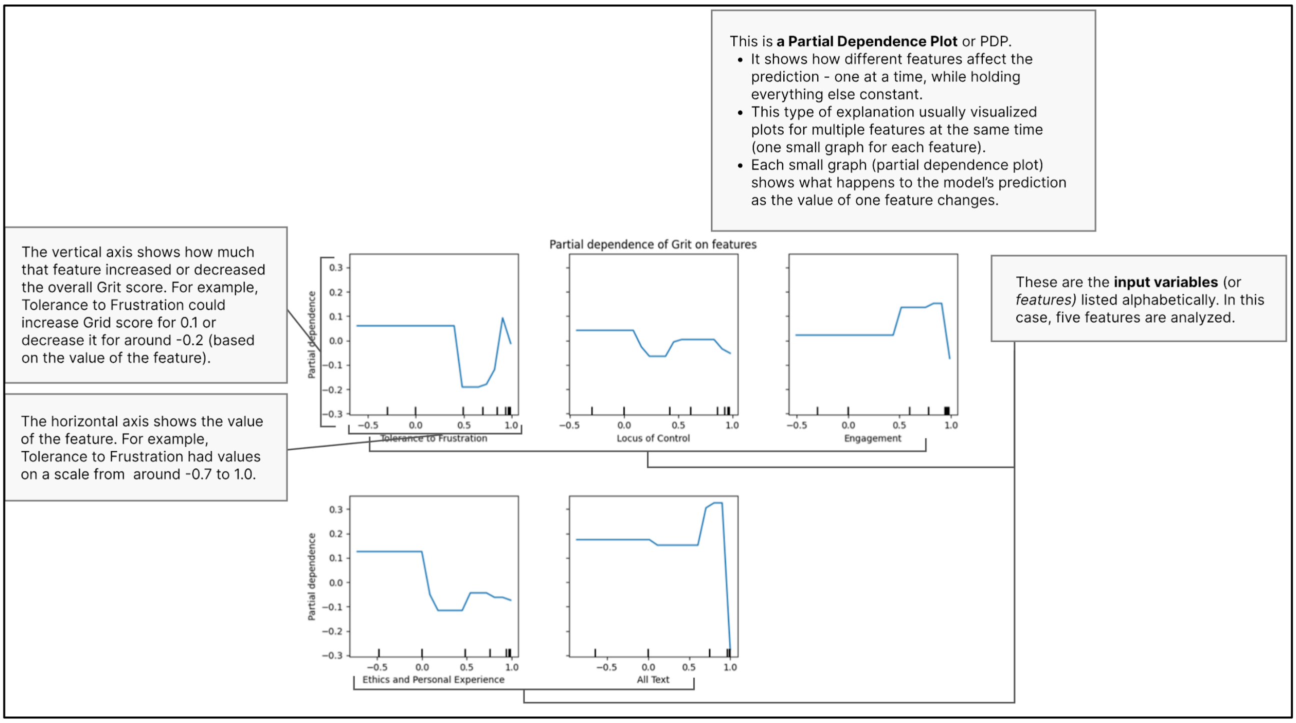
3.5 Final GRIT Model

The final model is a single, interpretable gradient boosting regressor trained on a curated subset of sentiment features extracted from the aggregated participant responses. It integrates emotion signals from multiple sources but prioritizes robust and interpretable indicators such as overall sentiment polarity and basic emotions (e.g., trust, fear, anticipation). It is capable of producing continuous grit scores that correlate well with the self-reported scores from participants.

The model is lightweight and suitable for real-time or offline integration into feedback systems or educational tools. It also enables individual-level interpretation, thanks to the SHAP framework, making it particularly useful for applications in student coaching, self-reflection environments, or adaptive learning platforms.



SHAP decision plots are particularly helpful in understanding individual predictions. These plots show the contribution of each feature for a specific participant, starting from the baseline (the average prediction for all participants). Features with positive SHAP values push the prediction higher, while those with negative SHAP values pull it lower. For example, a participant with high trust and positive sentiment may have a final grit prediction significantly above the mean, whereas another participant with high sadness and fear may see their grit estimate reduced. These individual visual explanations are especially powerful for educational feedback or interventions.



This modeling effort confirms that non-cognitive traits such as grit can be reliably inferred from the emotional and expressive content of narrative reflections. The combination of traditional sentiment analysis and modern explainable AI techniques allows for robust, interpretable prediction pipelines that preserve nuance without sacrificing transparency. Future extensions may explore longitudinal modeling, multi-language adaptation, or comparative analysis across institutions. The present model provides a reproducible, interpretable, and computationally efficient solution for assessing psychological attributes from natural language data.

4. Implications and application

This research confirms that sentiment and emotion features extracted from multilingual, open-ended student reflections can reliably predict grit levels, demonstrating a viable alternative to traditional psychometric surveys. By leveraging sentiment analysis and advanced interpretability methods such as SHAP and PDPs, the study provided transparent insights into not only which emotional features matter but precisely how they influence grit predictions.

The relationship between Grit and emotions is fundamental to understanding how people maintain motivation and long-term effort in demanding contexts. Evidence suggests that positive emotions encourage perseverance. Emotions such as enthusiasm, interest or hope act as an important emotional fuel that allows us to maintain our effort in a given task. In addition, positive emotions help us to reframe difficulties as opportunities, maintaining motivation. In turn, positive emotions are associated with greater emotional self-regulation, which promotes perseverance.

Unfortunately, people with high levels of grit are not immune to negative emotions such as frustration or anxiety. However, they are able to tolerate and manage them better. People with high levels of Grit tend to use more adaptive strategies of emotional regulation, such as cognitive reappraisal, rather than giving up or avoiding difficult situations. The personality factor "Grit" acts as an emotional regulator providing greater personal resilience, decreasing the impact of negative emotions on the person and helping them to reinterpret negative situations.

Focusing on our work, the data show a significant relationship between the personality model (Grit) and the manifestation of certain emotions. The results indicate that positive emotions (such as positivity and confidence) make a person score higher on the Grit scale (Datu, 2021; Musumari et al., 2018; Akbag & Ümmet, 2017; Lee, 2022). Conversely, negative emotions (such as sadness, fear and disgust) cause a person to score lower on the Grit scale. Extremely positive emotions and overly cheerful expressions seem to slightly reduce Grit scores. These results need further reflection and follow-up, as they could be caused by a social desirability effect.

The methodological integration of explainable AI significantly enhanced interpretability, providing granular, actionable insights that traditional assessment methods rarely deliver. Educational practitioners can leverage these findings to better understand the fine-grained impact of emotions influencing students' persistence and to develop targeted interventions aimed at fostering resilience and long-term goal commitment. Moreover, the ability to detect grit through sentiment analysis of naturally occurring textual data opens opportunities for scalable, unobtrusive assessment practices in educational settings, which can complement or partially substitute standard psychometric instruments. Ultimately, this study highlights the potential of sentiment analysis and XAI to transform the measurement and understanding of grit. By capturing emotional dynamics that underlie perseverance, this approach provides educators and researchers with deeper insights and practical tools to support students' academic success.

Conclusion

The relationship between Grit and emotions is fundamental to understanding how people maintain motivation and long-term effort in demanding contexts. Evidence suggests that positive emotions encourage perseverance. Emotions such as enthusiasm, interest or hope act as an important emotional fuel that allows us to maintain our effort in a given task. In addition, positive emotions help us to reframe difficulties as opportunities, maintaining motivation. In turn, positive emotions are associated with greater emotional self-regulation, which promotes perseverance. Unfortunately, people with high levels of grit are not immune to negative emotions such as frustration or anxiety. However, they are able to tolerate and manage them better. People with high levels of Grit tend to use more adaptive strategies of emotional regulation, such as cognitive reappraisal, rather than giving up or avoiding difficult situations. The personality factor "Grit" acts as an emotional regulator providing greater personal resilience, decreasing the impact of negative emotions on the person and helping them to reinterpret negative situations. Focusing on our work, the data show a significant relationship between the personality model (Grit) and the manifestation of certain emotions. The results indicate that positive emotions (such as positivity and confidence) make a person score higher on the Grit scale (Datu, 2021; Musumari et al., 2018; Akbag & Ümmet, 2017; Lee, 2022). Conversely, negative emotions (such as sadness, fear and disgust) cause a person to score lower on the Grit scale. Extremely positive emotions and overly cheerful expressions seem to slightly reduce Grit scores. These results need further reflection and follow-up, as they could be caused by a social desirability effect.

The findings of this research hold meaningful implications for educational practice, because leveraging sentiment analysis and XAI methods to measure grit, educators can more easily gain deeper insights into students' emotional states and their impact on perseverance. Specifically, recognizing that positive emotional expressions such as trust and anticipation strongly enhance grit offers practical opportunities for intervention. For example, educators might actively foster classroom environments that encourage trust-building activities, goal-setting discussions, and optimistic future-oriented feedback. On the other hand, understanding that negative emotions like sadness or fear diminish grit can help educators identify students who may require emotional or psychological support to build resilience. Incorporating sentiment analysis tools into routine formative assessments could allow for real-time, nuanced monitoring of students' grit, providing educators with actionable data to tailor individual support programs and ultimately improving academic persistence.

The GRIT project has demonstrated the feasibility and value of integrating psychological assessment, narrative reflection, and advanced AI techniques to explore non-cognitive skills in academic settings. By combining game-based learning with open-ended self-reflection and sentiment analysis, we have created a multidimensional framework capable of both engaging participants and extracting meaningful data on perseverance, emotional regulation, and personal motivation.

The sentiment-informed machine learning model developed is lightweight, interpretable, and suitable for use in educational tools or reflective environments. Importantly, the use of SHAP and

PDP techniques has ensured that each prediction can be explained, enhancing trust and usability. Beyond predictive modelling, this work underlines the importance of emotional insight and narrative depth in understanding student resilience and motivation. Future work may extend this approach to longitudinal tracking, multilingual support, and cross-cultural comparisons. Ultimately, GRIT WP5 has laid a solid foundation for integrating emotional analytics into human-centred educational technologies.

References

Allport, G. W. (1961). *Pattern and growth in personality*. Holt, Rinehart and Winston.
Caspi, A. (2000). The child is father of the man: Personality continuities from childhood to adulthood. *Journal of Personality and Social Psychology*, 78(1), 158–172.
Duckworth, A. (2016). *Grit: The power of passion and perseverance*. Scribner.
Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26–34.
McCrae, R. R., & Costa, P. T., Jr. (1985). Updating Norman's "adequate taxonomy": Intelligence and personality dimensions in natural language and in questionnaires. *Journal of Personality and Social Psychology*, 49(3), 710–721.
Norman, W. T. (1975). A five-factor taxonomy: Emergence of the Big Five. In D. W. Fiske & R. D. Williams (Eds.), *Annual Review of Psychology* (Vol. 26, pp. 1–26). Annual Reviews.
Plomin, R., & Caspi, A. (1999). Behavioral genetics and personality. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (2nd ed., pp. 251–276). Guilford Press.
Strelau, J. (1998). *Temperament: A psychological perspective*. Plenum Press.

Annexes

Glossary

Psychological Concepts

- **Grit**: A non-cognitive trait defined by perseverance and sustained passion toward long-term goals. It includes two components: (1) *Consistency of Interest*—maintaining the same goals and passions over time; and (2) *Perseverance of Effort*—the ability to continue working hard despite challenges or failure (Duckworth et al., 2007).

- **Growth Mindset**: The belief that abilities and intelligence can be developed over time through effort, strategies, and feedback (Dweck, 2006). A growth mindset supports persistence and resilience in the face of setbacks.

- **Perseverance**: Continued effort to do or achieve something despite difficulties, failure, or opposition. In psychology, it's closely related to conscientiousness and is a central part of grit.

- **Resilience**: The psychological capacity to adapt to stress and adversity. Resilient individuals are able to bounce back from failure or hardship without becoming overwhelmed or discouraged.

- **Self-Regulation**: The ability to monitor and control one's behavior, emotions, or thoughts, modifying them according to the demands of the situation to achieve long-term goals.

- **Emotional Regulation**: The strategies and processes people use to influence their own emotional experience. Adaptive strategies include *cognitive reappraisal* (reframing a situation), while maladaptive strategies include *suppression* or *avoidance*.

- **Cognitive Reappraisal**: A form of emotional regulation where a person changes their interpretation of a situation to alter its emotional impact.

- **Locus of Control**: A concept referring to individuals' beliefs about the causes of events in their lives. Those with an *internal locus* believe they control their own fate; those with an *external locus* attribute outcomes to outside forces like luck or authority figures.

- **Engagement**: A psychological state characterized by attention, curiosity, interest, and passion. In educational contexts, it's a predictor of motivation, achievement, and persistence.

- **Motivation (Intrinsic vs Extrinsic)**: *Intrinsic motivation* refers to doing an activity for its inherent satisfaction, while *extrinsic motivation* involves doing something to earn a reward or avoid punishment.

- **Frustration Tolerance**: The capacity to endure feelings of frustration or failure without giving up. High frustration tolerance is closely linked to perseverance and problem-solving ability.

- **Self-Efficacy**: The belief in one's own ability to succeed in specific situations or accomplish tasks. It strongly influences thought patterns, emotional reactions, and behavior.

- **Affect**: A broad psychological term referring to the experience of emotion, typically categorized as positive or negative. It underlies mood, emotional responses, and temperament.

- **Valence**: The intrinsic attractiveness (positive valence) or averseness (negative valence) of an emotional experience.

- **Arousal**: A physiological and psychological state of being alert or stimulated. High arousal emotions include excitement and anger; low arousal emotions include sadness and calmness.

- **Dominance**: The degree of control or influence one feels in a situation. In emotional modelling, dominance captures whether a person feels empowered or helpless.

- **Big Five Personality Traits**: A widely accepted model that describes personality in terms of five broad dimensions:

    - **Neuroticism**: Emotional instability, anxiety, moodiness.

    - **Extraversion**: Sociability, assertiveness, energetic behaviour.

    - **Openness to Experience**: Creativity, curiosity, openness to new ideas.

    - **Agreeableness**: Compassion, cooperation, trust.

    - **Conscientiousness**: Organization, dependability, goal-directed behaviour.

- **Trait vs State**: *Traits* are stable personality characteristics, while *states* are temporary emotional or psychological conditions.

- **Self-Reflection**: The process of introspection, where an individual examines their thoughts, feelings, and behaviours. In research, reflective tasks help participants articulate their inner experiences.

- **Narrative Psychology**: A field focusing on how individuals create meaning through personal stories. In the GRIT project, narrative responses were analyzed to understand emotional and motivational themes.

Research and Analytical Terms

- **Sentiment Analysis**: A computational method used to detect and classify emotional tone in textual data. It quantifies emotions like joy, anger, and sadness, or overall polarity (positive/negative).

- **VADER (Valence Aware Dictionary for Sentiment Reasoning)**: A rule-based model specifically attuned to sentiments expressed in social media and short text formats. Produces polarity scores and emotion probabilities.

- **VAD (Valence-Arousal-Dominance)**: A three-dimensional framework used to represent emotional states:

    - *Valence*: Degree of pleasantness.
    - *Arousal*: Intensity of emotion.
    - *Dominance*: Sense of control over the emotion.

- **Transformer Models**: Deep learning architectures (e.g., BERT, RoBERTa) used for natural language understanding. In this project, transformers were applied to classify emotions and extract semantic features from text.

- **Clustering**: A machine learning technique used to group data points based on similarity. In unsupervised learning, it helps uncover hidden structures in data.

- **Principal Component Analysis (PCA)**: A statistical method for reducing the dimensionality of data while preserving variance. Used for visualizing and interpreting patterns in participant responses.

- **SHAP (SHapley Additive exPlanations)**: A method for explaining the output of machine learning models. It assigns each feature a value representing its contribution to a prediction.

- **Partial Dependence Plot (PDP)**: A plot showing the marginal effect of a selected feature on the predicted outcome of a machine learning model.

- **MAPE (Mean Absolute Percentage Error)**: A performance metric that expresses prediction accuracy as a percentage. Lower values indicate better model performance.

- **MSE (Mean Squared Error)**: A common loss function in regression models that measures the average squared difference between actual and predicted values.

- **Likert Scale**: A psychometric scale commonly used in questionnaires to measure attitudes or opinions. In the GRIT scale, responses range from "Not like me at all" to "Very much like me."

- **Qualitative Coding**: The process of categorizing and interpreting narrative data based on themes, emotions, or concepts.

- **Multimodal Data**: Data collected from different formats or sources (e.g., text, gameplay interaction logs, emotional ratings), often used together for richer analysis.

Questionnaire: Angela Duckworth's Grit Scale

Responses: "Very much like me", "Mostly like me", "Somewhat like me", "Not much like me", "Not like me at all"

1. New ideas and projects sometimes distract me from previous ones.
2. Setbacks don't discourage me.
3. I have been obsessed with a certain idea or project for a short time but later lost interest.
4. I am a hard worker.
5. I often set a goal but later choose to pursue a different one.
6. I have difficulty maintaining my focus on projects that take more than a few months to complete.
7. I finish whatever I begin.
8. I am diligent.

Questionnaire: Reflective Open-Ended Game-Based Emotional and Strategic Reflection

Each answer to be written in free-form narrative based on the game experience

Frustration and Perseverance:

1. How did you handle setbacks during the game when puzzles took longer than expected to solve?

2. What emotions did you experience when faced with particularly difficult puzzles?

3. Can you describe a moment in the game where you felt frustrated but persisted? (e.g., with the Telegram bot, cyphers, or code combinations)

4. How did you manage emotions when the game didn't progress as you hoped?

5. What strategies did you use to manage your emotions while playing the game, especially during difficult moments?

6. Describe a moment in the game that made you feel particularly motivated or inspired. What emotions did you experience during this moment?

7. Reflect on a challenging puzzle you encountered during the game. How did you feel when you faced this challenge, and what motivated you to continue?

Locus of Control and Problem Solving:

8. How much control do you feel you have over the outcomes in the game?

9. To what extent do you believe your problem-solving skills contributed to your success in the game?

10. Did you feel external factors, such as luck or hints, played a significant role in your progress?

11. How did you adapt when circumstances in the game disrupted your strategy?

12. Reflect again on a challenging puzzle. How did you feel and what kept you going?

Engagement and Motivation:

13. Focus on a moment when you were completely immersed in the game. What factors contributed to your engagement?

14. Did the game's narrative make you feel emotionally connected to the characters? Why or why not?

15. How did exploring the game environment enhance your experience?

16. What motivated you to continue/complete the game's challenges and puzzles?

17. Did interactions with the Telegram Bot or open science tools increase your commitment to solving the mystery? How?

Ethical Reasoning and Reflection:

18. Did the game's narrative make you reflect on ethical dilemmas in research? Can you provide an example?
19. How did the portrayal of open science tools affect your understanding of ethical research practices?
20. Describe a situation in the game that challenged your sense of right and wrong. How did you resolve it?
21. What values do you think the game tried to communicate through its storyline and characters?
22. How did the game reinforce the importance of honesty and responsibility in research?
23. How did the game's narrative influence your emotional responses related to your own research experiences?
24. In what ways did the game encourage you to think about your long-term academic goals and aspirations?