



**Revealing the Secret to Successful Virtual Meetings:  
How Personality Traits Impact Conversational Involvement in a Group**

**Andy Bo Xuan Li<sup>1</sup>**

**Supervisor(s): Catholijn Jonker<sup>1</sup>, Masha Tsfasman<sup>1</sup>**

<sup>1</sup>EEMCS, Delft University of Technology, The Netherlands

A Thesis Submitted to EEMCS Faculty Delft University of Technology,  
In Partial Fulfilment of the Requirements  
For the Bachelor of Computer Science and Engineering  
June 23, 2023

Name of the student: Andy Bo Xuan Li  
Final project course: CSE3000 Research Project  
Thesis committee: Catholijn Jonker, Masha Tsfasman, Gosia Migut

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

## Abstract

With the rise in popularity of online conferencing platforms such as Zoom and Teams, there is an increasing need to understand the dynamics of virtual meetings to enhance productivity. One key aspect that can contribute to productive online meetings is conversational involvement. Personality traits influence the behaviour of participants within a group and therefore, have the potential to influence conversational involvement. This research paper aims to explore the impact of personality traits on conversational involvement in online meetings. To interpret personalities, the HEXACO model has been employed. Through a systematic investigation, this study provides valuable insights into how individual personalities influence engagement levels in a group during virtual meetings. The results of this study shed light on the relationship between personality traits and conversational involvement, thus contributing to the existing body of knowledge surrounding effective online meeting practices.

## 1 Introduction

Since the COVID-19 pandemic, video conferencing tools like Zoom and Teams have been used more frequently [1; 2]. Therefore, it is becoming more critical to understand the dynamics of virtual meetings between group participants. Kauffeld and Lehmann-Willenbrock [3] believe that participants within a group need to listen to each other and build upon each other's ideas to achieve productive meetings. Using conversational involvement in a group could help achieve such productivity [4]. In particular, verbal [5] as well as non-verbal signals, such as eye gaze [6], can provide valuable information about the level of conversational involvement of each participant.

Furthermore, personality influences how people interact with each other and can impact (social) engagement [7]. These (non-)verbal signals, and consequently conversational involvement, can be influenced by the personality of individuals within a group [8].

The structure of this paper is as follows. Section 2 provides background information supported by relevant literature, followed by the research questions. Next, in section 3, an explanation is provided about the MEMO corpus used for this research. Section 4 describes the methodology used to annotate conversational involvement and how the annotation and personality data have been processed. Furthermore, section 5 explains three different experimental setups, each corresponding to one of the three sub-questions. Section 6 presents the results, which is followed by the discussion in section 7 and the conclusion in section 8. Lastly, section 9 addresses the ethical aspects of conducting responsible research and the reproducibility of the research.

## 2 Background

Conversational involvement is defined as engagement in a conversational setting [9], which is why involvement and engagement will be used interchangeably. Gatica-Perez et al.

[10, p. 490] defined group involvement as “The perceived degree of interest or involvement of the majority of the group.” (Social) engagement has been defined by Sidner et al. [11, p. 123] as “The process by which interactors start, maintain, and end their perceived connections to each other during an interaction.” Based on these studies, it is known that conversational involvement is about the engagement of participants in social settings such as online conversation.

Besides, according to Maria et al. [12], individuals who share similar personality traits are more likely to establish relationships with similar others. The personality traits were: extroversion (X), Agreeableness (A), and Openness to experience (O). Maria et al. [12] have also found that people that scored high on Agreeableness (A) tended to communicate with people that scored high on extroversion (X). The dynamics of social signals, such as speech and eye gaze, are likely to be influenced by individual and social factors, such as personality traits [8; 12]. However, it is worth noting that the research conducted by Ivaldi et al. [8] focused on the context of Human-Robot Interaction (HRI) and examined the effect of personality on speech and eye gaze during assembly tasks.

This paper aims to investigate the influence of personality on conversational involvement in online meetings between human participants. To achieve this, we aim to answer the following questions:

- **Main research question:** *Does the personality of participants influence how involved they are in a conversation?*
- **Sub-question 1:** *What are the main personality traits that could influence the level of involvement in a conversation?* Given that extroversion is recognized as a fundamental dimension in shaping how individuals establish and maintain social relationships [13]. Thus participants with different levels of extroversion could potentially influence the conversational involvement in a group.
- **Sub-question 2:** *How does the composition of homogeneous or heterogeneous personality combinations influence conversational involvement in a group?* Homogeneous groups, such as those comprised entirely of introverted or extroverted individuals, exhibit different conversational styles [14]. This could influence conversational involvement, as this paper focuses on perceived conversational involvement.
- **Sub-question 3:** *How does the level of extroversion in a group relate to the level of involvement in a conversation?* Previous research suggests that individuals scoring high on extroversion (X) tend to be more talkative and active in a conversation [8; 15; 16].

## 3 MEMO Data Collection

Before analysing the effect of personality on the level of involvement in a conversation, video segments needed to be annotated on their conversational involvement. These video segments were taken from the corpus: multimodal longitudinal meeting corpus (MEMO) [17], which is a recently collected corpus on virtual group discussions on COVID-19.

The corpus consists of 53 participants ranging from ages 18 to 76 and 4 moderators ranging from ages 24 to 45, which are available in the data collection [17]. The purpose of the moderators was to keep the conversation flowing and encourage all participants to express their opinions as much as possible. Additionally, the participants were divided into 15 groups based on which COVID-19 risk group the participants were in, to maximize diversity in this aspect.

Furthermore, the participants filled in a pre-screening survey which contained the answers to a 24-item questionnaire. These answers were on a 5-point scale ranging from low with a score of 1 to high with a score of 5, which was used for the personality data.

Lastly, some groups had missing personality data for at least one participant. Additionally, group 1 was chosen not to be annotated, as the video segments were in speaker view instead of gallery view; meaning only the active speaker was visible during all the sessions. Therefore, those groups have been excluded from the analysis and results. Resulting in only ten of the fifteen groups being usable for research.

## 4 Methodology

### 4.1 Annotation Process

For the annotation of the video segments, they needed to be distributed among the researchers to annotate, on which the process will be discussed below. Followed by the annotation process in Elan (computer software program) [18]. Lastly, the inter-rater reliability scores will be discussed.

#### Distributing Annotation Load

There are approximately 32 hours of video content available for annotation, but due to time constraints, the annotators have decided to annotate only half of it, which is approximately 17 hours. However, assigning this workload to a single person within the given time constraint is impractical. As a result, it has been determined that the annotation work will be distributed among the four researchers. To facilitate this distribution, a Python script has been developed, which performs the following steps:

1. **Create five-second window samples from all the videos that have not been annotated before.**

Some video segments were excluded from annotation, as these annotations have been done by third parties before. These segments have been excluded based on a Comma Separated Values (CSV) file containing which time windows have been annotated before. Furthermore, the start and end are excluded based on a given timestamps CSV file, as these segments contain head-calibration instructions.

2. **Take a sub-sample of the entire list of time window samples.**

Randomly choose 12600 samples to distribute amongst the researchers. This corresponds to 5 minutes per researcher to annotate from each group session of which there are 42 group sessions to annotate.

3. **Randomly distribute these samples into four distinct sets.**

Each set corresponds to each researcher such that there is a 10% overlap in annotation, meaning 10% of all the samples are annotated by a pair of annotators.

4. **Create Elan files.**

Elan [18] has been used to annotate video segments. Using the four sets, the Python script creates Elan files for each annotator for all the time windows that correspond to the same group session.

#### Annotation in Elan

The level of involvement in a conversation will be ranked on a 5-point Likert scale based on the above-described definitions. As this was used during previous research [17]. The scores are the following:

- 1 - Very low group involvement.
- 2 - Low group involvement.
- 3 - Moderate group involvement.
- 4 - High group involvement.
- 5 - Very high group involvement.

A thing to note is that a higher score does not necessarily mean that more participants are involved. It is always about the level of involvement in a conversation for the vast majority of the group.

A screenshot of how the annotation process looks like in Elan [18] can be seen below in Figure 1.

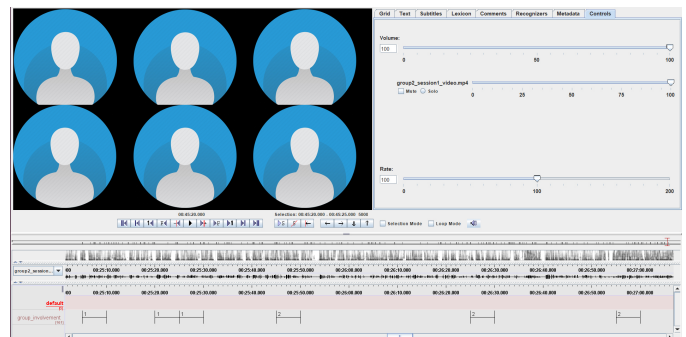


Figure 1: The screenshot showcases the annotation process in Elan, a computer software program. It displays the video in the top left corner, while the annotation intervals are located at the bottom.

On the top left in Figure 1, the video segment to annotate can be found of which the participants in a group can be seen. Furthermore, at the bottom, the timeline of the video can be seen of which some 5-second intervals contain the score of conversational involvement. These scores have manually been entered to annotate the conversational involvement of those specific 5-second video segments.

#### Inter-Rater Reliability Score

The inter-rater reliability score was calculated to measure the reliability of the ratings of the annotators. In Table 1, the inter-rater reliability scores can be found between pairs of annotators for each of the four annotators.

Raters	1	2	3	4
1	-	0.76	0.62	0.58
2	0.76	-	0.52	0.65
3	0.62	0.52	-	0.60
4	0.58	0.65	0.60	-

Table 1: The inter-rater reliability scores between different pairs of annotators using the Intraclass Correlation Coefficient (ICC).

According to Koo and Li [19], the ratings have moderate reliability as, on average, the Intraclass Correlation Coefficient (ICC) are within a range of 0.5 and 0.75. All sets have been used to include annotations from both lenient and strict annotators and not just the sets with the highest inter-rater reliability scores.

## 4.2 Data Processing of Conversational Involvement and Personality Traits

### Data Processing of Annotation Data for Conversational Involvement

The annotated Elan [18] files from all the researchers were converted to CSV files such that they can be processed with the use of a Python script. Furthermore, the inter-rater reliability score has been calculated to determine the reliability of the ratings. The way this has been calculated is with the use of the Intraclass Correlation Coefficient (ICC).

### Data Processing of Personality Traits Using the HEXACO Model

The personality of participants has been interpreted by using the HEXACO model [20]. The Brief HEXACO Inventory (BHI) has been used as an instrument that measures the six personality dimensions of the HEXACO model [21], as the BHI was used in previous research [17]. Other than BHI, there is the original instrument, the HEXACO-PI-R, which can also be used to measure the six personality dimensions of the HEXACO model [21]. The BHI has the same validity as HEXACO-PI-R as the BHI “re-estimates the original construct validity correlations of the HEXACO-PI-R with relatively great accuracy.” [21, p. 871].

Furthermore, the questions for the BHI have significantly fewer items with 24 items, compared to the 200 items for HEXACO-PI-R [21]. The ease of use is another advantage to using the BHI as an instrument as the questionnaire is self-assessed and shorter surveys are more likely to be filled in [22; 23].

## 5 Experimental Setup

In this section, the aspects that were considered for the analysis of conversational involvement will be described in section 5.1. Followed by a description of the three different experimental setups in section 5.2.

### 5.1 Conversational Involvement Aspects

When examining conversational involvement, several aspects can be considered. For this study, the aspect of frequency percentage was used. By focusing on frequency percentage rather than the overall average of conversational involvement

per group, we aimed to mitigate any bias that may arise from longer group sessions. Specifically, the frequency percentage of the involvement score of 5 was considered. This involvement score corresponds to the moments of highest conversational involvement.

## 5.2 Description of the Three Different Experimental Setups

### Experiment 1: Common Personality Traits for High Conversational Involvement

This experiment aimed to answer the sub-question: *What are the main personality traits that could influence the level of involvement in a conversation?* To set up this experiment, the HEXACO model [20] was used to calculate scores for all six dimensions of each participant. Subsequently, the median score was determined for each group, as opposed to using the mean, which could be skewed by extreme outliers and potentially distort the representation of the personality data. For each of the HEXACO dimensions, the mean and margin of error were computed to establish a 95% confidence interval. Afterwards, the HEXACO personality scores were examined for commonalities among the top-3 conversational involvement groups compared to the bottom-3 groups. To analyze the data, a Mann-Whitney U test was conducted for each of the HEXACO dimensions, as opposed to using a t-test, considering that the data did not satisfy the assumption of being normally distributed. Additionally, a Kruskal-Wallis test was not considered as the experiment involved comparing two means.

### Experiment 2: Groups with (Dis)similar Personalities

This experiment aimed to answer the sub-question: *How does the composition of homogeneous or heterogeneous personality combinations influence conversational involvement in a group?* To set up this experiment, only three of the six HEXACO scores, calculated in the previous experiment, were used. The specific dimensions that were looked at are the following three HEXACO dimensions: extroversion (X), Agreeableness (A), and Openness to experience (O). According to Maria et al. [12], the similarity-attraction paradigm could be verified for these personality traits. Maria et al. [12] have also found that people that scored high on Agreeableness (A) tended to communicate with people that scored high on extroversion (X).

Furthermore, the groups were ranked based on a similarity score. This similarity score was calculated using the Manhattan Distance (MD) formula:  $MD(p, q) = \sum_{i=1}^n |p_i - q_i|$ , which takes the sum of the absolute differences between two points. Based on this similarity score, the groups were ranked from low distance (similar) to high distance (dissimilar), of which the top-3 groups were compared to the bottom-3 groups in terms of their conversational involvement. Lastly, a t-test was performed to compare the distribution of the two groups.

### Experiment 3: Groups with High Levels of Extroversion

This experiment aimed to answer the sub-question: *How does the level of extroversion in a group relate to the level of involvement in a conversation?* To set up this experiment, the

median of each group has been ranked on their level of extroversion (X) from high to low to examine the difference in conversational involvement between the top-3 and bottom-3 groups ranked on extroversion (X). Finally, a t-test was performed to compare the distribution of the two groups.

## 6 Results

In this section, the results will be described for the three different experiments outlined in the previous chapter (section 5.2).

### Experiment 1: Common Personality Traits for High Conversational Involvement

For this experiment, we aimed to answer the sub-question: *What are the main personality traits that could influence the level of involvement in a conversation?* To do this, the top-3 groups were compared with the bottom-3 groups to determine if there are commonalities in personality traits among groups with high conversational involvement. These differences are illustrated in Figure 2. In this figure, each dimension of the HEXACO scores for the top-3 groups is plotted alongside the bottom-3 groups. The 95% confidence interval is represented for both groups, as well as the mean for each dimension of the HEXACO scores.

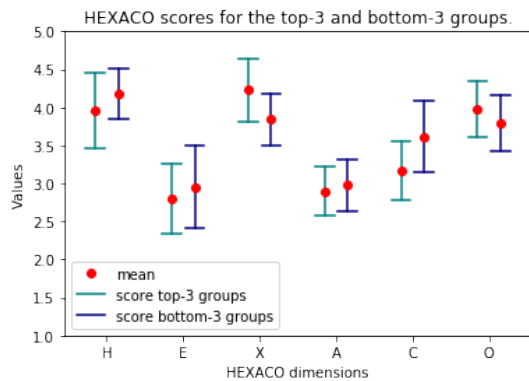


Figure 2: The HEXACO scores, along with their 95% confidence intervals, were compared between the top-3 and bottom-3 groups, which were ranked based on the highest frequency percentage for conversational involvement score 5.

Additionally, the disparity between the top-3 groups and the bottom-3 groups has been calculated. These differences are presented in Table 2 below.

For each dimension of the HEXACO model (Table 2), the confidence intervals were calculated to determine the 95% confidence interval within which the population mean lies. Additionally, the value of 0 is within the confidence intervals presented in Table 2, suggesting the possibility of no significant difference between the two groups [24]. This observation is further supported by the p-values in Table 2, which were obtained from the Mann-Whitney U test. All these p-values are greater than 0.05, indicating that the observed differences are not statistically significant.

Dimension	Difference	Conf. Interval	p-value
H	0.000	(-0.501, 0.038)	0.486
E	0.125	(-0.472, 0.163)	0.335
X	0.750	(0.143, 0.626)	0.054
A	-0.125	(-0.289, 0.134)	0.295
C	-1.000	(-0.712, -0.165)	0.663
O	0.125	(-0.057, 0.416)	0.185

Table 2: The difference in frequency percentages for the conversational involvement score 5, along with the 95% confidence interval and the corresponding p-value, was examined for each of the HEXACO dimensions between the top-3 and bottom-3 groups.

### Experiment 2: Groups with (Dis)similar Personalities

For this experiment, we aimed to answer the sub-question: *How does the composition of homogeneous or heterogeneous personality combinations influence conversational involvement in a group?* To do this, the top-3 most similar groups were compared to the bottom-3 most similar groups to examine the difference in conversational involvement between these two groups. Figure 3 displays the distribution of conversational involvement scores for the top-3 and bottom-3 most similar groups along with error bars for each score.

Distribution conversational involvement for the top-3 and bottom-3 groups.

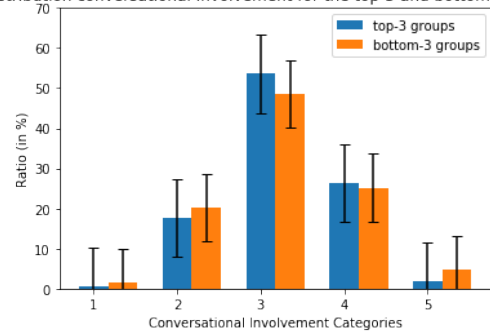


Figure 3: The distribution of conversational involvement scores for the top-3 and bottom-3 most similar groups in terms of the three HEXACO dimensions: extroversion (X), Agreeableness (A) and Openness to experience (O).

In Figure 3, it is apparent that the top-3 groups have a lower frequency percentage for the conversational involvement score of 5, which are the moments of highest conversational involvement, compared to the bottom-3 groups. The difference in this aspect is approximately 2.8%. Conversely, the top-3 groups exhibit a slightly higher frequency percentage (around 1%) for the conversational involvement score of 4 compared to the bottom-3 groups. The complete frequency percentage differences can be found in Table 3.

Involvement Score	1	2	3	4	5
Difference (in %)	-0.8	-2.6	5.0	1.3	-2.8

Table 3: The differences between the top-3 and bottom-3 most similar groups in terms of their conversational involvement scores are expressed in percentages.

Lastly, a t-test was used to compare the means of the two groups, resulting in a p-value of 0.966. This means that the results are insignificant since the p-value is greater than 0.05.

### Experiment 3: Groups with High Levels of Extroversion

For this experiment, we aimed to answer the sub-question: *How does the level of extroversion in a group relate to the level of involvement in a conversation?* To do this, the top-3 highest-scoring groups on extroversion (X) were compared with the bottom-3 groups on their frequency percentage of the conversational involvement score of 5. The conversational involvement score of 5 corresponds to the moments of highest conversational involvement. Figure 4 illustrates the distribution of conversational involvement scores for these groups.

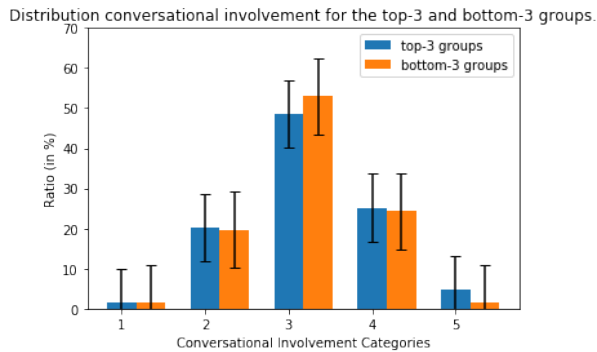


Figure 4: The distribution of conversational involvement scores for the top-3 and bottom-3 highest-scoring groups based on the level of extroversion (X).

In Figure 4, it is evident that the top-3 groups have a higher frequency percentage for the conversational involvement score of 5 compared to the bottom-3 groups. The difference between the two groups for the involvement score of 5 is approximately 3.1%, as indicated in Table 4. Table 4 displays the complete frequency percentage differences.

Involvement Score	1	2	3	4	5
Difference (in %)	-0.1	0.5	-4.3	0.8	3.1

Table 4: The differences in conversational involvement scores (expressed in percentages) between the top-3 and bottom-3 groups with the highest scores in extroversion (X).

Lastly, a t-test was used to compare the means of the two groups, resulting in a p-value of 0.552. This means that the results are insignificant since the p-value is greater than 0.05.

## 7 Discussion

In this chapter, section 7.1 presents possible explanations for the obtained results. These explanations delve into the factors that may account for the observed outcomes. Following that, a subsequent section (section 7.2) discusses the limitations of the conducted research. This section highlights the constraints and potential factors that may have influenced the findings of the research.

### 7.1 Possible Explanation for the Results

As mentioned in section 6, the results are statistically insignificant. Therefore, the findings may not hold significant meaning.

For the first experiment, we aimed to answer the sub-question: *What are the main personality traits that could influence the level of involvement in a conversation?* To do this, the top-3 groups were compared with the bottom-3 groups to determine if there are commonalities in personality traits among groups with high conversational involvement. Of which the three highest-scoring groups in terms of conversational involvement on the median had a 0.75 higher score in extroversion (X) compared to the three lowest-scoring groups. This observation may be attributed to how engagement or involvement has been defined as “the process by which interactions start, maintain, and end their perceived connections to each other during an interaction” [11, p. 123]. This definition aligns with the notion that extroversion is fundamental in shaping how individuals establish and maintain social relationships [13]. Additionally, during the annotation process, we observed practices exhibited by these extroverted individuals that may align with practices of high conversational involvement. As extroverted individuals tend to be more talkative [8; 15; 16], utilize hand gestures more frequently during communication compared to their introverted counterparts [25], and employ facial expressions, such as smiling [26], during communication.

Furthermore, when examining the median scores, the top three groups showed a 1.0 lower score in terms of Conscientiousness (C) compared to the bottom three groups. Conscientiousness (C) generally pertains to an individual’s social behaviour and inclination to act in various circumstances [27]. Higher levels of Conscientiousness (C) are often associated with characteristics such as being organized, self-disciplined and careful whereas low levels are associated with laziness, recklessness, irresponsibility and impulsiveness [15]. It is plausible that the lower score in Conscientiousness (C) can be explained by the nature of the three HEXACO dimensions: extroversion (X), Conscientiousness (C) and Openness to experience (O), as each of them influences the level of engagement within a social context [28].

In the second experiment, we aimed to answer the sub-question: *How does the composition of homogeneous or heterogeneous personality combinations influence conversational involvement in a group?* To do this, the top-3 most similar groups were compared to the bottom-3 most similar groups to examine the difference in conversational involvement between these two groups. Of which the average frequency percentage for the score of 5 for the three most similar groups was approximately 2.8% lower compared to the three most dissimilar groups. Perhaps the fact that homogeneous groups have different conversational styles [14] did not affect the perceived conversational involvement. Or perhaps the similarity-attraction paradigm is limited to the context of the conducted research, which was in the context of an online social network [12]. Alternatively, the phenomenon of individuals with low levels of engagement having the highest Openness to experience (O) score [29], combined with the possibility that highly extroverted individuals exhibit practices as-

sociated with high conversational involvement, as mentioned earlier, could explain why the top three similar groups had a lower frequency of high conversational involvement compared to the top three dissimilar groups.

In the third experiment, we aimed to answer the sub-question: *How does the level of extroversion in a group relate to the level of involvement in a conversation?* To do this, the top-3 highest-scoring groups on extroversion (X) were compared with the bottom-3 groups in the frequency percentage of the conversational involvement score of 5. Of which the average frequency percentage for the score of 5 was approximately 3.1% higher for the three highest-scoring groups on extroversion (X) compared to the three lowest-scoring groups. Perhaps, this observation could be explained by the fact that extroverted individuals tend to be more talkative and actively engaged in conversations [8; 15; 16]. Additionally, as mentioned earlier, extroverted individuals often exhibit practices that could be associated with high conversational involvement, further supporting the higher frequency percentage of score 5 in the groups that score highest in extroversion (X).

## 7.2 Limitations

In this subsection, some limitations will be discussed that may have influenced the results and should be taken into account.

Firstly, there is a limitation regarding the annotation process, as only a fraction (half) of the data was annotated. This limitation arose due to the time constraints of the research, which was conducted within a period of fewer than 10 weeks.

Furthermore, the inter-rater reliability score indicated that the ratings had moderate reliability instead of good or excellent reliability. As well as the fact that the results were statistically insignificant for the three different experiments. This raises questions about the validity of the results and suggests potential inconsistencies in the annotations.

Additionally, it is important to note that the findings of this research are limited to the conversational context, specifically focusing on COVID-19-related topics rather than encompassing a broader range of subjects.

Besides, the participating groups mostly consisted of extroverted individuals, as the vast majority of participants scored towards the upper end of the extroversion (X) scale. The lowest extroversion score available was 2.75, indicating slight introversion, while the highest score observed was 5.0, which represents extreme extroversion (X). This difference in extroversion (X) scores may have influenced the results.

Lastly, there are limitations associated with the accuracy of the HEXACO personality traits, as the questionnaires were self-assessed. This self-assessment method may introduce biases or inaccuracies in the measurement of personality traits.

## 8 Conclusions and Future Work

In this paper, we examined the influence of personality traits on the level of involvement in an online conversation. To answer this, we first needed to answer the following sub-question: *What are the main personality traits that could influence the level of involvement in a conversation?* The most

prominent differences between the top three and bottom three groups were that the level of Conscientiousness (C) was lower for the top three groups compared to the bottom three groups, and the level of extroversion (X) for the top three groups was higher compared to the bottom three groups. The second sub-question was: *How does the composition of homogeneous or heterogeneous personality combinations influence conversational involvement in a group?* We have shown that dissimilar groups had a slightly higher frequency percentage of the highest moments of conversational involvement. The similarity of the groups was determined by the similarity of three HEXACO dimensions: extroversion (X), Agreeableness (A), and Openness to experience (O). The third sub-question was: *How does the level of extroversion in a group relate to the level of involvement in a conversation?* We have shown that the top three highest-scoring extroversion (X) groups had a slightly higher frequency percentage for the highest moments of conversational involvement.

Overall, Conscientiousness (C) and extroversion (X) are identified as personality traits that could potentially influence conversational involvement. The groups with the highest conversational involvement frequency percentage for the score of 5 scored lower on Conscientiousness (C) and higher on extroversion (X) compared to the bottom groups. Conversely, the groups with the highest scores on extroversion (X) also had a higher frequency percentage for the score of 5 regarding conversational involvement. Thus these groups had a higher percentage of the highest moments of conversational involvement compared to the bottom three groups.

For future research, it would be interesting to explore the annotation of the perceived conversational involvement of each individual within a group to determine if personality traits can influence how we perceive individual conversational involvement. This would provide valuable insights into the subjective and individual aspects of involvement and its relationship with personality traits. By examining the perceived involvement of individuals, researchers can further understand the nuances of conversational dynamics and how personality traits may shape our perceptions of others' involvement. This line of inquiry could shed light on how personality traits influence our interpretations and evaluations of conversational dynamics.

## 9 Responsible Research

The provided data (MEMO corpus) was already collected from previous research as described in section 3. However, as this data has not been published yet and may contain sensitive information, the provided data can only be used during the ten weeks for the Research Project (CSE3000) course, which lasted from the beginning of April to the end of June. It is important to adhere to the General Data Protection Rules (GDPR) [30] when using this data. To gain access to the data for the research during the 10 weeks, adherence to the specified rules was formalized through a signed statement.

### 9.1 Ethics

Re-identification of the data can pose a possible risk for the video segments as well as the results from the questionnaire.

Therefore, any personal information that may lead to re-identification will not be included in this report.

To mitigate confirmation bias, great care needed to be taken to avoid annotating in favour of the research question. It was purposely chosen to first annotate the video segments according to the agreed-upon definition of conversational involvement. Only afterwards were the HEXACO personality traits of the participants analysed, to prevent favouring the research towards the research question(s).

Additionally, there is the risk of personal bias, as there are four different annotators. To mitigate this risk, it was decided to randomly select segments to annotate, which were then distributed among the four researchers, each of whom had different research questions.

It is also important to note the bias of Western culture, as the video segments annotated for conversational involvement mainly feature residents from the United Kingdom and were annotated by four researchers who are residents of Europe.

## 9.2 Reproducibility

The videos used to annotate conversational involvement and the Elan files with the annotations, as well as the code used to interpret personality traits cannot be shared, as they may contain sensitive information about the participants and could potentially enable re-identification of the participants. However, the methodology of how this process was carried out can be found in section 4 and can be used to reproduce the process with different data, such as annotations for different video segments and potentially other participants with different personalities.

## Acknowledgements

The data for the ICC was provided by Ana Hobai. I want to thank my colleagues within the research project group: Ana Hobai, Sebastiao Holtreman and Matthijs Rijm. Furthermore, I want to thank Masha Tsfasman for her feedback on the process of writing this research paper and Catholijn Jonker for their feedback during this research process.

## References

- [1] E. Hartsell and J. Lane, "Initial lessons learned during a remote drug court evaluation during the covid-19 pandemic," *Journal of Crime and Justice*, vol. 45, no. 5, pp. 627–645, 2022.
- [2] C. Tudor, "The impact of the covid-19 pandemic on the global web and video conferencing saas market," *Electronics*, vol. 11, no. 16, p. 2633, 2022.
- [3] S. Kauffeld and N. Lehmann-Willenbrock, "Meetings matter: Effects of team meetings on team and organizational success," *Small group research*, vol. 43, no. 2, pp. 130–158, 2012.
- [4] C. Oertel, "Towards developing a model for group involvement and individual engagement," *In Proceedings of the 2013 ACM on International Conference on Multimodal Interaction*, pp. 349–352, 2013.
- [5] B. Wrede and E. Shriberg, "Spotting "hot spots" in meetings: Human judgments and prosodic cues," *In Conference of the International Speech Communication Association*, vol. 2003, pp. 2805–2808, 2003.
- [6] C. Oertel and G. Salvi, "A gaze-based method for relating group involvement to individual engagement in multimodal multiparty dialogue," *In International Conference on Multimodal Interfaces*, pp. 99–106, 2013.
- [7] H. Salam, O. Celiktutan, I. Hupont, H. Gunes, and M. Chetouani, "Fully automatic analysis of engagement and its relationship to personality in human-robot interactions," *IEEE Access*, vol. 5, pp. 705–721, 2017.
- [8] S. Ivaldi, S. Lefort, J. Peters, M. Chetouani, J. Provasi, and E. Zibetti, "Towards engagement models that consider individual factors in hri: On the relation of extroversion and negative attitude towards robots to gaze and speech during a human-robot assembly task: Experiments with the icub humanoid," *International Journal of Social Robotics*, vol. 9, no. 1, pp. 63–86, 2017.
- [9] C. Oertel, G. Castellano, M. Chetouani, J. Nasir, M. Obaid, C. Pelachaud, and C. Peters, "Engagement in human-agent interaction: An overview," *Frontiers in Robotics and AI*, vol. 7, no. 92, 2020.
- [10] D. Gatica-Perez, L. McCowan, D. Zhang, and S. Bengio, "Detecting group interest-level in meetings," *In Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, vol. 1, pp. 489–492, IEEE, 2005.
- [11] C. L. Sidner, C. Lee, and N. Lesh, "Engagement when looking: behaviors for robots when collaborating with people," in *Diabruck: Proceedings of the 7th workshop on the Semantic and Pragmatics of Dialogue*, pp. 123–130, Citeseer, 2003.
- [12] J. Maria Balmaceda, S. Schiaffino, and D. Godoy, "How do personality traits affect communication among users in online social networks?," *Online Information Review*, vol. 38, no. 1, pp. 136–153, 2014.
- [13] D. W.-L. Wu, W. F. Bischof, N. C. Anderson, T. Jakobsen, and A. Kingstone, "The influence of personality on social attention," *Personality and Individual Differences*, vol. 60, pp. 25–29, 2014.
- [14] J. Tolins, K. Liu, Y. Wang, J. E. F. Tree, M. Walker, and M. Neff, "A multimodal motion-captured corpus of matched and mismatched extravert-introvert conversational pairs," 2016.
- [15] M. C. Ashton and K. Lee, "The hexaco model of personality structure and the importance of the h factor," *Social and Personality Psychology Compass*, vol. 2, no. 5, pp. 1952–1962, 2008.
- [16] M. Jensen, "Personality traits and nonverbal communication patterns," *International Journal of Social Science Studies*, vol. 4, no. 5, 2016.
- [17] M. Tsfasman, K. Fenech, M. Tarvirdians, A. Lorincz, C. Jonker, and C. Oertel, "Towards creating a conversational memory for long-term meeting support: predict-



ing memorable moments in multi-party conversations through eye-gaze,” *In International Conference on Multimodal Interaction*, pp. 94–104, 2022.

- [18] “Elan (version 6.4) [computer software].” <https://archive.mpi.nl/tla/elan>, 2022.
- [19] T. K. Koo and M. Y. Li, “A guideline of selecting and reporting intraclass correlation coefficients for reliability research,” *Journal of chiropractic medicine*, vol. 15, no. 2, pp. 155–163, 2016.
- [20] A. Feher and P. A. Vernon, “Looking beyond the big five: A selective review of alternatives to the big five model of personality,” *Personality and Individual Differences*, vol. 169, p. 110002, 2021.
- [21] R. E. De Vries, “The 24-item brief hexaco inventory (bhi),” *Journal of Research in Personality*, vol. 47, no. 6, pp. 871–880, 2013.
- [22] B. Burchell and C. Marsh, “The effect of questionnaire length on survey response,” *Quality and quantity*, vol. 26, no. 3, pp. 233–244, 1992.
- [23] P. Edwards, I. Roberts, P. Sandercock, and C. Frost, “Follow-up by mail in clinical trials: does questionnaire length matter?,” *Controlled clinical trials*, vol. 25, no. 1, pp. 31–52, 2004.
- [24] Statology, “Confidence Interval for the Difference Between Means.” <https://www.statology.org/confidence-interval-difference-between-means/>, 6 2020.
- [25] S. O’Carroll, E. Nicoladis, and L. Smithson, “The effect of extroversion on communication: Evidence from an interlocutor visibility manipulation,” *Speech Communication*, vol. 69, pp. 1–8, 2015.
- [26] M. Radeke and A. Stahelski, “You’re smiling, you must be an extrovert: The use of photographs to assess facial expressions and personality traits,” *National Social Science Journal*, vol. 44, no. 1, pp. 48–61, 2015.
- [27] B. W. Roberts, J. J. Jackson, J. V. Fayard, G. Edmonds, and J. Meints, “Conscientiousness,” 2009.
- [28] M. C. Ashton and K. Lee, “Objections to the hexaco model of personality structure—and why those objections fail,” *European Journal of Personality*, vol. 34, no. 4, pp. 492–510, 2020.
- [29] H. Yildiz Durak, “Role of personality traits in collaborative group works at flipped classrooms,” *Current Psychology*, pp. 1–21, 2022.
- [30] B. Wolford, “What is GDPR, the EU’s new data protection law?.” <https://gdpr.eu/what-is-gdpr/>, 2022.