Abstract

When confronted with the sentence *"Some parrots are birds"* the listener is forced to interpret the sentence either as *"Some but not all parrots are birds"* or *"Some and possibly all parrots are birds"*. This is known as a scalar implicature where you could either have a logical interpretation (true) or a pragmatic interpretation (false). That forces the listener to derive the enriched meaning of what is said and what is implied. Scalar implicatures are well-studied in the field of *Individual Differences*. This paper focuses on reanalyzing data from "What second-language speakers can tell us about pragmatic processing" [8] which investigated the scalar implicature processing among L2 English speakers and how it is influenced by their proficiency and Theory of Mind abilities. The reanalysis focuses on the Exploratory Analysis of individual differences and their effect on implicature derivation. This analysis focuses on whether I can replicate their findings or if I come to additional findings that were not in the scope of the paper. With this, I hope to gain more insights into the role of Individual Differences as this was not the main goal of the reference paper.

1 Introduction

In everyday conversation, speakers and listeners need to navigate the interplay of literal and implied meanings. Listeners often expect the speaker to be as informative as possible without misleading them. For example, if a professor tells their students *"Some of you passed the exam"* they would likely assume that not everyone passed, even though the word *some* technically includes the possibility that everyone could have passed [8]. This interpretation relies on scalar implicatures where listeners infer additional meaning based on the word choice [9, 10, 15]. Scalar implicatures are often associated with the cognitive abilities of the person, specifically *Theory of Mind* (ToM). ToM refers to the ability to attribute mental states to themselves and others and understand that they could differ from their own. This is crucial for interpreting not just the literal meaning of words but also the intended subtext. The listener needs to keep track of the beliefs of the other person and needs to combine that knowledge with their own beliefs and possibly other clues, whether they are from a linguistic or non-linguistic background.
The interplay of ToM with scalar implicature derivation has been demonstrated in studies such as those by Fairchild and Papafragou [6], who found that individuals with stronger ToM abilities are more likely to interpret statements pragmatically rather than literally. This raises the question: does this tendency hold in a second language context, and how does language proficiency interact with personality measures? With the given data, that includes 11 personality measures, this study aims to identify which traits are most significant and how these may vary with language proficiency.

2 Individual Differences in scalar implicature derivation in L2 context

The exploration of scalar implicatures in L2 speakers remains relatively unexplored. Slabakova was among the first to focus on this area [16]. She tested to which extent Korean L1 speakers and L2 learners with varying proficiencies would interpret under-informative sentences. The experiment was divided into two parts: contextualized and non-contextualized. The contextualized experiment consists of under-informative sentences, similar to scalar implicatures, that were provided with a context, like a picture story. Whereas the non-contextualized consisted just of under-informative sentences. In both cases, the participants were tasked with evaluating the truth value of these sentences. The results showed that L2 learners, no matter their proficiency, were more likely to reject under-informative sentences than their L1 counterparts. She proposed that L2 learners have enhanced executive functions and additional abilities and therefore exhibit better responses on metalinguistic tasks. Consequently, the study concluded that scalar implicatures are generated automatically and pose no significant challenges to L2 learners.
Building on this, Dupuy and colleagues [5] investigated the pragmatic inference derivation in French monolinguals and bilingual adults with upper-intermediate proficiency levels. Bilingual participants were tested in equal parts in their first and second languages. They used contextualized picture stories with under informative sentences similar to Slabakova [16]. Their results indicated that the bilingual participants had an identical proportion of pragmatic responses in their L1 and L2. They were also more likely to derive pragmatic meanings than the French monolinguals. They attributed these findings to the heightened awareness of pragmatic cues due to language switching, resulting in enhanced pragmatic abilities.
Those papers highlighted the role of proficiency in L2 but did not put them into perspective with ToM. This is what the original analysis aimed at and what I want to further explore with my analysis. Proficiency seems to play a significant role for L2 speakers, but we can't disregard the findings that were made for L1 speakers. It is expected that the ToM abilities not only affect L1 speakers but also L2 learners. Therefore this analysis aims at finding the degree to which ToM abilities interact with L2 proficiency level in implicature derivation.

3 Materials

3.1 Truth-value judgment task

The participants were given under-informative sentences like *"Some Parrots are Birds"* and control sentences:

Reference	Example sentence	Appropriate response
T1	Some parrots are birds	?
T2	Some birds are parrots	True
T3	Some parrots are fish	False
T4	All parrots are birds	True
T5	All birds are parrots	False
T6	All parrots are fish	False

Table 1: Example sentences used in the experiments

The participants were asked to answer as quickly as possible. First they were shown the fixation cross which then was replaced by the sentence, which was displayed word by word shortly after another.

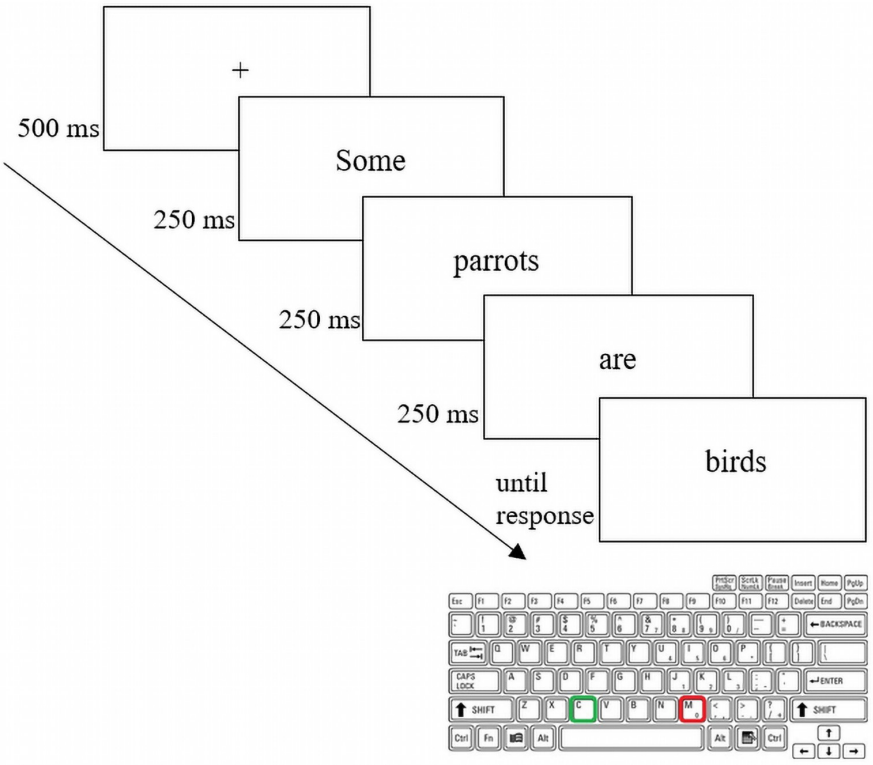

Figure 1: An example of the display of the test sentences including a fixation cross, a series of words making up a categorical sentence, and a response keyboard.

3.2 Personality Measurements

As explanatory measurements, Khorsheed et al. [8] used three self-assessment personality tests. These add up to 11 personality measures.

3.2.1 Autism Spectrum Quotient (AQ)

The AQ [2] measures the degree to which adults show autism-like traits. It consists of 50 questions about communication, social skills, strong attention to detail, and imagination. It is used to asses how Individual Differences on the social-skill subscale change the tendency to derive pragmatic interpretations. It is expected that the lower the social skill, the less pragmatic the person.

3.2.2 Systemizing questionnaire (SQ-R)

The SQ-R [18] measures Individual Differences in systemizing i.e. analyzing systems, extracting rules and predicting outputs. Systemizing individuals are associated with attention to detail and wanting to find the exact truth. We expect that listeners may base judgment on "statistical patterns" to find the likelihood that a potential interpretation is relevant to the speaker's intended meaning.

3.2.3 Big Five Inventory (B5)

The B5 [12] was built to model the relationship between personality and academic behavior. It measures 5 dimensions of personality:

Personality Dimensions	**Associated Traits**
Extraversion	talkativeness, assertiveness, energy
Agreeableness	sympathetic, helpful, forgiving, cooperative
Conscientiousness	Organized, thorough, cautious, responsible
Neuroticism	easily anxious, tense, unstable, emotional
Openness	imaginative, intelligent, artistic, curious

Table 2: Personality dimension defined by B5 and their associated traits

4 Analysis and results

4.1 Data handling and exclusion

For a good comparison, I analyzed the data in R, using the lme4 package [3] to fit linear regression models, mirroring the methodology of Khorsheed et al. [8]. Therefore, I excluded participants with less than 70% accuracy on the control conditions (T2-T6). For the analysis, regression models were fitted using only data from condition T1 and varying combinations of personality measures plus their proficiency levels (high vs. low). This approach is designed to comprehensively assess the interactions between personality traits, proficiency in a second language, and the ability to derive scalar implicatures.

4.2 Principal component analysis (PCA)

Given the extensive number of personality measures involved in this study (11 in total), it is plausible to expect some to be correlated. To mitigate multicollinearity in our regression models I used Principal Component Analysis (PCA) [13, 11] to reduce the number of variables. The initial examination of the correlation matrix suggests a potential reduction in variable size. The correlation matrix is illustrated in Figure 2 below.

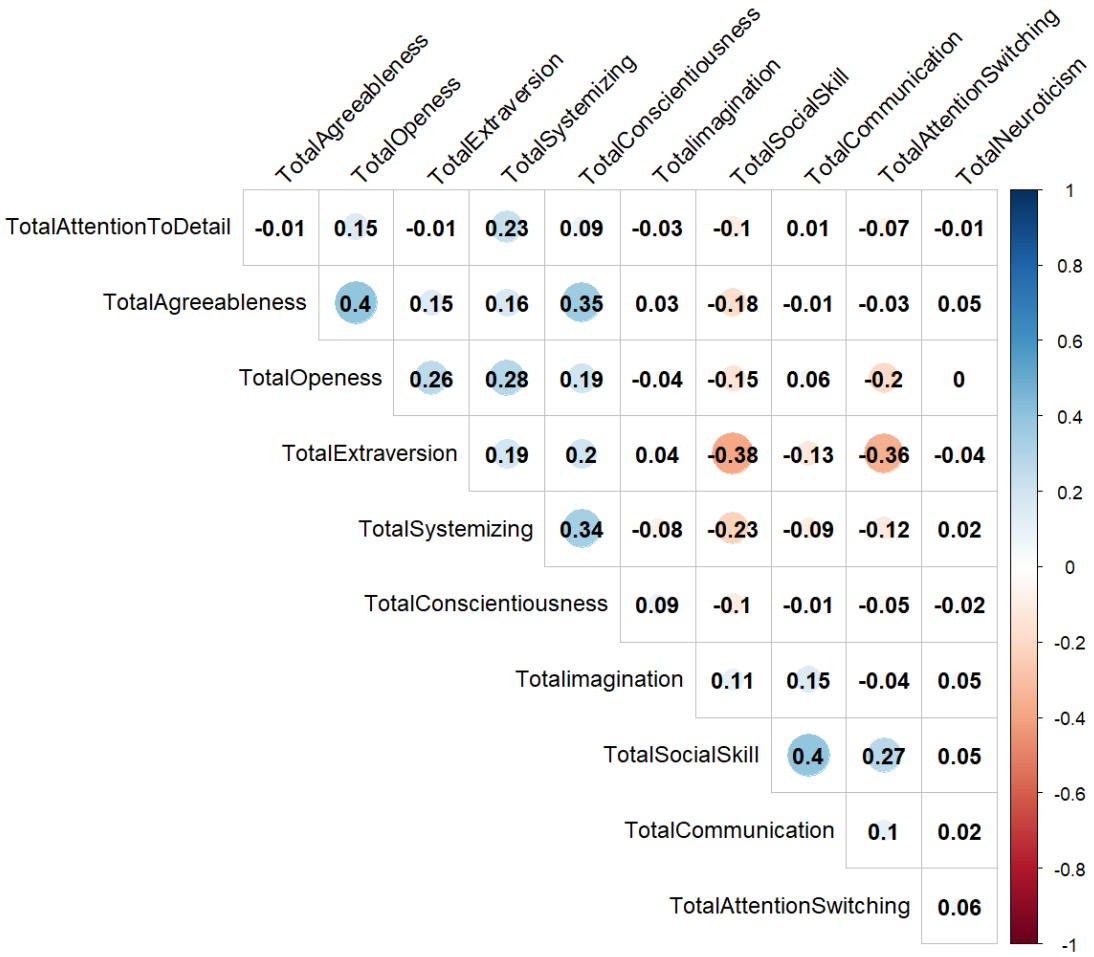

Figure 2: Correlation Matrix of all Personality measures

Although outlier exclusion influenced the directions of the eigenvectors, it did not significantly alter the principal components. I opted not to strictly follow the Kaiser criterion, which suggests retaining components with eigenvalues over 1.

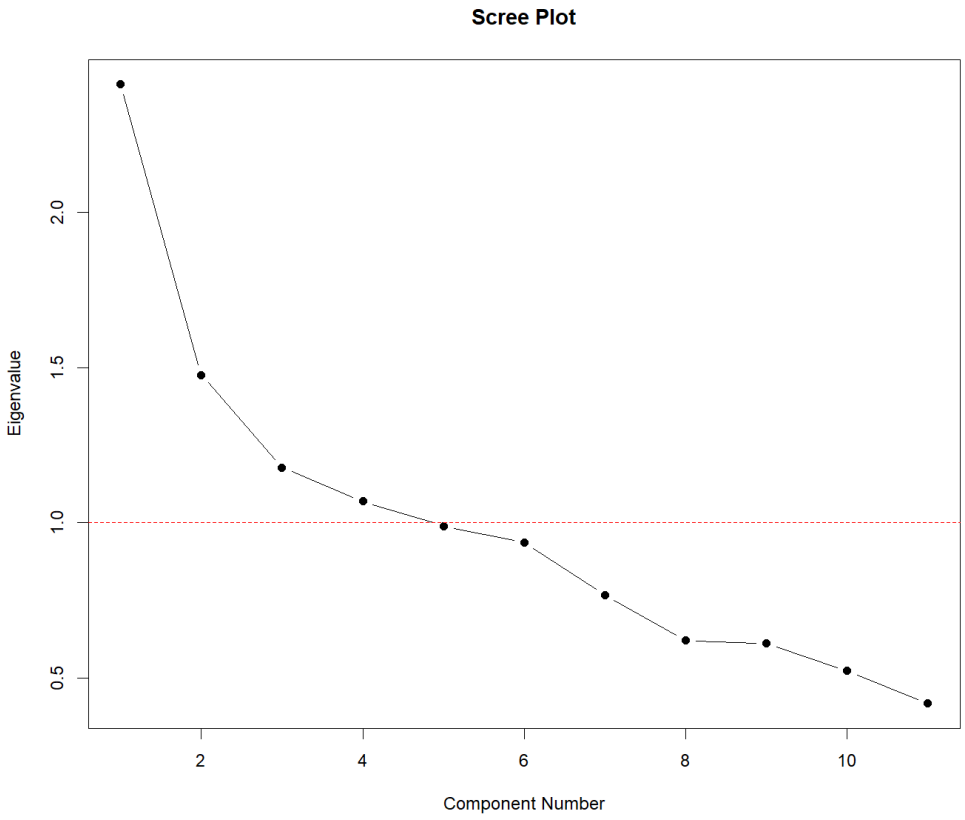

Figure 3: Scree Plot to visualize the Kaiser-Criterion and the Scree Test

Instead, all principal components with eigenvalues down to 0.94 were retained. This decision would align with the ambiguous results from the Scree test. Therefore I decided to use six principal components. With this approach, I obtained the loadings detailed in table 4.2. Notably, "Openness" appears in RC1 and RC6, suggesting residual correlations between these variables.

Principal Component	**Personality Measurement**
RC1	Negative Attention switching, Positive extraversion (openness)
RC2	Social Skill, Communication
RC3	Imagination
RC4	Attention to Detail, Systemizing
RC5	Neuroticism
RC6	Imagination, Agreeableness, Conscientiousness (openness)

Table 3: The table show which Principal Component (RC1-RC6) represent which personality measure

This provides insights into the interrelations among different personality traits, confirming some expected associations such as the link between social skills and communication, or systemizing and attention to detail.

4.3 K-Means Clustering

To identify groups of individuals with similar personalities and implicature derivation rates, I used the K-means clustering algorithm. To find the best number of clusters I utilized the Elbow method [17] and the Silhouette method. The Silhouette method [14] suggested two clusters, while the Elbow method, although somewhat ambiguous, generally supports this suggestion. Further testing validated that two clusters achieved well-separated groups (Figure 4), with no other number of clusters showing similar separation.
However, this clustering analysis did not reveal any new insights into the data. As shown in table 4, there weren't many notable personality differences and the rates of pragmatic derivation were similar.

	1	2
TotalSocialSkill	2.859.649	47.708.333
TotalAttentionToDetail	5.631.579	48.437.500
TotalAttentionSwitching	5.078.947	62.708.333
TotalCommunication	3.315.789	39.687.500
Totalimagination	3.763.158	37.500.000
TotalSystemizing	76.552.632	580.312.500
TotalExtraversion	27.447.368	229.062.500
TotalAgreeableness	34.166.667	294.479.167
TotalConscientiousness	30.728.070	259.479.167
TotalNeuroticism	23.061.404	240.208.333
TotalOpeness	37.692.982	326.666.667
Proficiency	1.561.404	15.104.167
Pragmatic=0,Logic=1	0.417154	0.4108796

Table 4: The table shows the mean values of each measurement in the respective Cluster. Note: these values are scaled and only show a relative difference.

In further experimentation, I attempted clustering using the principal components. However, this approach did not result in clearly defined clusters. This indicates that the personality traits of our participants are intertwined in a way that makes it difficult to find clear segmentation.

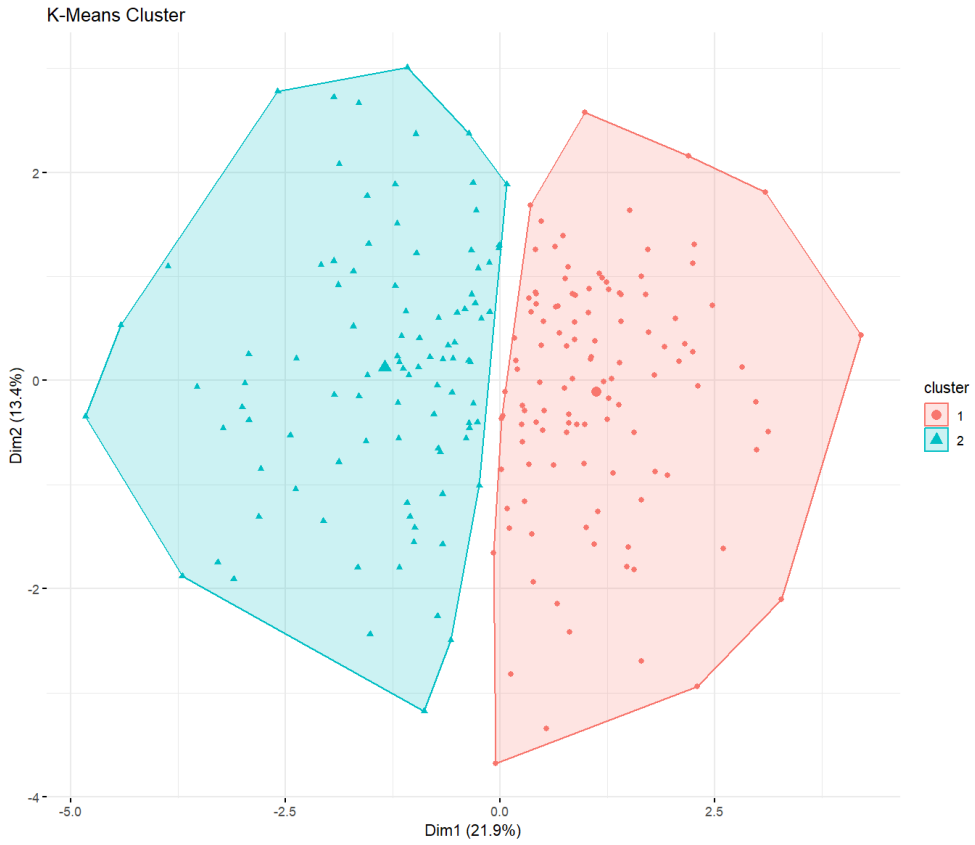

Figure 4: Two clusters computed my K-Means over the 11 personality measurements and their answer on the T1 condition

4.4 Interaction of Proficiency with Social Skill & Imagination

As mentioned in the Introduction, the role of proficiency and its interplay with Individual Differences remain underexplored. It has been suggested by Fairchild and Papafragou [6] that people with inferior ToM are more literal in their interpretation. In our dataset, ToM is represented in the Social Skill subscale of the AQ. For simplicity, the analysis was confined to T1 conditions. To explore further potential influences I modeled the accuracy with all 11 personality measurements as fixed effects. Of these, only 4 showed significance: Social Skill ($p = 0.000176$), Attention To Detail ($p = 0.001618$), Imagination ($p = 0.00000000336$) and Systemizing ($p = 0.000465$).
Given the numerous predictors, I employed a Bonferroni correction. This meant dividing the value of our alpha threshold by 11, i.e. the number of total comparisons. This left me with a new alpha threshold of 0.004. It's worth noting that Attention to Detail and Systemizing load into the same principal component and show a correlation in the matrix (Figure 2).
Thus I modeled two models: one using the 4 significant variables and one replacing Attention to Detail and Systemizing with their respective principal component (RC4). Surprisingly,

when using the principal component, these variables showed no significance ($p = 0.639585$), while they did individually (Attention to Detail $p = 0.002599$, Systemizing $p = 0.003518$). I followed this up by considering the interaction of each of these variables with proficiency. Surprisingly all of them were significant for our threshhold ($z = 2.772921, p = 0,005$).

4.4.1 Social Skill

As visualized in Figure 5 there is a difference among individuals with high proficiency: the more social the individual, the more likely they are to derive a pragmatic inference. In contrast, at low proficiency levels both groups tended to be more logical. I hypothesize that people with higher proficiency were able to rely on their intuition and the tone of the sentences, as their executive functions are not as taxed as those with lower proficiency.

4.4.2 Imagination

Individuals with high imagination skills tend to be more logical in their interpretation regardless of their proficiency. With low imagination, high-proficiency subjects are more likely to derive pragmatically compared to their counterparts. I hypothesize that the cognitive strategies employed by individuals with high proficiency and low imagination differ from those with low proficiency, resulting in variations in their pragmatic interpretation.

4.4.3 Systemizing

In high proficiency Systemizing appears to have no influence, both groups have no preference for logical or pragmatic inference. However, at low proficiency levels, lower Systemizing skills result in more logical responses. I hypothesize that to compensate for their lack of proficiency, they rely more heavily on their Systemizing skills and are thus more pragmatic as they compute the likelihood of what the sentence could mean. Alternatively, like the PCA analysis showed, Systemizing may not be significant, and the observed differences could just be noise.

4.5 Attention to Detail

Attention to Detail has the opposite effect on high proficiency levels than on low levels. In high proficiency, high Attention to Detail leads to a more pragmatic approach, whereas in low proficiency, it results in more logical responses. I propose that the subject of Attention to Detail shifts with proficiency. From social cues and tone at high proficiency to grammatical cues at low proficiency. Alternatively, Attention to Detail might not be significant, aligning with the results observed in the PCA analysis, and the observed difference could be merely noise.

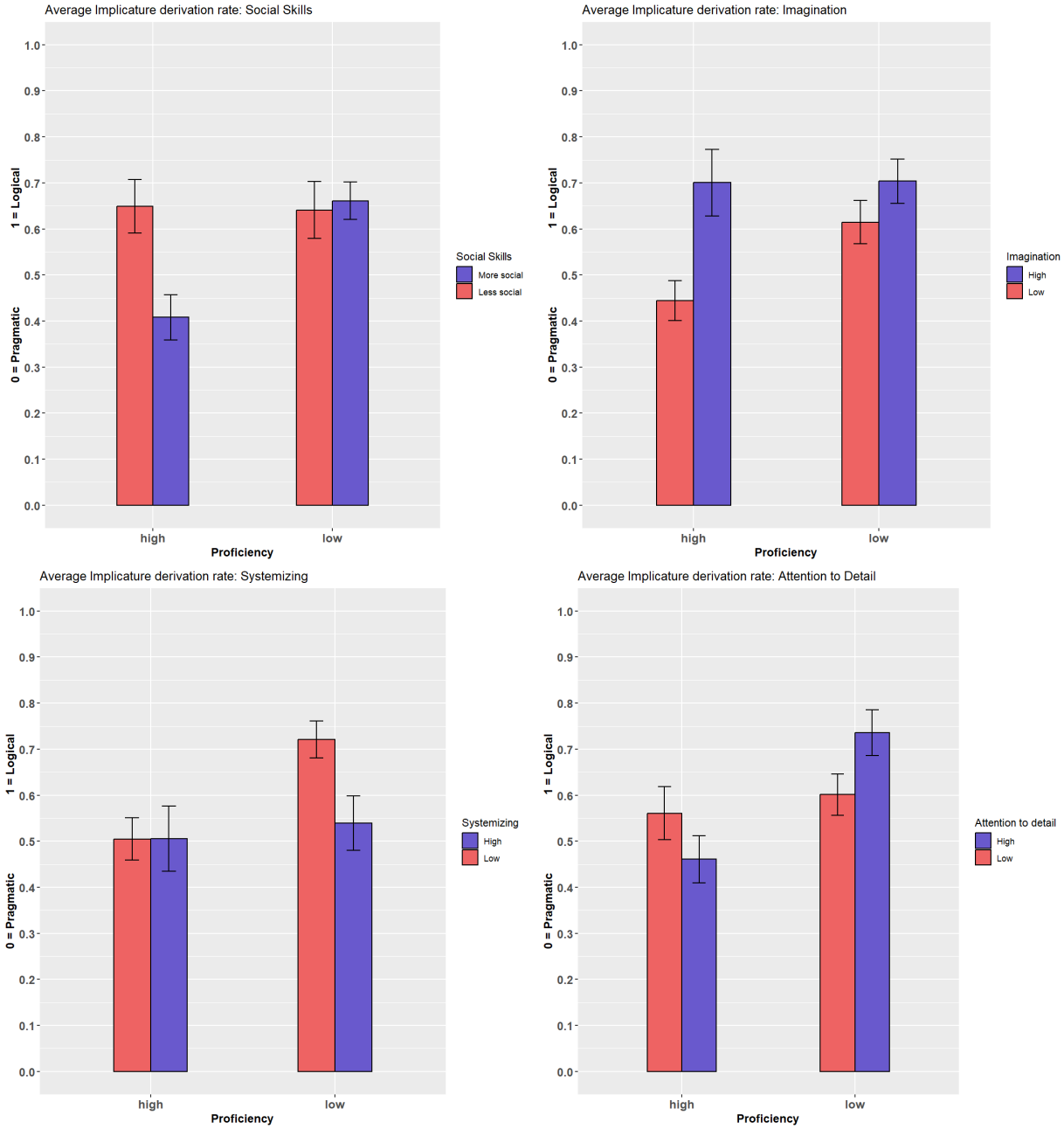

Figure 5: Implicature derivation rate

4.5.1 Comparative Analysis with the Original Study

In contrast to the original data analysis, I did not incorporate random effects into my model due to their slow convergence and issues with singularity. They accounted for by-item random effects, which I deemed negligible, as the items should have been designed to convey similar

messages with minimal variability. Including random effects on subjects caused my models to overfit, leading me to exclude them from my analysis. This methodological divergence influenced the outcomes compared to the original analysis.
For the first model containing all personality measurements, they obtained an alpha threshold of 0.002 whereas mine was 0.004. This did not change which predictors were significant in the model. Differences came with the subsequent analysis. In their following analysis, only Social Skill and Imagination met their threshold, whereas in mine Attention to Detail and Systemizing were also significant. After testing for the effect of imagination only Social Skills showed significant interaction in their analysis.
Due to the complexity of their models and ongoing convergence issues, they had to analyze both predictors separately. In the reanalysis, I combined their models and maintained their model complexity to test if a combined approach would change their outcome. The model, similar to theirs', has singularity issues but supports their findings. The interplay between Social Skills and proficiency does not meet the threshold but is still significant. Whereas the interplay with imagination is negligible.

4.5.2 Rewriting the Original Code

In an attempt to retain the flexibility of the original model while avoiding singularity issues, I rewrote their code. Singularity may occur due to multicollinearity in the data, so I replaced the predictors with the principle components. However, this adjustment did not mitigate the issue. It's possible that accounting for random effects on both items and subjects was overly complex, so I opted to remove the random effects for items based on the reasoning mentioned earlier. Unfortunately, even this modification did not mitigate singularity.

5 Discussion

This reanalysis aimed at finding new insights and validating previous findings on Individual Differences in L2 Learners. Previous studies suggest that Social skills improve the computing of pragmatic inferences [6] and that L2 learners face no significant challenges in deriving such inferences [16]. Building on these insights, Khorsheed et al.'s analysis [8] focused on how L2 learners process scalar implicatures, while my reanalysis aimed to delve deeper into the exploration of Individual Differences in L2 learners.
According to established theories in pragmatics(e.g. [7, 4]) , the computation of scalar implicatures relies on an evaluation of the speaker's epistemic state. Recent studies on scalar implicatures suggest that those with poor ToM abilities are less likely to compute pragmatic inferences and tend to be more literal in their interpretations of under-informative sentences [6, 1]. Khorsheed et al.'s [8] analysis supports this notion, revealing that participants with stronger ToM abilities, as measured by Baron-Cohen's Social-Skill subscale [2], were more likely to derive scalar implicatures compared to those with weaker ToM abilities. This effect was especially visible with increased L2 proficiency, but this interaction did not

meet their strict significance criterion. This reanalysis supports the effect of Social skills in high proficiency but not in low proficiency. It suggests that not only Social skill shows significance but also Imagination, Attention to Detail, and Systemizing skills. However, these findings lack support and I was not able to find satisfying explanations.
Also, it is important to note that this study was limited to female participants, and the distinction between low and high proficiency levels was based on adjacent levels of their language test scale. To strengthen our findings future studies should include a more diverse range of participants and consider a broader proficiency spectrum.

6 Conclusion

Overall, I was able to support Khorsheed et al.'s [8] findings and shed light on possible new research topics. This included replicating the interaction of Social skills with proficiency while using a less flexible analyzing approach. But also revealed the possibility that not only Social skills and proficiency influence pragmatic interpretations, but also Imagination, Attention to Detail and Systemizing possibly in combination with proficiency. My hypothesis suggests that the varying personality measures manifest differently on both proficiency levels, possibly because of limited executive functions.

References

[1] Clara Andrés-Roqueta and Napoleon Katsos. "The Contribution of Grammar, Vocabulary and Theory of Mind in Pragmatic Language Competence in Children with Autistic Spectrum Disorders". In: *Frontiers in Psychology* (2017). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2017.00996. URL: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2017.00996.

[2] Skinner R. Baron-Cohen S. Wheelwright S. "The Autism-Spectrum Quotient (AQ): Evidence from Asperger Syndrome/High-Functioning Autism, Males and Females, Scientists and Mathematicians". In: *Journal of Autism and Developmental Disorders* 01 (Feb. 2022). DOI: 10.1023/A:1005653411471. URL: https://doi.org/10.1023/A:1005653411471.

[3] Douglas Bates et al. *lme4: Linear Mixed-Effects Models using 'Eigen' and S4*. R package version 1.1-30. 2022. URL: https://CRAN.R-project.org/package=lme4.

[4] Robyn Carston. "Quantity maxims and generalised implicature". In: *Lingua* 96.4 (1995), pp. 213–244. ISSN: 0024-3841. DOI: https://doi.org/10.1016/0024-3841(95)00016-S. URL: https://www.sciencedirect.com/science/article/pii/002438419500016S.

[5] L Dupuy et al. "Pragmatic abilities in bilinguals". In: *Linguistic Approaches to Bilingualism* 9.2 (2019), pp. 314–340. URL: https://doi.org/10.1075/lab.17017.dup.

[6] Sarah Fairchild and Anna Papafragou. "The Role of Executive Function and Theory of Mind in Pragmatic Computations". In: *Cognitive Science* 45.2 (2021), e12938. DOI: https://doi.org/10.1111/cogs.12938. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/cogs.12938. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/cogs.12938.

[7] H. P. Grice. "Logic and Conversation". In: Leiden, Niederlande: Brill, 1975, pp. 41–58. ISBN: 9789004368811. DOI: 10.1163/9789004368811_003. URL: https://brill.com/view/book/edcoll/9789004368811/BP000003.xml.

[8] Ahmed Khorsheed et al. "What second-language speakers can tell us about pragmatic processing". In: *PLOS ONE* 17 (Feb. 2022). DOI: 10.1371/journal.pone.0263724. URL: https://doi.org/10.1371/journal.pone.0263724.

[9] Mikhail Kissine. "Pragmatics, cognitive flexibility and autism spectrum disorders". In: *Mind & language* 27.1 (2012), pp. 1–28.

[10] Ira Noveck. *Experimental Pragmatics: The Making of a Cognitive Science*. Cambridge University Press, 2018.

[11] Karl Pearson. "LIII. On lines and planes of closest fit to systems of points in space". In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2.11 (1901), pp. 559–572. DOI: 10.1080/14786440109462720. eprint: https://doi.org/10.1080/14786440109462720. URL: https://doi.org/10.1080/14786440109462720.

[12] Beatrice Rammstedt. "The 10-Item Big Five Inventory". In: *European Journal of Psychological Assessment* 23.3 (2007), pp. 193–201. DOI: 10.1027/1015-5759.23.3.193. eprint: https://doi.org/10.1027/1015-5759.23.3.193. URL: https://doi.org/10.1027/1015-5759.23.3.193.

[13] William Revelle. *psych: Procedures for Psychological, Psychometric, and Personality Research*. R package version 2.2.5. 2023. URL: https://CRAN.R-project.org/package=psych.

[14] Peter J. Rousseeuw. "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis". In: *Journal of Computational and Applied Mathematics* (1987), pp. 53–65. ISSN: 0377-0427. DOI: https://doi.org/10.1016/0377-0427(87)90125-7. URL: https://www.sciencedirect.com/science/article/pii/0377042787901257.

[15] Walter Schaeken, Marie Van Haeren, and Valentina Bambini. "The Understanding of Scalar Implicatures in Children With Autism Spectrum Disorder: Dichotomized Responses to Violations of Informativeness". In: *Frontiers in Psychology* 9 (2018). ISSN: 1664-1078. DOI: 10.3389/fpsyg.2018.01266. URL: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2018.01266.

[16] Roumyana Slabakova. "Scalar implicatures in second language acquisition". In: *Lingua* 120.10 (2010). Morphological variation in Japanese, pp. 2444–2462. ISSN: 0024-3841. DOI: https://doi.org/10.1016/j.lingua.2009.06.005. URL: https://www.sciencedirect.com/science/article/pii/S0024384109001727.

[17] Robert L. Thorndike. "Who belongs in the family?" In: *Psychometrika* (Dec. 1, 1953), pp. 267–276. ISSN: 1860-0980. DOI: 10.1007/BF02289263. URL: https://doi.org/10.1007/BF02289263.

[18] S. Wheelwright et al. "Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ)". In: *Brain Research* 1079.1 (2006). Multiple Perspectives on the Psychological and Neural Bases of Understanding Other People's Behavior, pp. 47–56. ISSN: 0006-8993. DOI: https://doi.org/10.1016/j.brainres.2006.01.012. URL: https://www.sciencedirect.com/science/article/pii/S0006899306000722.