

Tabel 3.3 Cohens' Theory of Effect Size

Upon conducting the t-test, the researcher proceeded with a supplementary non-parametric analysis. This decision was necessitated by the data's failure to meet the normality assumption, confirmed by the Kolmogorov-Smirnov test. Therefore, the Wilcoxon Signed-Rank Test was also administered to examine the pre-test and post-test differences in students' writing performance.

The concurrent application of both parametric and non-parametric tests, performed using SPSS, ensured the methodological rigor and robustness of the findings, especially given the inherent data characteristics.

## **CHAPTER FOUR**

### **FINDING AND DISCUSSION**

#### **4.1 Findings**

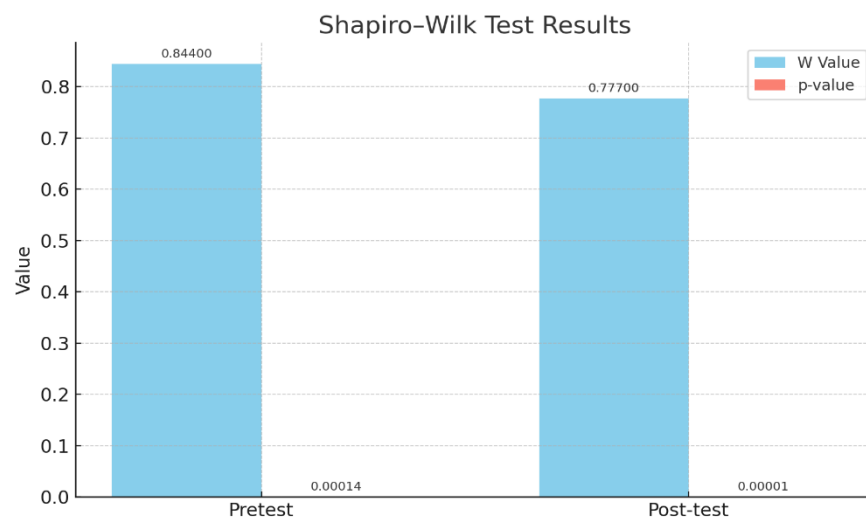
This section discusses the research findings and their validity, established through quantitative statistical tests and qualitative participant interviews. We collected data via pre-test and post-test administrations to sample groups, augmented by in-depth interviews. The sample comprised two groups: a control group that

learned procedure text writing through traditional face-to-face instruction, and an experimental group that used Discord as the primary teaching platform for this skill.

#### 4.1.1 The Quantitative Research

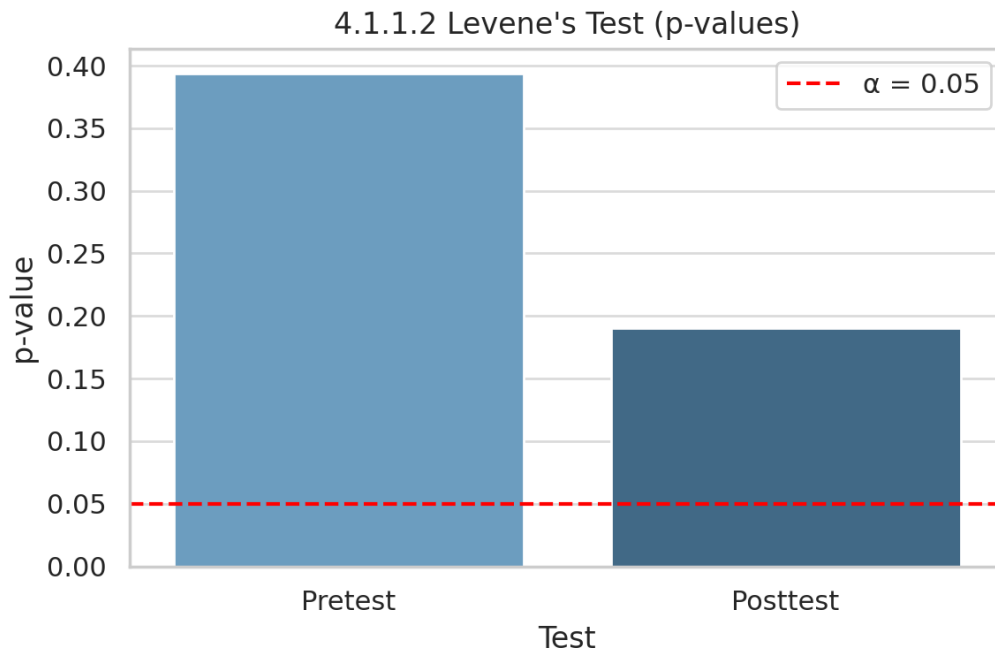
In analyzing the data obtained from the quantitative research, it is crucial to first ensure that the data distribution meets the assumptions required for statistical testing. One of the fundamental assumptions in parametric analysis is normality, which determines whether the data follow a normal distribution. To examine this assumption, the Shapiro-Wilk test is widely employed, as it is considered more reliable and powerful for small to medium sample sizes compared to other normality tests (Razali & Wah, 2011). Therefore, before proceeding to further statistical analysis, the normality of the data needs to be tested using the Shapiro-Wilk method.

##### 4.1.1.1 Shapiro-Wilk Test (Normality Test)

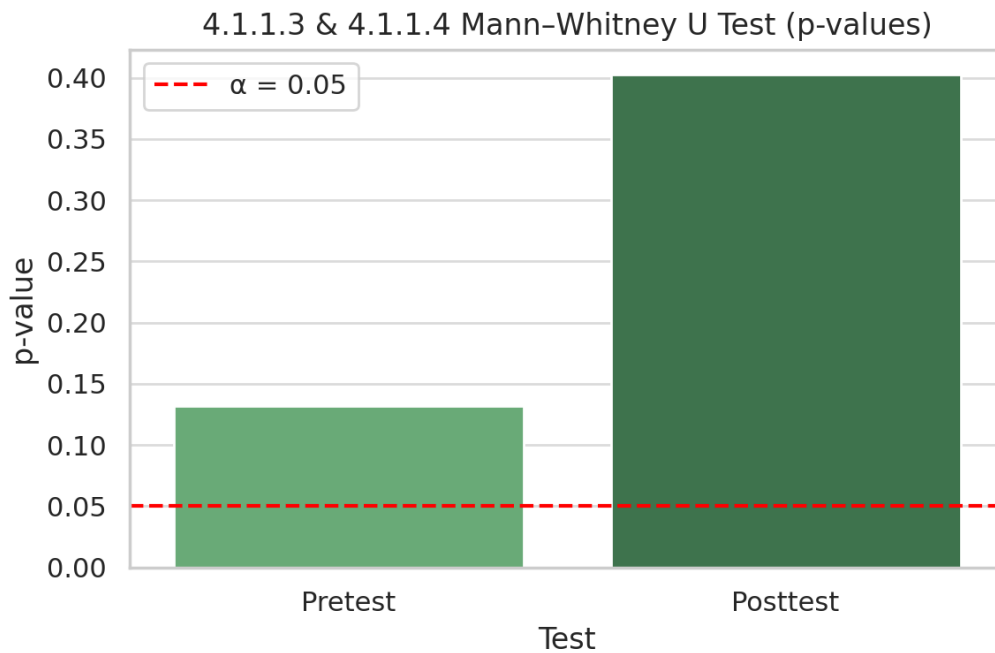


The normality assumption of the data was tested using a Shapiro-Wilk test on both pretest and post-test scores. The analysis revealed that the data were not normally distributed, with the pretest scores yielding a W value of 0.844 ( $p=0.00014$ ) and the post-test scores a W value of 0.777 ( $p=0.0000059$ ). As both p-values were less than 0.05, the null hypothesis of a normal distribution was rejected for both variables. Therefore, in adherence to statistical recommendations for non-normally distributed data

(Ghasemi & Zahediasl, 2012; Razali & Wah, 2011), non-parametric tests were employed in all further analyses to ensure the strength and accuracy of the statistical conclusions. 4.1.1.2 Levene's Test (Homogeneity of Variance Test)



Levene's Test was used to determine if the variances of the pretest and post-test scores were comparable between the experimental and control groups. The analysis revealed no statistically significant differences in variances. Specifically, for the pretest,  $W = 0.746$  ( $p=0.394$ ), and for the post-test,  $W = 1.789$  ( $p=0.190$ ). As both p-values were greater than the conventional alpha level of 0.05, the assumption of homogeneity of variance was satisfied, thus supporting robust group comparisons in subsequent statistical procedures. 4.1.1.3 Mann-Whitney U Test (Pretest Comparison Between Groups)

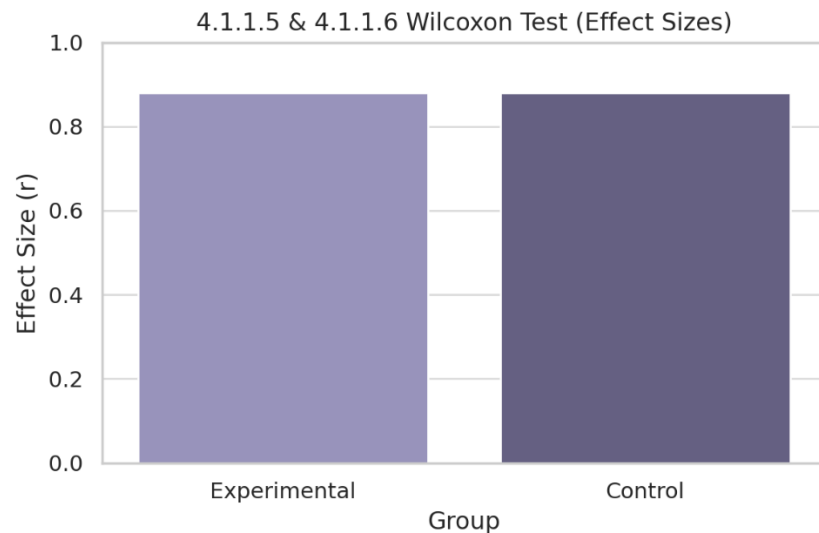


A Mann-Whitney U test was conducted to determine if there were significant differences in pretest scores between the experimental and control groups. The analysis yielded  $U=207.5$ ,  $Z=1.44$ , and a p-value of 0.132, along with an effect size of  $r=0.24$  (indicating a small to medium effect). As the p-value was greater than 0.05, the difference in pretest scores was not statistically significant. This finding is crucial as it suggests that the experimental and control groups were comparable at the outset, allowing for a more confident attribution of any subsequent post-intervention changes to the treatment.

#### 4.1.1.4 Mann-Whitney U Test (Post-test Comparison Between Groups)

To determine the intervention's effectiveness, a Mann-Whitney U test was employed to compare the post-test scores between the experimental and control groups. The results showed no meaningful difference in post-test performance ( $U=138.0$ ,  $Z=-0.76$ ,  $p=0.403$ ), and the effect size was small ( $r=0.13$ ). Given that the p-value was greater than 0.05, it can be concluded that the intervention did not lead to a statistically significant improvement in the experimental group relative to the control group.

#### 4.1.1.5 Wilcoxon Signed-Rank Test (Pretest vs Posttest Within Experimental Group)

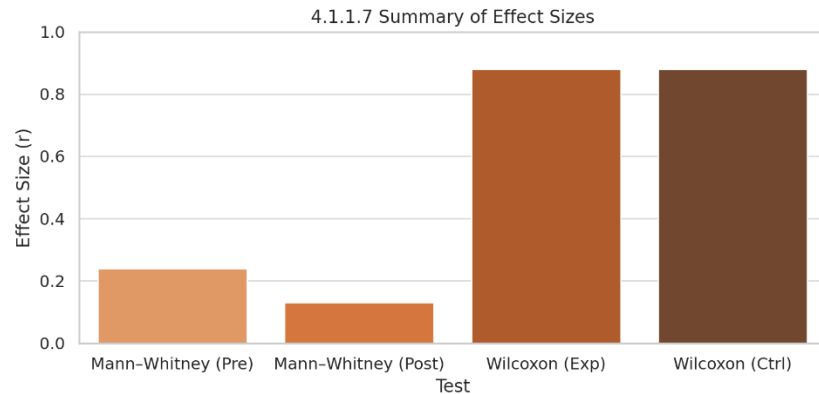


To evaluate within-group improvement, a Wilcoxon Signed-Rank test was performed on the experimental group's pretest and post-test scores. The analysis revealed a highly significant difference ( $W=0.0$ ,  $Z=-3.72$ ,  $p<0.001$ ), with a very large effect size ( $r=0.88$ ). This indicates that the experimental group experienced substantial improvement following the intervention, demonstrating a considerable impact of the treatment on participants' performance within this group.

#### 4.1.1.6 Wilcoxon Signed-Rank Test (Pretest vs Post-test Within Control Group)

Surprisingly, the control group also demonstrated statistically significant improvement from pretest to post-test, as revealed by a Wilcoxon Signed-Rank test. The results were identical to the experimental group:  $W=0.0$ ,  $Z=-3.72$ ,  $p<0.001$ , and  $r=0.88$ . This indicates that even without the intervention, the control group experienced a substantial positive change, with a very large effect size. This unexpected progress suggests the possible influence of external learning factors or uncontrolled variables that contributed to their development during the study period.

#### 4.1.1.7 Effect Size Calculation



An overview of the effect sizes reveals mixed results across the analyses. For the Mann-Whitney U tests comparing the experimental and control groups, the pretest yielded a small to medium effect ( $r=0.24$ ), while the post-test showed only a small effect ( $r=0.13$ ). In stark contrast, both Wilcoxon Signed-Rank tests, examining within-group improvement for the experimental and control groups, produced very large effect sizes ( $r=0.88$ ). These substantial within-group improvements confirm that both groups made significant progress over time. However, the statistical similarity in improvement between the groups limits the exclusive attribution of observed gains to the intervention administered to the experimental group.

According to Cohen (1988), an effect size of 0.8 or greater signifies a large effect, indicating substantial practical significance. This reinforces that despite the lack of statistical differences between the groups, the learning context itself contributed meaningfully to individual student progress. Such interpretation of effect sizes is highly recommended in applied language research (Lakens, 2013; Plonsky & Oswald, 2014).

#### 4.1.2 The Qualitative Finding

In addition to the quantitative findings, this study also incorporates qualitative data to provide a more comprehensive understanding of the research

problem. Qualitative findings allow the researcher to explore participants' experiences, perspectives, and responses in greater depth, offering insights that numbers alone cannot capture. To begin presenting the qualitative results, it is essential to first describe the research participants. This description provides important contextual information about who the participants are, their background, and their relevance to the study, which helps to interpret the subsequent qualitative data more effectively.

#### 4.1.2.1 Research Participants Description

The experimental group included six purposively selected students, chosen to represent a diverse range of writing proficiencies. This selection encompassed both high-achieving and lower-achieving students, determined by their post-test writing scores. The aim was to ensure a balanced and comprehensive understanding of student perceptions and experiences regarding Discord's use as a writing tool.

#### 4.1.2.2 Data Collection Process

To gather comprehensive qualitative data, in-depth, semi-structured interviews were conducted directly with participants. These sessions took place face-to-face in a quiet room within the school premises, ensuring an optimal environment free from distractions. Each interview typically spanned between 20 and 30 minutes. The core of these interviews involved a series of questions aimed at uncovering participants' personal experiences, specific preferences, and any challenges they faced when integrating Discord into their process of writing procedure texts.

Furthermore, the interviews sought to identify any perceived enhancements or improvements in their writing abilities attributable to using the platform. All verbal data collected was subsequently transcribed and anonymized to safeguard participant privacy and uphold ethical research standards. As supported by prominent qualitative research methodologies (Kvale & Brinkmann, 2009; Yin, 2016), this interview format was chosen

for its proven capacity to generate rich, detailed, and authentic participant narratives.

#### 4.1.2.3 Data Analysis Technique

The qualitative data collected from the interviews were analyzed using thematic analysis, a systematic method for identifying, analyzing, and reporting patterns (themes) within data. This process began with an inductive coding approach, allowing themes to emerge naturally and directly from the students' interview responses rather than imposing preconceived categories. Recurring statements were meticulously identified, coded, and then grouped into broader categories, ultimately leading to the identification of key overarching themes. As highlighted by Braun and Clarke (2006), thematic analysis is a particularly suitable methodology for uncovering prevalent patterns within participant perspectives, making it highly effective for qualitative research in educational contexts.

#### 4.1.2.4 Emerging Themes

Students' use of Discord for writing revealed two significant benefits: increased motivation and improved collaboration. The platform's informal and interactive environment led to higher writing motivation among students. Furthermore, Discord's features, such as threads and replies, made it easy for students to engage in structured peer feedback and collaborative revision.

The study also found that Discord positively influenced students' understanding of the writing process. Participants reported that using the platform helped them better grasp the different stages of writing, including planning, drafting, and revising. This suggests that the interactive nature of Discord can serve as a valuable tool for teaching and learning the writing process.

Despite these advantages, some challenges were noted. Several students faced technical and adaptation issues, such as difficulty with the interface or problems caused by unstable internet connections. These issues highlight the importance of addressing technical barriers to ensure all students can benefit from using the platform for their writing activities.

These emerging themes reflect key affective and cognitive factors known to shape writing development in digital settings (Hyland, 2019; Warschauer, 2010).

#### 4.1.2.5 Interview Results by Gender

This section analyzes interview responses by categorizing participants based on gender: female (Participants 1–3) and male (Participants 4–6). This classification helps explore potential gender-related differences in perceptions and experiences using Discord during writing activities. Where applicable, direct quotes are included to enhance qualitative depth.

##### A. Female Participants (P1–P3)

Participant 1 (Highly Receptive): Already familiar with Discord; appreciated voice and screen sharing features. No significant challenges aside from internet issues. Suggested increasing live sessions. Revised score: 75 (Pretest) → 77 (Posttest)

"At first, it felt a bit confusing because I wasn't familiar with the app. But after a while, I got used to it. I liked how it was organized with different channels for different topics. "Sometimes I missed some messages or links when there were too many chats going on. Also, the notifications overwhelmed me."

"I think it's doable and helpful, especially for group work or quick updates. But I still prefer face-to-face classes."

Participant 2 (Highly Receptive): Comfortable due to prior use; liked using text channels for reference. Suggested using grammar bots; no major

issues experienced. Revised score: 76 (Pretest) → 77 (Posttest)

"It was okay, I think. I liked the voice channels, they felt more personal than zoom."

"Mostly technical ones. Like sometimes my mic didn't work or I didn't know how to use some features."

"It works. I can see the benefits. But still, I think offline classes are better for learning."

Participant 3 (Neutral): Initially confused; valued the chat feature for asking questions. Reluctant to join voice channels. Requested tutorial and summary materials.

"It's quite practical in some ways. I can access the materials anytime and join discussions easily."

"Sometimes the group got too quiet, or people just texted without really engaging."

"I would [use Discord again], especially for group assignments. But if I had the choice, I would go for face-to-face classes."

#### B. Male Participants (P4–P6)

Participant 4 (Neutral): First-time Discord user; appreciated file sharing for draft exchange. Felt overwhelmed by too many notifications. Suggested a simpler channel layout. Revised score: 75 (Pretest) → 78 (Posttest)

"Honestly, I enjoyed it. I already use Discord for gaming and chatting with friends."

"Maybe just when others didn't respond quickly. But the app itself? No problem."

"It's more relaxed and familiar for me. It's actually kind of fun."

Participant 5 (Hindered by Device): Used an outdated phone which limited Discord usage; relied more on WhatsApp. Experienced lag and missed

several sessions. Recommended a lighter alternative platform. Revised score: 76 (Pretest) → 78 (Posttest)

"It was convenient and felt less formal than other platforms."

"Maybe when the internet was bad. But overall, I think Discord made communication easier."

"It matches how our generation communicates."

Participant 6 (Hindered by Device): Never used Discord before; only followed announcements. Experienced laptop crashes while using the platform. Requested PDF versions of materials and offline follow-ups. Revised score: 76 (Pretest) → 79 (Posttest)

"I'm already familiar with Discord, so using it for study made everything easier."

"Not really [any problems]. Maybe just some classmates who weren't active."

"It's flexible and fits how I already use the internet."

#### 4.1.2.6 Participants' Score Data (Pretest and Posttest)

All participants showed improvement in writing scores after utilizing Discord for collaborative writing activities.

Participant	Gender	Pretest	Posttest
P1	Female	75	77
P2	Female	76	77
P3	Female	75	77
P4	Male	75	78
P5	Male	76	78
P6	Male	76	79

#### 4.1.2.7 Male Students' Interest in Technology

Male students in this study appeared to show higher initial interest in exploring Discord's technical features—such as file uploads, asynchronous communication, and screen sharing. Despite facing technical

limitations (e.g., older devices or unstable internet), their curiosity and willingness to engage with new technology were evident.

This observation aligns with research suggesting that male students tend to demonstrate more positive attitudes toward technology integration in educational settings. According to Kay et al. (2012), males often report greater confidence and enthusiasm when engaging with digital tools, especially in informal or semi-formal learning environments.

## **4.2 Discussion**

This section delves into the research findings presented previously, interpreting them in light of the study's core research questions. This study explored the impact of using Discord as a platform for teaching procedure text writing, integrating both quantitative data (pre-test and post-test scores) and qualitative insights gleaned from student interviews. These findings are now discussed in relation to existing theoretical frameworks and prior research.

The initial key finding, derived from the Wilcoxon Signed-Rank tests, highlights a statistically significant improvement in writing scores across both the experimental and control groups. Notably, no participant experienced a decrease in their scores. This strongly suggests that both instructional approaches—the Discord-based method and traditional face-to-face instruction—positively contributed to students' overall writing development. The experimental group, specifically, demonstrated a robust within-group improvement ( $Z=-3.72$ ,  $p<0.001$ ,  $r=0.88$ ). This substantial effect size ( $r=0.88$ ) indicates that Discord, as a digital platform, likely offers a highly conducive environment for enhancing writing skills. These results align with Cohen's (1988) interpretation, which considers an effect size above 0.8 to reflect a strong practical impact.

Moreover, these findings underscore the importance of reporting effect sizes in applied language research for meaningful interpretation, a practice advocated by scholars like Lakens (2013) and Plonsky & Oswald (2014).

While significant improvement was observed within both the experimental and control groups, the Mann-Whitney U test comparison revealed no statistically significant difference in post-test performance between them ( $U=138.0$ ,  $p=0.403$ ,  $r=0.13$ ). This crucial finding implies that while students undoubtedly improved their writing, the Discord-based instruction didn't lead to demonstrably superior outcomes compared to the conventional teaching method. This aligns with findings from Qureshi and Aljanadbah (2022), who also reported no significant difference in reading ability between groups using distinct translanguaging strategies. This reinforces the idea that learning outcomes are multifaceted, influenced by a combination of student engagement, instructional design, and the broader learning context.

The qualitative data significantly enriches these statistical findings, offering deeper insights into participant experiences. Interviews consistently revealed that many students perceived Discord as a motivating and flexible environment for their writing activities. Features like threaded replies, asynchronous communication, and its inherently informal tone fostered a relaxed atmosphere that notably boosted student confidence. One student's comment, "It felt more like a group chat than an assignment," vividly illustrates Discord's potential to alleviate performance anxiety, echoing Hyland's (2019) research on how low-pressure, collaborative settings can enhance engagement.

Beyond engagement, students also reported a clearer understanding of the writing process—including planning, drafting, and revising—when using Discord. This suggests the platform may have supported their metacognitive development, aligning with Oakhill et al. (2015), who emphasize metacognitive strategies as essential for comprehension and writing tasks.

Despite these benefits, technical and adaptation issues emerged as recurring challenges, particularly for students with limited device access or no prior Discord experience. For example, Participants 5 and 6 faced significant barriers due to outdated devices and unstable internet connections, which undeniably hindered their full participation. These observations resonate with Warschauer's (2010)

argument regarding the critical role of digital literacy and robust infrastructure in successful technology integration. They also underscore Mabena's (2023) emphasis on the imperative for equitable access to learning tools in digital environments.

Additionally, while generally positive about Discord, some students voiced a need for more structured guidance, especially during the initial stages of implementation. This finding parallels Sugiarto et al. (2023), who highlight that multimodal platforms require clear instructions and scaffolding to maximize their instructional potential.

Interestingly, the control group also exhibited a statistically significant gain ( $Z=-3.72$ ,  $p<0.001$ ,  $r=0.88$ ). This improvement, occurring without the Discord intervention, might be attributed to various factors such as consistent teacher quality, positive classroom dynamics, or even students' incidental exposure to similar writing tasks outside the study's formal intervention. This outcome underscores the multifaceted nature of language acquisition and reinforces the idea that technological interventions should be contextualized rather than viewed as universally superior alternatives.

Overall, while this study supports previous research (e.g., Hyland, 2019; Warschauer, 2010; Mabena, 2023) on the potential of digital tools to enhance student learning, it also affirms the findings of Qureshi and Aljanadbah (2022) by demonstrating that not all technology-based interventions automatically lead to superior outcomes. Therefore, the true effectiveness of Discord as a writing tool extends beyond its inherent features; it is fundamentally intertwined with the quality of its pedagogical integration, equitable access, and user digital fluency.

In conclusion, the combined quantitative and qualitative analyses suggest that Discord can be an effective platform for enhancing students' writing performance and engagement. However, its successful implementation is contingent upon adequately addressing critical factors such as infrastructure, comprehensive training, and ongoing support. Educators are encouraged to leverage Discord's