What makes a video game localisation tester? Exploring the effect of individual factors on error detection
José Ramón Calvo-Ferrer* and Antonio Jesús Muñoz Villena**, Universidad de Alicante

ABSTRACT

Video game quality assurance (QA) is vital to ensuring a high-quality gaming experience across different languages and regions. Competent localisation testers are expected to have advanced language skills, exceptional attention to detail, and familiarity with the game they are testing. However, individual testing performance differs, prompting the question of what constitutes a proficient localisation tester. In this study, students (N = 43) from the degree in Translation and Interpreting from the University of Alicante and the Specialist degree in Translation for the Video Game Industry from the University of Vigo were asked to test a localised video game which contained deliberate errors introduced for the purpose of this study. Subsequently, they completed questionnaires regarding localisation-related attitudes, personality traits, and trait anxiety levels. The results of various bivariate correlation and multivariate linear regression analyses indicate that translation expertise, low emotional stability, and a preference for screen reading were significant predictors of higher error detection rates. Furthermore, other variables such as attitudes towards video games, punctiliousness, and reading speed correlated with error detection in video game localisation testing. These findings may contribute to the enhancement of tester selection and training processes, resulting in superior localisations and more satisfying player experiences across diverse languages and cultures.

KEYWORDS

Localisation, video games, testing, language testing, errors.

1. Introduction

1.1. Video game localisation and testing

As businesses expand globally, they need to make their products or services accessible to customers worldwide. Localisation, the process of adapting such commodities to meet the language, cultural, and other specific requirements of a particular country or region, emerged along with the necessity “to create target market versions (also known as locales) of content in electronic form, notably computer software” (Mangiron & O’Hagan, 2006, p. 11). Localisation has been tackled from different fields — Bernal-Merino (2006), Muñoz Sánchez (2017), and Méndez González and Calvo-Ferrer (2018) have looked into video game localisation, whereas other authors have focused on other areas such as webpage localisation (Jiménez-Crespo, 2013) and software localisation (Esselink, 2000).

Video games are localised to provide similar game experiences in the target market (Fry & Lommel, 2003), allowing more players to enjoy games in their own language. Bernal-Merino (2006, p. 31) defines localisation as “the process of making a product linguistically and culturally […] appropriate to the target country and language”, suggesting that it goes beyond the translation of its language assets (see Chandler & Deming, 2012, p. 4). Furthermore, Bernal-Merino (2014, p. 98) proposes the term TMIES (Translation of Multimedia Interactive Entertainment Software) “to describe the

* ORCID: 0000-0002-6440-1730; email: jr.calvo@ua.es
** ORCID: 0000-0003-4205-9062; email: munnozvillena@ua.es

https://doi.org/10.26034/cm.jostrans.2024.5978
© The Authors
purely linguistic and cultural issues involved in the translation of video games”. Video game localisation differs from other types of localisation due to the interactive nature of video games, which presents unique challenges for localisation professionals (Dietz, 2007). Méndez González and Calvo-Ferrer (2018) use the term ‘paratranslation’ to include all texts likely to be translated in connection with a video game, other than ingame text, such as manuals, cheat books, legal texts, merchandising, etc. All in all, as Costales (2012, p. 389) suggests, adapting a game into a different target market “goes beyond the simple language transfer from one language into another”. Video game localisation has been discussed from multiple angles — O’Hagan and Mangiron (2013) yield a comprehensive viewpoint on its evolving nature, approaching localisation from the perspective of cultural studies, game reception, and accessibility, among others. Similarly, authors analyse cultural differences between locales (Consalvo, 2006) and how games need to be effectively localised (Méndez González, 2015), whereas Mangiron (2012) discusses the integration of features in localised games that cater for the needs of users with disabilities.

Localisation is typically broken down into three stages: familiarisation, translation, and testing. Quality assurance (QA) testing is an essential aspect of video game localisation, as localised games may include language-related errors that affect the overall gaming experience. In line with this, Fernández-Costales (2016) investigates the influence of localisation on game reception, exploring users’ perceptions of localised games and providing insights into how cultural adaptations influence player satisfaction and engagement. Deckert and Hejduk (2022) state that spelling errors in video game localisation detract from users’ quality estimations. QA testing in video game localisation traditionally encompasses functional and linguistic testing (Mangiron & O’Hagan, 2006). Functional testing ensures that the game operates correctly, while linguistic testing involves reviewing the game for spelling, grammar, and punctuation errors, ensuring consistency and coherence in the translations, and checking that the translations make sense in context, are culturally appropriate, and the game is easy to understand and play. For the purposes of this study, we propose an additional category in connection with language errors, termed ‘visual errors.’ These include issues such as character truncation, font issues, string length, and visual placement. Although intrinsically linked to linguistic aspects, visual testing focuses on the graphical representation of text rather than language use per se.

Video game localisation testers require specific personal qualities and job-related skills, including high proficiency in the target language, a solid understanding of the language’s conventions and rules, a deep knowledge of the game and its mechanics, and an awareness of how the language relates to the player’s experience. Furthermore, they must be detail-oriented, with a great capacity for attention to detail and problem-solving skills. Strong communication skills are also necessary, as the tester must report problems to the development team and provide solutions. Additionally, cultural sensitivity is critical, as testers must be able to recognise cultural references, imagery, and themes that may or may not be appropriate for the target audience (Méndez González, 2013).

All in all, localisation is an important process in the gaming industry, the effectiveness of which lies greatly on the capabilities of the testers, who are responsible for ensuring its overall quality. However, significant differences arise in error detection activities, which may stem from individual factors such as personality, attitudes towards games
or translation expertise (Calvo-Ferrer, 2022). As a result, identifying the traits that predict success in these roles is vital for maintaining high standards in localisation practice. Recognising the role of these testers and their diverse skillset needed, our study seeks to fill a significant gap in the current research by examining how these various personal qualities and professional skills correlate with successful localisation outcomes. Given the limited research in this area, it adopts an exploratory approach to investigate the impact of individual factors on video game localisation testing performance. This approach allows us to explore a range of factors, including personality traits and trait anxiety, to better understand what contributes to effective localisation testing in video games.

1.2. Personality traits and translation competence

Quality assurance testing, a crucial aspect of the localisation workflow, is a complex cognitive activity that involves a range of individual factors, including personality traits and emotional intelligence. In broader terms, several studies have explored the effect of individual factors on translation quality and performance. For instance, Gojian (2014) argues that anxiety may negatively affect learning and performance in translation tasks, whereas Cifuentes-Férez and Fenollar-Cortés (2017) indicate that translation students who inhibit their emotional state and response may produce higher-quality translations. Hubscher-Davidson (2009) reported significant correlations between personality traits and translation quality. In a similar vein, Hubscher-Davidson (2016) found that emotional intelligence was positively associated with job satisfaction, career success, and literary translation experience. These findings suggest that individual differences in personality and emotional intelligence may influence various aspects of translation performance, although the nature and extent of these effects remain to be fully understood and require further investigation. In line with this, Bolaños Medina (2014) emphasises the importance of examining individual skillsets in Translation Studies, especially in educational and employment contexts, as it may enhance translator training and facilitate better matching between job candidates and the required profiles.

Personality traits are inherent qualities that shape a person’s behaviour, thoughts, and feelings. One popular model for understanding personality is the Five Factor Model (McCrae & John, 1992), which includes five broad dimensions: openness, conscientiousness, extraversion, agreeableness, and neuroticism. Openness is a personality trait that reflects a person’s curiosity and willingness to embrace new and unconventional ideas. As McCrae and Greenberg (2014) suggest, people who score high in openness tend to be imaginative and creative, while those who score low are more practical and conventional. Conscientiousness, on the other hand, refers to the degree to which a person is organised, efficient, and goal-oriented. Highly conscientious individuals are typically diligent, disciplined, and reliable, as opposed to those who are more spontaneous and flexible (Kern, 2020). Extraversion reflects a person’s sociability and outgoing nature. Walker (2020) posits that people who score high in extraversion tend to be talkative, energetic, and assertive, while those who score low are more reserved and introverted. Agreeableness is a personality trait that reflects a person’s tendency to be cooperative, compassionate, and empathetic towards others. Agreeable individuals are typically kind, sympathetic, and accommodating, while those who score low may be more critical and sceptical of others’ intentions (Tobin & Graziano, 2020). Finally, Zhang (2020) explains that
neuroticism reflects a person’s tendency to experience negative emotions such as anxiety, fear, and sadness. Highly neurotic individuals are more prone to stress and may have difficulty coping with challenging situations, while those who score low tend to be more emotionally stable and resilient.

Similarly, anxiety is a common emotional response to stressful or threatening situations (Sagardoy Muniesa & Romeo Miguel, 2013). It is important to differentiate between state anxiety and trait anxiety when examining its impact. State anxiety is a temporary emotional condition that varies in intensity and over time, often fluctuating with changes in the perceived threat level. It is situationally induced and can be a direct response to an immediate stressor, such as the anxiety a player might feel during a particularly challenging level in a game. In contrast, trait anxiety refers to a more stable aspect of personality, indicating a general tendency to respond with anxiety across a variety of situations. This is a more enduring quality, which can predispose individuals to perceive a wider range of circumstances as threatening, and it can result in more consistently heightened levels of cognitive and somatic anxiety. Exploring trait anxiety aligns with the need for a comprehensive personality profile that may predict performance in localisation tasks. Unlike state anxiety, which might fluctuate from day to day, trait anxiety offers a more stable predictor of an individual's approach to work, their stress management strategies, and their overall job performance.

Cognitive anxiety and somatic anxiety are two distinct types of anxiety that can affect individuals in different ways. Cognitive anxiety refers to the psychological symptoms of anxiety, such as excessive worry, negative thoughts, and fear of failure. Somatic anxiety, on the other hand, refers to the physical symptoms of anxiety, such as increased heart rate, muscle tension, sweating, and shaking. Cognitive anxiety can be a major obstacle to success in many areas of life, including academic, athletic, and professional performance. Excessive worry and self-doubt can lead to negative thinking patterns that can hinder a person’s ability to concentrate, make decisions, and perform well under pressure. Somatic anxiety can also have a significant impact on a person’s performance, particularly in high-pressure situations. Physical symptoms such as increased heart rate, sweating, and shaking can interfere with fine motor skills and cognitive functioning, leading to poor performance and increased anxiety. However, somatic anxiety can also be used as a tool for enhancing performance. For example, athletes who experience a moderate level of somatic anxiety before a competition may perform better than those who are completely relaxed or extremely anxious. This is known as the ‘inverted-U hypothesis,’ which suggests that performance increases with arousal up to a certain point, beyond which it begins to decline (Arent & Landers, 2003).

It is our belief that, along with other factors, both personality and anxiety have a significant impact on localisation testing performance, as it can affect a tester’s cognitive and somatic abilities. High levels of trait anxiety can impair concentration, decision-making skills, and attention, all of which are crucial for video game localisation testing tasks. Understanding how these factors may influence individual performance can help profiling the proficient video game localisation tester and contribute to filling the research gap regarding the influence of individual factors on video game localisation testing.
1.3. Research questions and hypotheses

Based on the topics discussed in the introduction, the following research questions can be formulated:

1. To what extent do individual differences in personality traits, as described by the Five Factor Model, impact error detection in the context of video game localisation testing?
2. Does trait anxiety influence error detection during video game localisation testing?
3. What individual traits are significative of error detection in video game localisation testing? In other words, what makes a localisation tester?

Given the background information and research questions, the following hypotheses can be formulated for this study:

H1. There is a significant effect of personality, as described by the Five Factor Model, on error detection in the context of video game localisation. Specifically, traits such as conscientiousness, openness, and emotional stability (low neuroticism) may positively influence performance.
H2. Higher levels of cognitive anxiety may negatively impact performance during video game localisation testing, as excessive worry and self-doubt can hinder concentration, decision-making, and working memory. In contrast, moderate levels of somatic anxiety may enhance testing performance by increasing arousal and focus.
H3. Positive attitudes (such as testers’ interest in and positive perceptions of video games) and skills related to video game localisation (such as reading speed or translation expertise) may positively influence the identification of errors in video game localisation testing.

1.4. The current study

In the present study, the hypotheses were investigated through an experimental approach involving students from diverse Translation and Interpreting programmes. Participants engaged in a video game testing task designed to assess their ability to identify both linguistic and visual errors. Furthermore, participants completed a questionnaire addressing factors such as opinions on languages and translation, reading attitudes, language punctiliousness, perceived language skills, preferences in employing print or electronic materials, and attitudes towards video games. The findings of this study are anticipated to impact translator training in video game localisation testing and to provide insights into the characteristics that may contribute to proficiency in this field. This understanding is valuable from both the industry’s perspective in identifying suitable candidates for the workforce and the students’ perspective in considering such a career path. Additionally, this study aims to fill the research gap regarding the influence of individual factors, such as personality traits and anxiety, on video game localisation testing.
2. Materials and method

2.1. Sample

The study’s participants comprised students from different degrees in Translation and Interpreting ($N = 43$). Specifically, twenty-six of them ($n = 26$) were undergraduate students enrolled in the four-year degree program in Translation and Interpreting at the University of Alicante, encompassing students from various academic years within the programme. Conversely, seventeen ($n = 17$) were postgraduate students pursuing the one-year Specialist degree in Translation for the Video Game Industry at the University of Vigo. Among the participants, 31 identified as female (72.1%) and 12 as male (27.9%), with ages ranging from 18 to 32 ($M = 20.98$, $SD = 3.356$).

2.2. Materials

To assess QA testing performance, the educational video game *Subtitle Legends* (Calvo-Ferrer et al., 2020) was deliberately altered by incorporating the most prevalent linguistic and visual errors (see Vázquez-Rodríguez, 2018, p. 240). The XML files of the game were adjusted to include a range of errors, which were categorised into nine distinct types. Linguistic errors encompassed four types: punctuation inaccuracies, spelling mistakes, incorrect or absent diacritics, and duplicate text. Visual errors included five types: untranslated text, appearing codes that should not be visible, appearing variables that should not be visible, text overflow, and blank spaces. A total of 27 errors, with 3 instances for each of the 9 error types were introduced in the Spanish version of the game, which participants were instructed to detect (although received no indication as to the number or type of errors appearing in the game), enabling the researchers to evaluate their performance in video game localisation testing. Figure 1 provides examples of these errors.

2.3. Instruments

Three instruments were used in this study to gather data on the participants’ individual factors and their impact on video game localisation testing performance: the Big Five Questionnaire, the Competitive State Anxiety Inventory-2D (CTAI-2D), and a questionnaire developed from consultation with experts in video game localisation. The Big Five Questionnaire (Caprara et al., 1993) was employed to assess personality

![Figure 1. Examples of errors (presence of variables and blanks) introduced in the game *Subtitle Legends* (Calvo-Ferrer et al., 2020)](image-url)
traits. This questionnaire comprises 132 items and utilises a 1 to 5 Likert scale to measure Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism. On the other hand, the CTAI-2D questionnaire (Muñoz-Villena et al., 2022), consisting of 27 items, was utilised to assess trait anxiety, specifically evaluating the intensity and direction of somatic and cognitive anxiety, as well as self-confidence. In addition to these established instruments, an expert group questionnaire was developed based on the opinions of professionals in the video game localisation industry. A total of 15 professionals — including video game developers, localisers, and testers — and academics specialising in video game localisation were asked to suggest characteristics that, in their opinion, define a proficient video game localisation tester. Their responses were processed and reformulated to create a questionnaire consisting of 35 items designed to be rated using a Likert scale ranging from 1 (strongly disagree) to 7 (strongly agree) and assessing hypothesised high-performance traits related to attitudes towards translation and languages, reading, video games, punctiliousness, self-assessed language skills, and preferences for using screen or paper materials. A reading speed test was also provided via a link at the end of the questionnaire, further contributing to the comprehensive assessment of the participants' individual factors. The questionnaire development process followed the guidelines outlined by Martín Arribas (2004), particularly regarding the evaluation of the metric properties of the scale. As the Big Five Questionnaire (Caprara et al., 1993) and the CTAI-2D (Muñoz-Villena et al., 2022) are established instruments we maintained their 1 to 5 Likert scale to preserve the integrity of their validated constructs. Conversely, for the expert group questionnaire, specifically developed for our study, we chose a 1 to 7 scale based on research suggesting that a seven-point scale can provide greater sensitivity and nuanced responses (Cai et al., 2016). Participants in the study were asked to complete the questionnaires through a dedicated web page, as explained in the next section. A detailed overview of these instruments, including their purpose, is provided in Table 1.

### 2.4. Procedure

Participants took part in a combination of in-class and out-of-class activities for the study. In the initial in-class phase, which lasted two hours, students were asked to complete the aforementioned expert questionnaire and a reading speed test via the online link at bit.ly/SLTESTING. After completing these tasks, they were directed to bit.ly/SLTESTINGFORM to access the game and a reporting spreadsheet for documenting any errors encountered during gameplay. Participants were instructed to

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big Five Questionnaire (Caprara et al., 1993)</td>
<td>To assess personality traits including Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism.</td>
</tr>
<tr>
<td>Competitive State Anxiety Inventory-2D (CTAI-2D) (Muñoz-Villena et al., 2022)</td>
<td>To evaluate the intensity and direction of somatic and cognitive anxiety, as well as self-confidence.</td>
</tr>
<tr>
<td>Expert Group Questionnaire</td>
<td>To assess high-performance traits related to attitudes towards translation and languages, reading, video games, punctiliousness, self-assessed language skills, preferences for screen or paper materials.</td>
</tr>
<tr>
<td>Reading Speed Test</td>
<td>To measure reading speed as part of the comprehensive assessment of participants' individual factors.</td>
</tr>
</tbody>
</table>

Table 1. Overview of instruments used
play through the game and report any identified errors, without any indication regarding how many of such instances had been intentionally introduced, to avoid potential bias. Upon completion of the testing, students submitted the provided spreadsheet file to one of the researchers and were dismissed from the in-class session. Subsequently, students received an email inviting them to participate in the completion of the Big Five Questionnaire and the Competitive State Anxiety Inventory-2D (CTAI-2D), with a one-week deadline. However, this approach resulted in significant experimental mortality, as only 43 out of 100 participants submitted the completed questionnaires. Administering these questionnaires outside the classroom setting was deemed necessary due to time limitations that would have hindered the completion of all tasks during the in-class session. Participants were provided with a brief explanation of the study’s purpose and procedures prior to their involvement, and informed consent was obtained from each participant before they engaged in the study. The study was conducted in accordance with the Declaration of Helsinki and following the regulations in force at the University of Alicante (Spain) for studies involving humans: https://web.ua.es/en/vr-investigacio/comite-etica/regulations.html.

2.5. Data analysis

The data collected allowed for the creation of several variables, including opinions on translation, attitudes towards reading and video games, punctiliousness, perceived language skills, screen vs paper preferences, reading speed, translation expertise, different personality traits, and cognitive and somatic anxiety. Additionally, the number of errors identified by the students was computed, with, as discussed earlier, a distinction made between linguistic and visual errors. Following this, various dependent variables were computed, such as the number of visual errors spotted, the number of linguistic errors spotted, and the total number of errors spotted. Table 2 provides an example of each error type.

To analyse the data, bivariate correlations were first computed to explore relationships between variables. Several statistically significant correlations between the ability to identify errors and other measured variables were found. Multivariate linear regressions were then conducted to investigate the predictive strength of these variables. All analyses were performed using the SPSS 22.0 statistical software, with

<table>
<thead>
<tr>
<th>Error type</th>
<th>Correct text</th>
<th>In-game text</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Linguistic errors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Punctuation</td>
<td>¡Genial! Lo has hecho muy bien.</td>
<td>¡Genial!. Lo has hecho muy bien.</td>
</tr>
<tr>
<td>Spelling</td>
<td>Vamos a ver los distintos tipos de errores.</td>
<td>Vamos haber los distintos tipos de errores.</td>
</tr>
<tr>
<td>Diacritics</td>
<td>¿A quién vas a llamar?</td>
<td>¿A quien vas a llamar?</td>
</tr>
<tr>
<td>Duplicates</td>
<td>La verdad es que no parece que haya ningún error ortográfico.</td>
<td>La verdad es que no parece que haya ningún error ortográfico.</td>
</tr>
<tr>
<td><strong>Visual errors</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Untranslated</td>
<td>No hay ofertas disponibles.</td>
<td>No offers available.</td>
</tr>
<tr>
<td>Codes</td>
<td>Según el diccionario bilingüe debería ser...</td>
<td>Según el diccionario bilingüe debería ser...</td>
</tr>
<tr>
<td>Variables</td>
<td>Cuando hayas completado al menos 5 trabajos podrás ver tu nivel de conocimientos.</td>
<td>Cuando hayas completado al menos @number@ trabajos podrás ver tu nivel de conocimientos.</td>
</tr>
<tr>
<td>Overflow</td>
<td>B1</td>
<td>Nivel de conocimientos B1</td>
</tr>
<tr>
<td>Blank</td>
<td>Ramón</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 2. Error types and examples
a significance level of 0.05. Key assumptions of the linear regression model, including normality, linearity, homoscedasticity, and absence of multicollinearity and autocorrelation, were checked and confirmed using the approach recommended by Vilà Baños, Torrado Fonseca, and Reguant Álvarez (2019).

3. Results

3.1. Errors identified

Participants identified an average of 30.70% (M = 4.6047, SD = 3.05614) of the 15 visual errors deliberately incorporated into the game Subtitle Legends. Specifically, they detected 25.58% of non-translated text errors (M = 0.7674, SD = 0.94711), 39.53% of appearing codes errors (M = 1.1860, SD = 0.85233), 18.60% of text overflow errors (M = 0.5581, SD = 0.66556), 12.40% of blank spaces errors (M = 0.3721, SD = 0.61811), and 57.36% of appearing variables errors (M = 1.7209, SD = 1.27850). Similarly, participants identified an average of 26.74% of the 12 linguistic errors (M = 3.2093, SD = 1.72597) in the game. Specifically, they detected an average of 23.26% of punctuation inaccuracies (M = 0.6977, SD = 0.83195), 15.5% of incorrect or absent diacritics (M = 0.4651, SD = 0.63053), 58.91% of spelling mistakes (M = 1.7674, SD = 0.99612), and 9.3% of duplicate text errors (M = 0.2791, SD = 0.45385). All in all, the participants detected an average of 28.94% (M = 7.8140, SD = 3.75598) out of the total 27 errors introduced by the researchers.

3.2. Relations between the variables explored

The expert questionnaire, the Big Five Questionnaire, and the CTAI-2D questionnaire were employed in the present study to explore a range of variables that may impact error detection in video game localisation testing. These instruments exhibited satisfactory to excellent reliability, with Cronbach’s Alpha coefficients of 0.744, 0.908, and 0.804, respectively, indicating consistency in measuring the same construct and supporting their use in research contexts. As previously discussed, the variables investigated included opinions on translation, attitudes towards reading and video games, punctiliousness, perceived language skills, screen vs paper preferences, and reading speed (as determined by the expert questionnaire). Additionally, we considered translation expertise, which for the purpose of this study refers to the level of academic training in translation and interpreting, as indicated by the participants’ year of study, rather than professional experience. This was complemented by an assessment of personality traits (as determined by the Big Five Questionnaire), and cognitive and somatic anxiety (as determined by the CTAI-2D questionnaire).

To identify potential predictors of error detection in video game localisation testing, a correlation analysis was conducted. The analysis explored the associations between the computed variables and the number of visual, linguistic, and total errors detected, using Spearman’s rho correlation coefficients. For visual errors, two statistically significant correlations emerged. First, a positive correlation between somatic anxiety and visual errors detected (ρs = 0.309, p = 0.044) indicated that higher levels of somatic anxiety were associated with detecting a greater number of visual errors. Second, a negative correlation between neuroticism and visual error detection (ρs = -0.323, p = 0.035) suggested that individuals with higher emotional stability tended to spot fewer visual errors. Similarly, a significant negative correlation between
neuroticism and the number of total errors detected ($\rho_s = -0.369, p = 0.015$) indicated that participants with higher emotional stability spotted fewer errors in general. However, no statistically significant correlations were observed for linguistic errors.

The analysis revealed further significant correlations. These included a positive correlation between the year of study and visual errors ($\rho_s = 0.459, p = 0.002$), suggesting that as translation expertise increased, so did the number of visual errors detected. Additionally, a positive correlation between attitudes towards video games and visual errors ($\rho_s = 0.313, p = 0.041$) indicated that those with a more favourable view of video games tended to detect more visual errors. Furthermore, a series of positive correlations between punctiliousness and visual errors ($\rho_s$ ranging from 0.425 to 0.531, $p < 0.01$) suggested that individuals who pay greater attention to punctuation and grammar detected more visual errors. Positive correlations between screen vs paper preferences and visual errors ($\rho_s = 0.347, p = 0.023; \rho_s = 0.361, p = 0.017$) implied that participants who prefer reading from paper or proofreading translations on screen without printing them detected more visual errors. Lastly, a negative correlation between reading speed and visual errors ($\rho_s = -0.356, p = 0.019$) revealed that those with a higher reading speed detected fewer visual errors, which suggests that faster readers may overlook certain visual errors due to their rapid processing of information, or that efficient readers have developed strategies or skills that enable them to focus on the content while not being as susceptible to other types of errors having to do with format.

Regarding the number of total errors detected, similar significant correlations were also observed. A positive correlation between the year of study and total errors was also identified ($\rho_s = 0.323, p = 0.035$), suggesting that an increase in translation expertise is associated with a higher number of total errors detected. Moreover, a positive correlation between attitudes towards video games and total errors ($\rho_s = 0.330, p = 0.031$) suggested that individuals with a more positive outlook on video games were inclined to identify a greater number of total errors. As detailed in Table 4, these attitudes encompass aspects such as enjoyment of playing video games, self-identification as a regular player, and the extent of lifelong engagement with video games, as measured by specific items in the expert questionnaire. Also, a series of positive correlations between punctiliousness and total errors ($\rho_s$ ranging from 0.413 to 0.399, $p < 0.01$) suggested that individuals who pay greater attention to punctuation and grammar detected more total errors. Finally, a positive correlation between screen
vs paper preferences and total errors ($\rho_s = 0.357$, $p = 0.019$) implied that participants who proofread their translations on screen detected more total errors.

3.3. Effect of computed variables on the detection of visual errors

The variables that yielded significant correlations were included in the regression analyses to identify their contributions to the prediction of error detection in video game localisation testing. To investigate which of these variables potentially influence the detection of visual errors, a stepwise regression analysis comprising three models was conducted.

The first model solely integrated the predictor ‘Translation expertise’ ($\beta = 0.464$), accounting for 21.6% of the variance in visual error detection ($R^2 = 0.216$). This result implies that the participants’ year of study exerted a substantial impact on the dependent variable ($F[1,40] = 10.991$, $p = 0.002$) — participants are more likely to identify visual errors in the video game localisation testing as they advance in their Translation and Interpreting studies. The second model built upon the foundational model by incorporating ‘Screen vs Paper preferences’ as an additional predictor. The inclusion of this variable yielded an increased $R^2$ value of 0.305, signifying that both ‘Translation expertise’ ($\beta = 0.486$) and ‘Screen vs Paper preferences’ ($\beta = -0.300$) collectively explained 30.5% of the variance in visual error detection. The enhancement of the model’s explanatory power emphasises the considerable effect of the participants’ preference for screen reading on their visual error detection capabilities ($F[2,39] = 8.554$, $p = 0.001$). Finally, the third model further refined the analysis by integrating ‘Emotional volatility’ as a variable. This final model exhibited an

<table>
<thead>
<tr>
<th>Item</th>
<th>$M$</th>
<th>$SD$</th>
</tr>
</thead>
<tbody>
<tr>
<td>I like languages.</td>
<td>6.651</td>
<td>0.650</td>
</tr>
<tr>
<td>I like translating.</td>
<td>6.256</td>
<td>0.978</td>
</tr>
<tr>
<td>I see myself working as a translator in the future.</td>
<td>5.977</td>
<td>1.422</td>
</tr>
<tr>
<td>Reading is fun.</td>
<td>6.163</td>
<td>1.067</td>
</tr>
<tr>
<td>I like reading.</td>
<td>6.140</td>
<td>1.207</td>
</tr>
<tr>
<td>I see myself as an eager reader.</td>
<td>4.907</td>
<td>1.688</td>
</tr>
<tr>
<td>I have been reading out of pleasure for most part of my life.</td>
<td>5.070</td>
<td>1.981</td>
</tr>
<tr>
<td>Video games are fun.</td>
<td>6.140</td>
<td>1.441</td>
</tr>
<tr>
<td>I like playing video games.</td>
<td>5.744</td>
<td>1.677</td>
</tr>
<tr>
<td>I see myself as a regular video game player.</td>
<td>4.442</td>
<td>2.164</td>
</tr>
<tr>
<td>I have played video games for most part of my life.</td>
<td>4.930</td>
<td>2.063</td>
</tr>
<tr>
<td>I pay attention to punctuation when I am reading.</td>
<td>5.860</td>
<td>1.457</td>
</tr>
<tr>
<td>I pay attention to punctuation when I am writing.</td>
<td>6.419</td>
<td>0.763</td>
</tr>
<tr>
<td>I could not go out with a person who makes lots of mistakes in writing.</td>
<td>4.279</td>
<td>2.175</td>
</tr>
<tr>
<td>I cannot send a message or an email which contains incorrect grammar.</td>
<td>5.186</td>
<td>1.803</td>
</tr>
<tr>
<td>I see myself as a ‘grammar nazi’.</td>
<td>3.767</td>
<td>2.091</td>
</tr>
<tr>
<td>I have excellent first language skills.</td>
<td>5.744</td>
<td>0.978</td>
</tr>
<tr>
<td>I have excellent second language skills.</td>
<td>5.186</td>
<td>0.880</td>
</tr>
<tr>
<td>I am very good at languages.</td>
<td>5.419</td>
<td>0.879</td>
</tr>
<tr>
<td>I do not like working with in teams because I typically have to correct people’s mistakes.</td>
<td>3.558</td>
<td>2.085</td>
</tr>
<tr>
<td>In general, people have worse language skills than me.</td>
<td>3.698</td>
<td>1.567</td>
</tr>
<tr>
<td>I usually print my notes to study from them.</td>
<td>4.279</td>
<td>2.453</td>
</tr>
<tr>
<td>I feel it is more comfortable to read from paper than from a screen.</td>
<td>5.488</td>
<td>1.737</td>
</tr>
<tr>
<td>I usually proofread my translations on screen, without printing them.</td>
<td>5.651</td>
<td>1.646</td>
</tr>
</tbody>
</table>

Table 4. Descriptive statistics for the items in the Experts’ questionnaire

Note: $N = 43$
\( R^2 \) of 0.386, indicating that the combined influence of ‘Translation expertise’ (\( \beta = 0.498 \)), ‘Screen vs Paper preferences’ (\( \beta = -0.246 \)), and ‘Emotional volatility’ (\( \beta = -0.291 \)) accounted for 38.6\% of the variance in visual error detection (\( F[3,38] = 7.976, p = 0.000 \)). The increase in the \( R^2 \) value accentuates the effect of emotional volatility — a sub-trait of the variable ‘Neuroticism’ — in detecting visual errors.

3.4. Effect of computed variables on the detection of linguistic errors

To ascertain whether any of the variables that yielded significant correlations could predict the detection of linguistic errors in video game localisation testing, a second regression analysis was performed, which constructed a single model incorporating the variable ‘Translation expertise.’ This model indicated that this variable accounted for a mere 2.6\% of the variance in linguistic error detection (\( R^2 = 0.026 \)), suggesting a relatively limited influence on the dependent variable. Furthermore, the ANOVA results revealed that the relationship between ‘Translation expertise’ and linguistic error detection lacked statistical significance (\( F[1,40] = 1.055, p = 0.311 \)).

Likewise, none of the excluded variables exhibited a significant association with linguistic error detection, as demonstrated by their elevated p-values and diminished partial correlations. Some excluded variables encompassed ‘Attitudes towards video games’ (\( \beta = 0.075 \)), ‘Screen vs Paper preferences’ (\( \beta = 0.158 \)), ‘Punctiliousness’ (\( \beta = 0.296 \)), ‘Reading speed’ (\( \beta = -0.198 \)), and all dimensions of neuroticism and somatic anxiety. In conclusion, the regression analysis revealed a tenuous and statistically non-significant relationship between the participants’ year of study and the detection of linguistic errors, while none of the other excluded variables demonstrated a substantial influence on the detection of this type of error.

3.5. Effect of computed variables on the detection of the total number of errors

To investigate whether any of the variables that yielded significant correlations could predict the detection of errors in video game localisation testing (i.e., linguistic and visual errors combined), a third regression analysis was performed. In this case, two models were constructed to evaluate the contributions of different predictors.

The first model incorporated the variable ‘Translation expertise,’ which accounted for 9.2\% of the variance in total error detection (\( R^2 = 0.092 \)). The ANOVA results signified that the relationship between ‘Translation expertise’ (\( \beta = 0.304 \)) and the total number of errors detected was statistically significant (\( F = 4.059, p = 0.051 \)). The second model introduced the variable ‘Emotional volatility’ into the analysis using a stepwise method. This inclusion enhanced the explained variance, with the model accounting for 20.8\% of total error detection (\( R^2 = 0.208 \)). The ANOVA results validated the model’s statistical significance (\( F = 5.108, p = 0.011 \)), denoting the contributions of both ‘Translation expertise’ (\( \beta = 0.322, p = 0.030 \)) and ‘Emotional volatility’ (\( \beta = -0.340, p = 0.022 \)) in predicting the detection of all errors included in the video game. An examination of excluded variables revealed that none exhibited a significant association with total error detection and were not included in the final model due to their failure to meet the stepwise criteria for inclusion (probability of F-to-enter <= 0.05 and probability of F-to-remove >= 0.10). Among the excluded variables were ‘Attitudes towards video games’ (\( \beta = 0.193 \)), ‘Screen vs Paper preferences’ (\( \beta = -0.179 \)), ‘Punctiliousness’ (\( \beta = 0.227 \)), ‘Reading speed’ (\( \beta = 0.104 \)), and other dimensions of neuroticism and somatic anxiety.
In summary, the regression analysis disclosed that the combination of ‘Translation expertise’ and ‘Emotional volatility’ significantly predicts total errors, explaining 20.8% of the variance. In contrast to visual errors, the variable ‘Screen vs Paper preferences’ is not significant in predicting the total number of errors detected, implying that the preference for reading on screen does not exert a substantial influence on the detection of errors when considering both visual and linguistic errors combined.

4. Discussion

The present study aimed to investigate the impact of individual traits on student video game localisation testing proficiency, focusing on the detection of visual and linguistic errors. A range of variables were examined, including opinions on translation, attitudes towards reading and video games, punctiliousness, perceived language skills, screen vs paper preferences, reading speed, translation expertise, personality traits, and cognitive and somatic anxiety. The results provided insight into the relationships between these variables and error detection capabilities, as well as the potential predictors of proficiency in video game localisation testing.

The first hypothesis, which anticipated a significant effect of personality traits on error detection in video game localisation testing, was partially corroborated by the study’s findings. Although a significant relationship emerged, it manifested in a manner that contradicted the initial expectations — emotional stability, as measured by the neuroticism dimension of the Big Five personality traits, was a negative predictor of the detection of both visual errors and the total number of errors. This indicates that individuals with higher emotional stability may be less adept at detecting errors, possibly due to lower levels of stress leading to reduced vigilance during the testing process. Higher emotional stability may be linked to a more relaxed approach to task processing (Meinhardt & Pekrun, 2003), causing emotionally stable individuals to fail to perceive the task as highly challenging or crucial, potentially leading to less effort or motivation to identify errors. Similarly, individuals with lower levels of confidence may display an overall heightened strain and sensitivity to discrepancies, which could contribute to better error detection capabilities. In line with this, previous research suggests that higher emotional stability leads to overconfidence (Logg et al., 2018), which may even relate to inadequate adaptation to task demands (de Craen et al., 2007), resulting in a reduced perception of errors or an inclination to assume that detected errors are insignificant.

The second hypothesis, which postulated that cognitive anxiety might negatively influence performance while moderate levels of somatic anxiety could enhance video game localisation testing by increasing alertness and attentiveness, could not be supported. Neither cognitive nor somatic anxiety showed any significant effect on performance, although the results revealed a significant positive correlation between elevated levels of somatic anxiety and the detection of visual errors. This may be attributable to individuals with higher somatic anxiety experiencing increased alertness due to physical symptoms such as elevated heart rate, rapid breathing, and muscle tension (Belem Da Silva et al., 2014). As components of the body’s natural stress response mechanism, these symptoms can heighten one’s senses and improve environmental awareness (Dierolf et al., 2018), thereby augmenting the capacity to identify visual errors. Furthermore, higher levels of somatic anxiety may foster a more
vigilant and detail-oriented approach to tasks. In accordance with this notion, anxious individuals are often more sensitive to potential threats, which may lead to enhanced attention to detail and a propensity to scrutinise the visual aspects of a game more closely. Additionally, those with elevated somatic anxiety may experience an intensified need to perform well in tasks due to fear of negative consequences or judgments upon failure. This heightened motivation could prompt them to invest more effort and time in identifying visual errors, ultimately resulting in better error detection.

The third hypothesis, which anticipated that positive attitudes and greater skills related to video game localisation have a positive impact on the identification of errors, could only be partially supported. Our findings revealed a positive correlation between the year of study and the detection of visual errors and total errors. This suggests that as students progress in their Translation and Interpreting studies, they become more adept at identifying errors in video game localisation. This result is in line with previous research, which has shown that experience and expertise in translation can improve error detection skills (Redelinghuys & Kruger, 2015). It must be noted, however, that translation expertise is defined in the context of this study based on academic training — the assumption that translation training inherently improves proofreading abilities needs careful consideration. The data show that screen versus paper preferences emerged as a significant predictor of error detection, with participants who exhibit a preference for on-screen reading detecting a greater number of visual errors and total errors. This finding implies that familiarity with and preference for on-screen reading may enhance error detection capabilities in a digital context, such as video game localisation testing, thus supporting previous research which identifies that proofreading translations on screen may lead to greater error detection, particularly of the visual kind (Calvo-Ferrer, 2022). Several factors may underlie this phenomenon. First, individuals who are accustomed to reading on screens might have developed an enhanced ability to process visual information in a digital format, such as video games. This proficiency may be particularly pronounced in those who are regular gamers. Second, on-screen readers may be more comfortable navigating digital interfaces and adept at identifying issues unique to digital platforms, such as text overflow or misplaced tags. Their gaming experience and familiarity with digital environments could provide them with an advantage in detecting errors specific to video game localisation. Lastly, on-screen readers may exhibit a higher level of digital literacy, which may be augmented by regular gaming activities, facilitating the ability to effectively identify errors during video game localisation testing.

Although not predictors of performance in error detection, several statistically significant relationships between proficiency in video game localisation testing and two dimensions from the expert questionnaire were observed. First, the results revealed a significant positive correlation between attitudes towards video games and the detection of visual errors and total errors. This finding, as indicated by responses to specific questions outlined in Table 3, suggests that greater familiarity and experience with video games may enhance error detection capabilities in localisation testing. Experience with video games could lead to a better understanding of the gaming environment, mechanics, and user interface, allowing individuals to recognise inconsistencies or errors more easily. Familiarity with common gaming conventions and practices may enable testers to anticipate potential issues, while experience with a variety of game genres could increase their adaptability to different testing scenarios. This is in line with the broader themes in video game localisation theory, as discussed
by Bernal-Merino (2015) or Mangiron (2018), and aligns with recent studies examining the impact of errors on video game localisation reception (Kudła, 2021). Furthermore, a positive attitude towards video games correlated with the detection of visual errors, offering an insight into the potential impact of personal interests on task performance. Second, punctiliousness correlated positively with error detection, implying that individuals who exhibit a heightened focus on punctuation and grammar are more adept at identifying errors in video game localisation. This increased attention to linguistic nuances may be attributed to a solid foundation in grammar and syntax, or an innate inclination towards linguistic precision. Such individuals are likely to possess a discerning eye for detail, enabling them to detect inconsistencies or errors in textual content more effectively. As a result, their propensity for identifying linguistic errors in localised video games may be enhanced, leading to more accurate and thorough localisation testing. This finding highlights the importance of involving testers with robust linguistic backgrounds and a meticulous approach to language when evaluating the quality of localised video games (Popović et al., 2006; Prieto Ramos, 2015; Toma et al., 2017).

However, several limitations must be taken into consideration when interpreting these results. First, the sample size was relatively small given the novelty of the research area and its exploratory nature, which, along with the differences in curricula between the institutions, may limit the generalisability of the findings. Additionally, the study focused on a specific type of video game localisation testing, which may not necessarily reflect the range of challenges encountered in different game genres or platforms. In line with this, the errors introduced for testing primarily included typographical, grammatical, and other technical inaccuracies, as categorised by Vázquez-Rodríguez (2018). While significant for assessing linguistic competencies, these errors do not fully encompass all potential localisation errors, particularly those involving cultural nuances and contextual appropriateness. As already stated, high-quality localisation involves more than just linguistic accuracy; it demands a comprehensive understanding of cultural contexts to ensure that content is contextually suited to target audiences. This aspect of localisation, requiring cultural sensitivity, was underrepresented in the error types used in our study. Third, the current study relied on self-report measures to assess individual traits, which may be subject to response biases. Finally, the sample was primarily composed of students, which might not accurately represent the skills and behaviours of professional localisation testers, raising questions about the generalisability of the results to the broader field of professional video game localisation testing. While acknowledging this limitation, it is important to note that Translation and Interpreting programmes are designed to develop skills relevant to the localisation industry — skills such as linguistic accuracy, cultural sensitivity, and technical proficiency. Also, to better align the study with industry standards, several localisation professionals were consulted during the experimental design phase to ensure that both the competencies evaluated and the errors introduced in the game are reflective of professional settings. All in all, the study’s experimental design could not fully emulate the complex professional environment of video game localisation.

Thus, the findings are suggestive rather than definitive and should be interpreted with caution. To address these limitations, future research could benefit from larger and more diverse samples that includes the participation of professional localisation testers. Additionally, a longitudinal component could be introduced to observe how the
development of the traits identified in this study may influence localisation testing performance over time in a professional context, perhaps to also explore the impact of individual traits on error detection in different types of video games. Additionally, incorporating objective measures of these variables, such as actual language proficiency tests, could enhance the robustness of future studies and provide a more comprehensive understanding of the factors influencing error detection in video game localisation.

5. Concluding remarks

In response to the question driving this study, a proficient localisation tester is likely to exhibit certain characteristics that enhance their ability to detect errors in video games. First, although seemingly evident, the data suggest that individuals with greater experience in translation, as determined by their year of study, tend to detect more errors. As participants progressed through their academic years, they demonstrated a greater proficiency in identifying inaccuracies, indicating that experience and education in translation and interpreting play a crucial role in honing the necessary skills for efficient video game localisation testing. Second, neuroticism emerged as another relevant factor for proficient video game localisation testers, as participants with higher emotional stability detected fewer visual errors and total errors overall. Thus, meticulous, vigilant individuals who remain focused and perceive assigned tasks as challenging exhibit better performance in detecting errors in video game localisation. Third, participants who prefer reading from screens seem better suited to the digital nature of video game localisation testing and more proficient at detecting errors. Finally, while not predictors of higher error detection rates per se, traits such as interest in and positive attitudes towards video games, punctiliousness, greater attention to punctuation and grammar, and careful, unhurried, content-focused reading positively correlated with error detection in video game localisation testing.

Taken together, the results of this study provide valuable insights into the factors that may influence error detection proficiency in video game localisation testing, suggesting that variables such as the year of study, attitudes towards video games, punctiliousness, screen vs paper preferences, and emotional stability can impact error detection capabilities. These results can inform the development of targeted training programs and strategies to enhance error detection proficiency in video game localisation, as well as to inform the recruitment and selection of localisation testers, as companies may prioritise candidates with certain traits that have been shown to correlate with better error detection capabilities.

References


Internacional de Medicina y Ciencias de La Actividad Física y Del Deporte, 22(85), 107–127. https://doi.org/10.15366/RIMCAFD2022.85.008


Data availability statement

The datasets generated and analysed during the current study are available in the institutional repository of the University of Alicante, accessible via http://rua.ua.es/dspace/handle/10045/142499.