Journal on Interactive Systems, 2025, 16:1 doi: 10.5753/jis.2025.5644

A gamified solution to promote positive habits in children and adolescents with Intellectual Disabilities

Pedro Afonso F. Michalichem [**University of São Paulo (USP)** | ***darknighmare1@usp.br***]
Leonardo Tórtoro Pereira [**São Paulo State University (UNESP)** | ***leonardo.t.pereira@unesp.br***]
Kamila Rios da Hora Rodrigues [**University of São Paulo (USP)** | ***kamila.rios@icmc.usp.br***]

Institute of Mathematical and Computing Sciences, University of São Paulo, Avenida Trabalhador São-carlense, 400 - Centro, São Carlos, SP, 13566-590, Brazil.

Received: 05 March 2025 • **Accepted:** 19 September 2025 • **Published:** 22 September 2025

Abstract: This paper outlines the development and assessment of a gamified solution called *Task Complete*, designed to enhance productivity and promote positive habits, particularly among individuals with Intellectual Disabilities. The solution converts daily tasks into motivational elements, motivating users through virtual rewards. It was created using User-Centered Design principles and was developed and tested with a focus on accessibility and usability. The project incorporated the PENS model and Web Content Accessibility Guidelines (WCAG) to ensure the solution is inclusive and suitable for a wide range of user profiles. We also performed a heuristic evaluation aided with Large Language Models in some tasks to identify issues and create a better solution.

Keywords: Gamified Web Solution, Task Complete, Usability, Game Evaluation, Accessibility, Intellectual Disability.

1 Introduction

Intellectual Disability (ID) is characterized by substantial limitations in cognitive functioning and behavioral adaptation, affecting areas such as practical, interpersonal, and conceptual skills. This condition generally manifests itself in the early stages of development, before the age of 18 [AAIDD, 2021]. ID involves restrictions in the development of essential functions for understanding and interacting with the environment and is observed in conditions such as Autism Spectrum Disorder (ASD) and Down Syndrome, among others [Organization *et al*., 1992].

Gamification, in turn, involves using game design elements in non-game contexts [Deterding *et al*., 2011; Marczewski, 2015], has been researched a lot in different fields like health [Lewis *et al*., 2016], software creation [Pedreira *et al*., 2015], as well as teaching [Bai *et al*., 2020]. It attempts to increase user motivation and engagement, as incorporating enjoyable elements has an immediate effect on these users.

It is widely viewed as a method to make various tasks or activities more enjoyable as well as appealing by incorporating many game elements like scoring, rewards, challenges, and progression [Marczewski, 2015].

Planning and developing gamification is indeed a complex task, which has ultimately led to the emergence of a field dedicated to carefully studying multiple methods for gamifying both virtual and non-virtual environments. To help developers, technicians, specialists, and researchers plan and implement gamification in certain situations, many frameworks, methods, processes, and other similar strategies have been mindfully created over time [Mora *et al*., 2017].

This document introduces the *Task Complete* Web solution, considerately created as a particular option to especially help children and teens with Intellectual Disabilities (ID)[1] in completely working on good, helpful behaviors as they make their everyday activities somewhat more fun and interesting. The method involves obtaining virtual coins for performing tasks, and then using those coins to acquire virtual avatar accessories.

This project comes from working with the ACORDE Institution[2], which helps people with ID. ACORDE has tried to improve its educational work by using things like digital games along with gamified setups.

When designing and developing *Task Complete*, we carefully considered accessibility requirements based on the Web Content Accessibility Guidelines (WCAG) and the GAIA Recommendations [Pichiliani, 2020], which are tailored to support children on the autism spectrum. By integrating these guidelines and standards, we aimed to improve the solution's accessibility and create a more engaging and inclusive experience for the user—referred to as the player in this context.

This paper is an extended version of the work presented in the SBGAMES'24: Proceedings of the XXIII Brazilian Symposium on Games and Digital Entertainment, titled – *Task Complete: A gamified solution to exercise positive habits in players with Intellectual Disabilities* [Michalichem *et al*., 2024]. We update the study with the heuristic evaluation of the usability of the task "add task", aided by Large Language Model (LLM) solutions, to find issues and recommendations for a better solution while providing a case study for the use of these applications to help in usability evaluation, as proposed by some recent works [Duan *et al*., 2024; Schmidt *et al*., 2024; Duan *et al*., 2023; Meinecke *et al*., 2025].

[1]Term used according to the Protocol for the Etiological Diagnosis of Intellectual Disability. Available at: `https://www.gov.br/conitec/pt-br/midias/relatorios/2020/20201203_relatorio_572_pcdt_deficiencia-intelectual_.pdf`.

[2]Available at: `https://institutoacorde.org.br/`

This analysis not only aids the quality of our solution, but is also a novel study case in the literature for generative AI-assisted heuristic evaluation using Nielsen's heuristics, with the only similar work we found being [Meinecke *et al*., 2025]. However, our analysis' novelty lies in that their study gave the LLMs access to the application's code, while we provided images like human evaluators are often provided. Furthermore, we propose a collaborative approach, showing how the LLM models may help humans to locate new issues. Finally, we also compare the efficiency of different models in our use case.

We also provide additions to the project's adherence to GAIA recommendations [Pichiliani, 2020] and which skills can be improved in the target audience with its implementation. We also provide new contrast tests, new system images, updated figures, and supplementary material with all steps from the heuristic evaluation.

The paper is divided as follows: Section 2 describes the related work, Section 3 describes the process of building the *Task Complete* gamified application, Section 4 brings a discussion about the importance of thinking about accessibility in computational solutions, and Section 5 describes the final considerations and future works.

2 Related Works

The following works explore gamified platforms in diverse settings or games created for individuals with intellectual disabilities. It's worth highlighting that our solution was tailored specifically for children and teenagers at the ACORDE institution, an organization dedicated to providing afternoon educational programs for individuals with Down Syndrome and Autism Spectrum Disorder [Michalichem *et al*., 2024].

Hosseini *et al*. [2022] examine the impact of gamification on task performance using a between-group experimental design in the context of the Covid-19 pandemic. Participants were asked to complete tasks related to: a) hygiene and infection (e.g., handwashing, maintaining distance, etc.); b) routines (e.g., walking daily, socializing with friends, cleaning the house, etc.); and c) personal matters (e.g., learning something new, checking in with a friend, etc.).

According to the authors, gamification improved the quality of work in task performance and subsequent submissions over time. Furthermore, it had a positive impact on timely deliveries.

Giacobo and de Souza [2023] introduce a gamified Web solution designed to engage and motivate students to complete assigned activities both inside and outside the classroom, within the set deadlines, in an enjoyable, fun, and competitive manner. The solution incorporates gamification elements such as levels, rewards, and competitions. Preliminary results from a pilot study conducted with the solution indicated that it was well received by students, who showed increased motivation and engagement in the proposed activities.

Domingos Filho and Vale [2017] introduce a gamified solution aimed at teaching physics to young people and adults. The project incorporates gamification elements, such as levels, rewards, and competitions, to boost students' interest and motivation in learning.

Simões-Silva *et al*. [2022] highlight that gamification can be beneficial for individuals with Autism Spectrum Disorder (ASD), "which is a persistent neurodevelopmental disorder that can be characterized briefly by deficits in verbal and non-verbal communication, difficulties in interaction, and manifestation of stereotyped movements or interests". According to the authors, programs, software, or mobile applications designed for individuals with ASD should focus on developing intrapersonal skills (such as motivation) and interpersonal/social skills.

da Cruz Netto *et al*. [2020] introduce the virtual environment "Our Life", created to help children with Down Syndrome memorize action sequences in their daily routines. The project involved a multidisciplinary team, and the effectiveness of the tool was evaluated with 30 children with Down Syndrome from a special education school for individuals with intellectual disabilities (APAE, acronym in Portuguese).

The children were divided into two groups: experimental (EG) and control (CG). The results showed that the EG outperformed the CG, with the mean progress of the children in the EG being 81.82% higher. According to the authors, the playful activities incorporated into the virtual environment captured the children's interest, encouraging them to have fun, test hypotheses, and question the sequences of actions in their daily routines.

Neves and Kanda [2016] present the findings of research that developed two educational games for individuals with intellectual disabilities. The games were evaluated through usability tests conducted with students from an APAE. The results showed that the games were well received by the students and are considered a promising tool for reinforcing learning among individuals with intellectual disabilities.

Venturelli and Ferraz [2019] present the findings of research that explored the perspectives of teachers who work with children with intellectual disabilities regarding the use of digital games to mediate mathematics instruction. The research suggested that teachers hold positive views on incorporating digital games into mathematics teaching for students with intellectual disabilities.

Mori *et al*. [2017] present research investigating the impact of games on the development of memory and attention in students with intellectual disabilities. The study found that students who participated in recreational activities performed better on memory and attention tasks compared to those who did not engage in these activities. The authors conclude that games can be an effective tool for enhancing these skills in students with intellectual disabilities.

Finally, Jadán-Guerrero *et al*. [2023] conducted a review of 66 studies on gamification in inclusive education for children with disabilities. They found that terms such as game-based learning, educational games, e-learning, gamification, and serious games were commonly used in these studies. Additionally, they observed that this research is being conducted globally, with contributions from nearly 200 authors across various universities in regions such as the USA, Europe, Brazil, and Mexico, and published in a wide range of outlets. Their review highlights that this is a

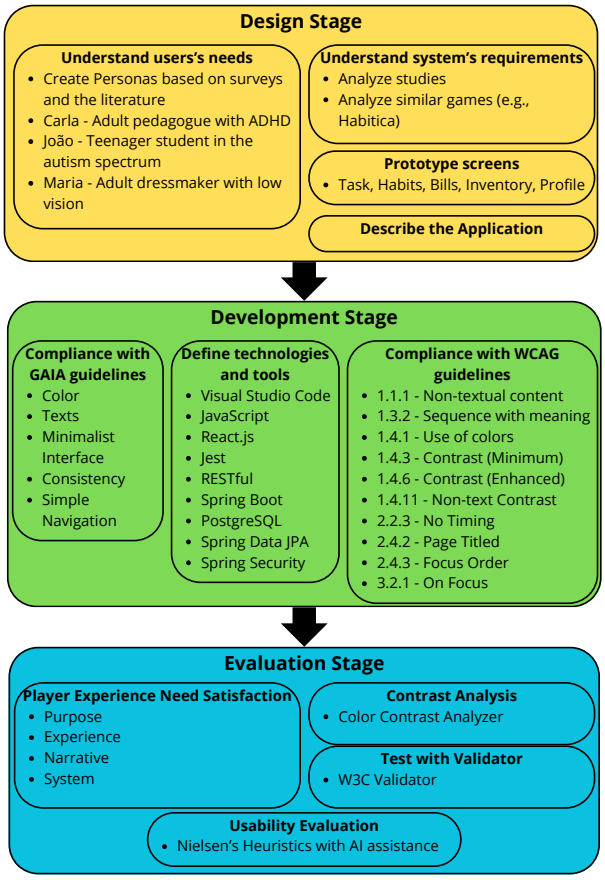


Figure 1. Overview of the design, development, and evaluation stages of *Task Complete*. **Source:** Adapted from Michalichem *et al*. [2024].

significant and varied area of research.

There are also several other studies focused on using games to promote healthy and positive habits in children with ASD and Down Syndrome [Viveiros *et al*., 2023], serious games as a therapeutic tool for individuals with intellectual disabilities [Martins *et al*., 2011], and serious games designed to help children with Intellectual Disabilities understand healthy eating habits using an iPad [Isasi *et al*., 2013].

Although the studies mentioned above demonstrate positive results, many of the solutions are no longer accessible or fail to meet crucial accessibility requirements for the context, which was a key demand from our partner institution. Therefore, we applied User-Centered Design (UCD) techniques, specifically tailored for games [Pagulayan *et al*., 2002], to design and evaluate a gamified solution aimed at promoting positive habits in children and adolescents with intellectual disabilities. This solution is intended to be used with the guidance of an education professional from the partner institution.

3 Task Complete: An Accessible Gamified Solution

This section outlines the design, development, and evaluation phases of the *Task Complete* solution. An overview of the process is provided in Figure 1, with detailed descriptions in the following sections.

3.1 Design Stage

During this stage, as outlined in the Human-Computer Interaction (HCI) literature [Helen *et al*., 2019], designers must assess the users' needs and define the system's intended requirements.

To gather the application requirements, the studies discussed in Section 2 were reviewed, along with similar commercial games, such as Habitica[3] and Uno's colorblind version[4].

In addition, based on the insights gained from the literature review, Personas [Cooper and Saffo, 1999; Pruitt and Grudin, 2003] were developed to represent the solution's target users. Personas are fictional characters created from various sources, helping designers identify the demographic traits, preferences, and behaviors of the target users [Pruitt and Grudin, 2003].

The Persona technique is often used in academic settings when direct access to the target audience isn't feasible during the initial stages of requirements gathering and problem definition. In this project, since approval from the Research Ethics Committee was still pending at that time, personas were developed to help us better understand the context and progress with the development phase. These personas provided valuable insights that shaped key design decisions, informed the mechanics of the gamified solution, and guided the creation of the narrative and accessibility features.

As a result, three personas were created to represent the key stakeholders: children or adolescents with ASD, caregivers, and education professionals. These personas were developed based on data and surveys about the characteristics of individuals with disabilities, as outlined by Valle and Connor [2014]. It's worth noting that the primary focus of this work is on individuals with ASD who are supported by our partner institution.

The created personas were:

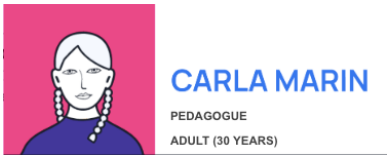
- **Carla Marin:** *represented in Figure 2a, she is a pedagogue who greatly values inclusive education. She is a 30-year-old who constantly seeks to strengthen her skills, actively participating in activities at the school where she teaches in the special education class. She also participates in courses offered by the education department in her city in this context. Carla also engages in creative activities such as community theater, adapted dance, and crafts and explores innovative ways to make the teaching-learning process more accessible, inclusive, and engaging. She was diagnosed with Attention Deficit Hyperactivity Disorder (ADHD) as a teenager and has difficulty focusing on activities that are not engaging her. This characteristic has helped her to get closer to her students.*
- **João Silva:** *represented in Figure 2b, is a curious teenager with a 14-year-old, diagnosed on the autism*

[3]Available at: `https://habitica.com/`. A productivity and task management app that uses gamification to motivate users to reach their goals and develop healthy habits.

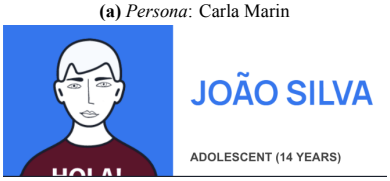
[4]Available at: `https://shop.mattel.com/products/uno-coloradd-hpp33`, a special edition of Uno designed for people with color blindness, featuring cards marked with the universal ColorADD color code for color blindness.

spectrum, who likes digital games. João has difficulty following routine activities and does not like environments with many people.

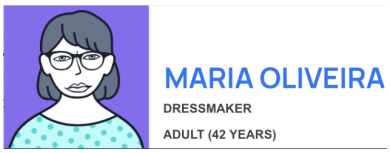
- **Maria Oliveira:** *represented in Figure 2c, is a 42-year-old, a self-employed seamstress, and a single mother. She lives in a poor community in Rio de Janeiro/Brazil and has a 6-year-old son who was recently diagnosed on the autism spectrum. Maria seeks ways to help her son communicate at home and regular school. She learned that a social institution in her community was carrying out activities with educational games and was interested in taking her son to participate. Maria has low vision and uses special glasses to carry out her daily activities.*



(a) *Persona*: Carla Marin



(b) *Persona*: João Silva



(c) *Persona*: Maria Oliveira

Figure 2. *Personas* representing the main target-users of the application. **Source:** Michalichem *et al*. [2024].

The personas assisted the team in making design decisions, such as:

- For Maria Oliveira, who has low vision and wants to monitor her son's activities within the gamified solution, a high-contrast color palette was developed. The visual elements were selected to enhance the identification and interaction with interface components, ensuring a more accessible and welcoming experience. Additionally, the application was designed with options to adjust font size and change screen contrast;
- For João Silva, the solution had to be intuitive and comfortable for users on the autism spectrum. João will use the gamified solution to assist with his daily activities. Therefore, the interface was designed with a cleaner and more simplified visual layout, incorporating rounded edges. This approach was intended to reduce cognitive overload, making the platform easier to understand and use for this type of user;
- For Carla Marin, the same accessibility elements mentioned earlier were considered. Given her ADHD, it's essential that she not only understands the solution but also finds ways to engage and motivate her students to use it. The solution was designed to be enjoyable and cost-effective, ensuring that Carla could easily implement it with her special education students, while helping them stay focused and follow the task routine playfully without distractions.

3.1.1 Application Description

The solution was designed after gaining a more in-depth understanding of the context and users.

Task Complete is a gamified Web platform that combines task management and gamification to enhance productivity and foster the development of positive habits. Its goal is to transform daily task management into an engaging gaming experience. Users guide avatars that progress and improve as they finish tasks, develop routines, and accomplish their objectives, creating a more engaging and rewarding experience.

The platform features four main areas of functionality. In the **Tasks section**, users can add activities to be completed in the future. The **Daily section** allows players to add tasks they need to accomplish each day. The **Habits section** helps users track tasks they want to repeat in a specific sequence to build lasting habits. Finally, the **Bills section** helps players manage their monthly bills. By completing these tasks, players earn money to purchase clothing for their avatars, allowing them to customize and personalize their appearance. Additionally, there is a leveling system that advances as tasks are completed.

A set of screens for the application was prototyped using the Figma tool. Figure 3 showcases some of these screens.

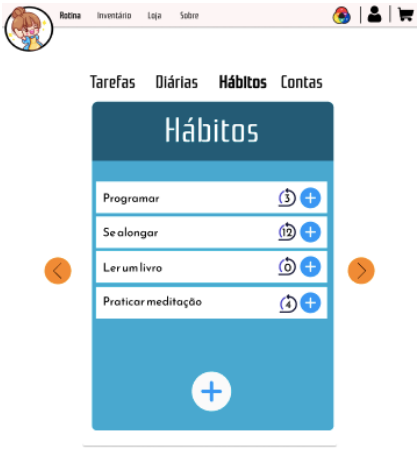
Figure 3a shows the **Task** section, which is the first option in the secondary menu. Here, specific tasks are created based on the player's needs, and the player can add as many tasks as desired. Once completed, the task will be marked as finished, and the player will earn coins as a reward for completing the tasks they set for themselves. In the top menu, next to the Task section, you'll find the **Daily tasks**, which are checked and updated every day. The **Habits** section (see Figure 3b), also located in the menu, keeps track of the sequence in which the player completes a habit. The sequence number increases each time the habit is successfully maintained, but it resets if the player misses a step. Finally, the **Bills** section (the last item in the menu—see Figure 3c) helps players organize and keep track of their monthly bills.

The inventory acts as a customization area (see Figure 4), allowing players to personalize their avatar within the app. It provides a clear view of all the items available on the platform, giving users the tools to make their avatar uniquely their own.

The player profile (see Figure 5) includes basic details like the player's name and email. It also features a progress bar that tracks the player's advancement, showing the avatar's current level and highlighting their progress within the platform.

A feature for increasing and decreasing font sizes was also designed to offer a more accessible and personalized experience for users who may have difficulty fully visualizing the components and text. This option is represented by the first icon on the main menu, located on the right side (see Figure 3[a]).

(a) Task Screen



(b) Habits Screen.

(c) Bills Screen.

Figure 3. Some prototyped screens for the *Task Complete* application (in Portuguese) - secondary menu items. **Source:** Michalichem *et al*. [2024].

Figure 4. Inventory screen (in Portuguese) - Main menu. **Source:** Michalichem *et al*. [2024].

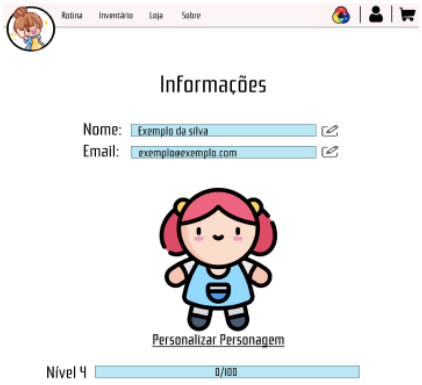


Figure 5. Profile screen (in Portuguese) - Main menu. **Source:** Michalichem *et al*. [2024].

A dark mode feature is also included, located in the top-right corner of the screen (it's the second icon on the main menu, on the right side—see Figure 3[a]). This option changes the color scheme from lighter to darker tones, improving both the user experience and accessibility by boosting the contrast of the elements shown on the screen.

3.2 Development Stage

The application was developed using web development technologies. The primary tool used for programming and structuring the code was *Visual Studio Code*, an Integrated Development Environment (IDE) known for its efficiency and advanced features. It was selected for its popularity and versatility. One of its key advantages is the extensive library of available extensions, allowing developers to customize VS Code by adding support for various programming languages, frameworks, and tools.

The programming language selected for implementation was *JavaScript*, a versatile and widely used language in Web development. Its dynamic, object-oriented nature provided a strong foundation for building an interactive and responsive application, perfectly suited to the project's objectives.

We chose the *React.js* library as the Front-End framework, as it offers a declarative and efficient approach to building complex user interfaces. Our decision was driven by its component-based architecture, enabling a modular and scalable code organization, enhancing the system's maintainability and extensibility. By using *React.js*, the development of an engaging user experience was streamlined

while ensuring an efficient and sustainable implementation of *Task Complete*.

For testing, we used testing tools integrated into VS Code, such as Jest. Unit tests were developed for each component, and they were run directly within the editor.

We implemented the *RESTful* architecture for the back-end, focusing on stateless communication and resource manipulation through standard HTTP operations. We chose *Spring Boot* due to its native support for *REST*, which streamlines the creation of *endpoints* and simplifies the handling of requests and responses.

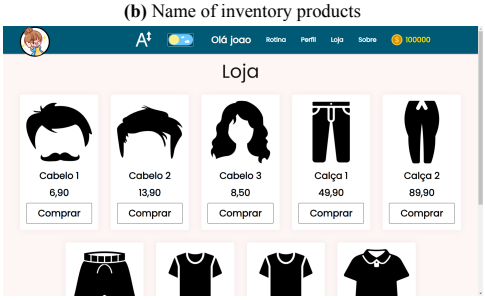
Moreover, we chose *PostgreSQL* as our database due to its reliability, robustness, and seamless integration with the *Spring* ecosystem. This integration was made easier through *Spring Data JPA*, which provides an abstraction layer for data access, simplifying interactions with the relational database.

Application security is ensured by *Spring Security*, which enables the implementation of authentication and authorization in a flexible and customizable manner. Additionally, *Spring Boot* provides production-ready features like metrics and monitoring, aiding in the efficient management of the application in a real environment.

Figure 6 illustrates several implemented screens, including the Habits Screen, the names of inventory products, and the names of products in the store.

(a) Habits Screen

(b) Name of inventory products



(c) Name of products in the store

Figure 6. Some screens developed for the *Task Complete* application (in Portuguese). **Source:** Michalichem *et al*. [2024].

3.2.1 Project adherence to WCAG guidelines - Accessibility

We incorporated several WCAG guidelines throughout the development of the project to enhance accessibility for players with varied disabilities. These guidelines are categorized into three compliance levels: A, AA, and AAA. These levels reflect a website's or an application's degree of accessibility based on the guidelines. In the *Task Complete* solution, the majority of the implemented guidelines fall under Level A, with additional guidelines from Levels AA and AAA. Some implemented guidelines include:

- **Success Criterion 1.1.1 - Non-text Content [Level A]:** Any "non-text" content that is essential for understanding the information must have a text alternative (visible or hidden) to describe the content, such as with captchas, for example.
  In the code snippet shown in Figure 7, an example of the *ALT tag* (alternative text) is provided, which is used to describe one of the images implemented in the solution;

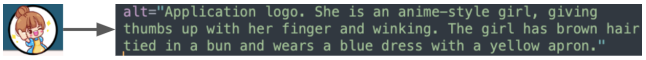


Figure 7. The code snippet demonstrates the use of the ALT tag to provide a description of the logo image. This feature is specifically used by screen readers to convey the content to users with visual impairments. **Source:** Michalichem *et al*. [2024].

- **Success Criterion 1.3.2 - Meaningful Sequence [Level A]:** Regardless of the interaction method, the information displayed on the screen must follow a logical and consistent sequence. For instance, when an error occurs in a form, it should include an alert icon, a message explaining how to correct the error, and a color change that highlights the information (not relying on color alone).
  In Figure 8, an example is shown where an incorrect email is entered, triggering an error message. This message alerts the user that valid information is required to proceed;

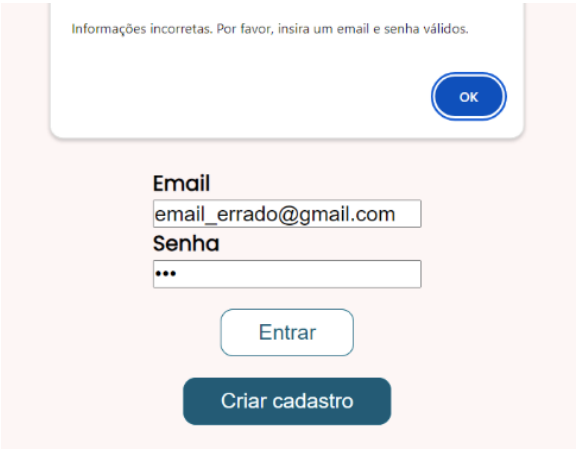


Figure 8. Input error and a clear message to the user (in Portuguese). **Source:** Michalichem *et al*. [2024].

- **Success Criterion 1.4.1 - Use of Color [Level A]:** Color should not be the sole visual means of conveying information, indicating an action, prompting a response, or distinguishing a visual element.

Figure 9 illustrates how this strategy is applied in the example of purchasing an item from the store. In this case, green is used to indicate a successful purchase, while a label is added to clearly represent the button's action;

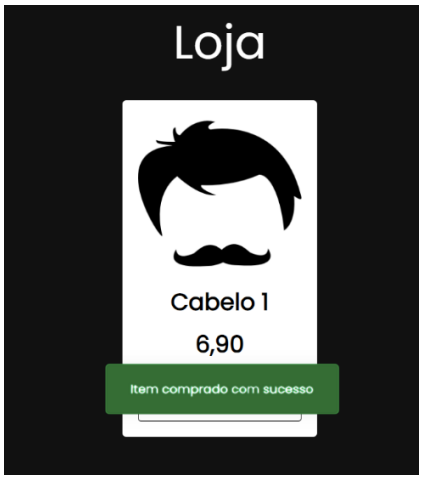


Figure 9. Example of a purchase confirmation message for an item in the store. (in Portuguese). **Source:** Michalichem *et al*. [2024].

- **Success Criterion 1.4.3 Contrast (Minimum) [Level AA] and Success Criterion 1.4.6 Contrast (Enhanced) [Level AAA]:** Text must have a contrast ratio of at least 4.5:1 between the foreground and background. **Note:** If the text font size is at least "18pt" or "14pt bold", the contrast ratio can be reduced to 3:1. Illustrations related to contrast, as per this guideline, will be provided in the following section (see Section 3.3.2). This includes contrast testing using the Color Contrast Analyzer[5] tool;
- **Success Criterion 1.4.11 Non-text Contrast [Level AA]:** Interface components (e.g., buttons) and images crucial for understanding the content must have a contrast ratio of at least 3:1 between the foreground and background.
  Illustrations related to contrast, in accordance with this guideline, will be provided in the following section (see Section 3.3.2), along with contrast testing using the Color Contrast Analyzer tool;
- **Success Criterion 2.2.3 No Timing [Level AAA]:** No functionality on the screen should require completion within a specific time limit. **Note:** Real-time events are exceptions to this rule.
  In *Task Complete*, there are no features that are time-limited;
- **Success Criterion 2.4.2 Page Titled [Level A]:** Every screen must have a clear main title that accurately describes its purpose.
  In Figure 6[a], the **Habits** screen is shown with a clear and descriptive title;
- **Success Criterion 2.4.3 Focus Order [Level A]:** The focus order of interactive elements on the screen must always be sequential and logical, based on the content presented;
  In Figure 10, the application menu is displayed, showing the logical sequence in which the elements are interacted with. It begins with general tasks, followed by daily tasks, habits, and finally, bills. These elements are all related to the player's routine, ensuring a logical and coherent flow of content;

Figure 10. Example of the content's logical sequence (in Portuguese). **Source:** Michalichem *et al*. [2024].

- **Success Criterion 3.2.1 On Focus [Level A]:** No unexpected contextual changes, such as opening a modal window, should occur when focusing on any element in the interface without direct confirmation (e.g., a confirmation button).
  One way to ensure consistency is by ensuring that an image displays the same text on every page where it appears. In *Task Complete*, for instance, item names are identical in both the inventory and the store, as shown in Figure 6[c].

The gamified application also complies with other WCAG guidelines, including 1.3.3 (Sensory Characteristics), 1.3.4 (Orientation), 2.4.11 (Focus Not Obscured - Minimum), 2.4.12 (Focus Not Obscured - Enhanced), 3.1.3 (Unusual Words), and 3.1.5 (Reading Level).

3.2.2 Project adherence to GAIA Recommendations

The project also adheres to the GAIA recommendations [Pichiliani, 2020]. GAIA is an open and collaborative set of 28 web accessibility guidelines focused on autism, covering areas from content writing to programmable resources. It is important to note that the GAIA recommendations were developed in alignment with the WCAG guidelines.

Below are some of the main recommendations we implemented, and the respective skills that can be improved in the target audience with its implementation [Pichiliani, 2020]:

- G01, which focuses on **Color** by ensuring it isn't the only way information is conveyed and that there's enough contrast between background and foreground elements to make content easily distinguishable:
  - Skills that can be improved: Attention, Reading, Verbal or linguistic comprehension, and Visual comprehension;
- G02, which deals with **Texts**, emphasizing the use of clear, straightforward language and avoiding unnecessary jargon:
  - Skills that can be improved: Attention, Reading, Verbal or linguistic comprehension;
- G04, which focuses on **Compatibility with the real world** in icons, images, and nomenclature of actions and menus to represent concrete actions and daily life activities that are easily recognized:

[5] Available at: `https://chromewebstore.google.com/detail/color-contrast-analyzer/dagdlcijhfbmgkjokkjicnnfimlebcll?pli=1` or `https://dequeuniversity.com/color-contrast`

  - Skills that can be improved: Memorization, Attention, Reading, Verbal or linguistic comprehension, Visual comprehension;
- G06, which focuses on **Informational Customization** by offering options to customize the display of information, such as player avatar images, with the player's individual preferences (see Figure 6[c]):
  - Skills that can be improved: Attention, Reading, Verbal or linguistic comprehension, Visual comprehension, Dealing with change or transitions, Sensory integration;
- G10, which highlights a **Minimalist Interface**, keeping the design simple and only including elements and content essential to the task at hand:
  - Skills that can be improved: Problem-solving, Memorization, Attention, Reading, Verbal or linguistic comprehension, Visual comprehension;
- G22, which promotes **Consistency**, ensuring that similar elements and interactions behave in predictable ways:
  - Skills that can be improved: Memorization, Visual comprehension, Dealing with changes or transitions;
- G23, which focuses on **Clickable appearance** by using larger icons, buttons, and form controls that provide adequate click/tap area and ensure they appear clickable:
  - Skills that can be improved: Attention, Visual comprehension, Sensory integration;
- G25, which supports **Simple Navigation**, offering a smooth and consistent way to move through different pages. These are just a few of the many guidelines we followed to improve accessibility and usability:
  - Skills that can be improved: Memorization, Attention, Visual comprehension, Dealing with changes or transitions;
- G27, which focuses on **Confirmation of actions** by offering feedback confirming correct actions or alerting about possible errors:
  - Skills that can be improved: Problem-solving, Attention, Verbal or linguistic comprehension, Visual comprehension, Dealing with change or transitions, Sensory integration.

3.3 Evaluation Stage

The inspection and checklist evaluations were conducted within the *Task Complete* interface. The following sections provide details of these evaluations.

3.3.1 Player Experience of Need Satisfaction - PENS

This model was the initial one used to assess the gamified solution developed. The goal of the evaluation was to determine whether the solution could create a positive experience for the player [Ryan *et al.*, 2006].

To apply the PENS model, the solution was examined through the lens of the gamification concepts outlined by Schell [2008], which are based on four dimensions: purpose, experience, narrative, and system. Each of these dimensions was then analyzed in relation to the *Task Complete* gamified solution:

- **Purpose:** The purpose of *Task Complete* is to support individuals with intellectual disabilities in managing their daily tasks related to personal or well-being activities. The solution integrates different gamification mechanics to encourage motivation and help players stay focused. For example, players can earn experience points and level up by completing tasks, setting goals, and overcoming challenges. They can also earn rewards, such as coins, which can be used in the store. These mechanics aim to keep players engaged and rewarded for their efforts, as illustrated in Figure 11 (areas marked with red rectangles);

Figure 11. Level and money mechanics (in Portuguese). **Source:** Michalichem *et al.* [2024].

- **Experience:** The solution incorporates several gamification mechanics to enhance user engagement. For instance, *Task Complete* features a user interface designed like a game, with a playful aesthetic that utilizes blue, orange, black, and white colors throughout the application. Players can personalize their avatars and equip them with items. Additionally, the solution uses border styles with rounded elements and fonts. These design choices are more visually accessible, as straight lines and sharp corners can be too heavy for individuals with autism, while curved lines and rounded edges offer a softer and less stimulating experience. Figure 12 illustrates these design features implemented in the solution;

Figure 12. Player interface with rounded border (in Portuguese). **Source:** Michalichem *et al.* [2024].

- **Narrative:** The narrative of *Task Complete* adds meaning to the gamification experience. In the solution,

the player controls an avatar that evolves as tasks are completed, gaining experience, leveling up, and unlocking new clothing items that can be personalized according to the player's preferences.
This narrative helps players connect with the application and boosts their motivation. By tracking their progress, users can feel a sense of accomplishment as they complete tasks and achieve goals. For example, players can measure their progress through their inventory, where they can see the items they've earned by consistently completing routine activities. This is illustrated in Figure 13;

Figure 13. Inventory interface with player progress (in Portuguese). **Source:** Michalichem *et al*. [2024].

- **System:** *Task Complete* is a direct, fair, and consistent solution. The rules are simple to understand and apply. The system is designed to be easy to use and intuitive, eliminating the need for a tutorial. Players can easily understand how to use the application. Additionally, the system is fair, ensuring that all users have an equal opportunity for success.

3.3.2 Contrast analysis - Test 1

The digital accessibility criterion of minimum contrast (1.4.3 in WCAG) is essential for supporting individuals with low vision and color blindness. Adhering to these guidelines enhances accessibility and improves the visual quality of the digital project for all users.

The free tool *Color Contrast Analyzer*, developed by Deque University[6], simplifies compliance with this criterion by enabling the analysis of the contrast between background and foreground colors for both large and small text. These principles can also be applied to presentation slides, social media graphics, e-books, and other digital content.

To conduct this test, a success criterion is needed to determine if the contrast is adequate for individuals with vision impairments to perceive the message. Therefore, the following success criterion, as recommended by the Color Contrast Analyzer, was adopted:

- **Success criterion:** Ensure a color contrast ratio of at least 4.5:1 for small text and 3:1 for large text, including text that is part of an image. Large text is defined in the requirements as 18pt (24 CSS pixels) or 14pt bold (19 CSS pixels).

In tests using the tool, the colors of the interface elements were assessed to ensure they met the following standards: 1.4.3: Minimum contrast (AA) and 1.4.6: Enhanced contrast (AAA). These standards apply to both small and large text, UI components, and graphic elements.

For the tests, the base colors used in the prototyping created in the Figma graphic editor were applied. These colors included shades of white (#FEF5F5, #FFFFFF), shades of blue (#005A75, #1DA7CF, #BAE7F4), black (#000000), and orange (#FF8C00).

Figure 14 displays the software used for conducting the tests. In this example, the contrast test met the WCAG AA and AAA standards for small text, large text, UI components, and graphical elements. The test involved the blue color (#FEF5F5) on the menu and white (#FFFFFF) text on that same menu.

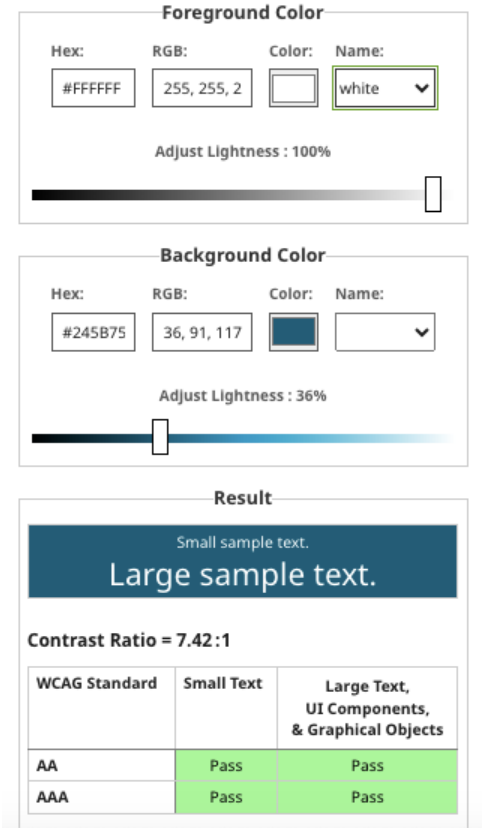


Figure 14. Use of the Color Contrast Analyzer - Approved menu color palettes. **Source:** Michalichem *et al*. [2024].

3.3.3 Contrast analysis - Test 2

We also evaluated the contrast between the foreground (font) and the background (solid colors, buttons, etc.), to analyze whether the elements allow people with low vision to identify the elements and texts on the screen. Figure 15 illustrates the test carried out, on two of the screens, which shows that the icons and fonts are perceptible when evaluated at Level AA - Medium Bold and Large Non-bold text (3:1).

In conclusion, the tests conducted using the tool, which covered the header with the menu, task topics, task item titles, arrow components, footer items, and text boxes, were all successful. The base colors from the Figma prototype, implemented in the Front-End, were applied correctly. As a result, the findings show that the project complies with the accessibility criteria for color.

[6]Available at: `https://dequeuniversity.com/`

Figure 15. Use of the Color Contrast Analyzer - Low vision

.

3.3.4 Test with Validator tool

The W3C Markup Validation Service, commonly referred to as the W3C Validator or Validator, is an online tool offered by the World Wide Web Consortium (W3C)[7]. This service enables the validation of HTML documents to ensure they comply with the specifications and standards set by W3C and WCAG.

Web developers can use the Validator by visiting its official website, where they can either input a page's URL (Uniform Resource Locator) or submit the code directly for validation. By following the Validator's suggestions, developers play a role in promoting a more standardized and accessible Web.

In this manner, *Task Complete* was also validated using the Validator. The errors that were identified are described below.

For instance, we received a warning: "Trailing slash on void elements has no effect and interacts poorly with unquoted attribute values". This typically happens when a slash (/) is added at the end of a void element tag, such as <img>,
, <hr>, or <input>. These elements do not have content and do not require a closing tag or trailing slash. This error was resolved in the code by checking for other void element tags with a trailing slash and removing the unnecessary slash or closing tag.

We encountered the error: "Bad value image for attribute type on element link: Subtype missing". This usually occurs when there is an issue with the declaration of a link (<link>) in the <head> section of the HTML document, specifically with the type attribute. The type attribute specifies the media type of the linked resource.

After examining the HTML code where the error was found, it was confirmed that the <link> tag was correctly implemented and that the type attribute's value was appropriate for the component being linked.

[7]Available at: `https://validator.w3.org/`

All the errors identified by the Validator tool were addressed, and none of them were related to accessibility; they were all coding issues.

3.3.5 Nielsen's Usability Heuristic Evaluation - Add Task

As a final evaluation step, we also conducted a heuristic evaluation using Nielsen's usability heuristics [Nielsen and Molich, 1990]. For this paper, we will discuss the heuristic evaluation of the task related to the "Add Task" procedure, and the "Add Habit" interface. We selected these as they are core functionalities and representative of the system as a whole. However, a careful evaluation of all other interfaces is required before releasing the final product. For the "Add Task", we use Figure 12 as the starting point, together with the desktop version of Figure 3a, illustrated in Figure 16.

Figure 16. Prototype screen for the *Task Complete* website. It shows The "Add Task" screen, with the "Close" red button above, and two text fields below: "Task Name" and "Date", respectively. Then, a checkbox with the "Important" text is illustrated, and the "Save Task" button.

We conducted the evaluation with the help of one of the authors who was not directly involved in the programming process (Person 1, 31 y.o., male), to allow for a less biased perspective, and two other specialists (graduate students) that were not involved in the project but have knowledge and experience on Heuristic Evaluation (Person 2, 27 y.o., male and Person 3, 28 y.o., female). Furthermore, following some recent trends, we also conducted the evaluation with four large language models in their free versions: ChatGPT (GPT-4o)[8], DeepSeek (V3)[9], Claude (3.7 Sonnet)[10], and Gemini (2.0 Flash)[11] [Duan *et al.*, 2024; Schmidt *et al.*, 2024; Duan *et al.*, 2023]. To avoid bias, the human evaluation was conducted before getting the results from the AI models.

The input for all models was the same, with a slight adjustment for Gemini, which can only receive a single image per input. In this case, we uploaded the first image, with the prompt "*Hold this image in your memory and wait for the next input*", and then sent the second image with the main prompt:

> *You are a PhD in Computer Science, a specialist in Human-Computer Interaction, and well-versed in Jakob Nielsen's works. Conduct Jakob Nielsen's heuristic usability evaluation on the Task*

[8]Available at: `https://chatgpt.com/`
[9]Available at: `https://chat.deepseek.com/`
[10]Available at: `https://claude.ai/chat`
[11]Available at: `https://gemini.google.com/`

> *Complete website, using the provided images that are, in order of upload: 1 - the main screen of the website, 2 - the page where you can add tasks, which appears after clicking the "Adicionar" button in image "1".*
>
> *For the heuristic evaluation, consider that the evaluated product is Task Complete. It is a website that is targeted at people with autism so that they can write down tasks, habits, bills to pay, and daily tasks. The website's user interface was developed with usability and accessibility heuristics, and guidelines focused on helping this group of users. The evaluated task is "Add a Task". You must follow Jakob Nielsen's heuristic usability evaluation guidelines, which can be checked on the following link, enclosed in <>: <https://www.nngroup.com/articles/ten-usability-heuristics/> Tips on conducting a good usability evaluation can be checked on the following website enclosed in {}: {https://www.nngroup.com/articles/how-to-conduct-a-heuristic-evaluation/.}*
>
> *For each of the 10 usability heuristics, propose a set of issues and recommendations, considering the 2 annexed images.*

The complete tables containing all issues and recommendations for each model and our human evaluation can be accessed in our Open Science Framework (OSF) project[12]. We maintained the original answers, even when they were irrelevant to a heuristic evaluation, and only added translations where the Portuguese terms were used without one.

Overall, DeepSeek and ChatGPT were able to provide a good evaluation, although sometimes with generic answers. Claude's answers sometimes did not follow the desired template, possibly due to it not having internet access and not being able to access the websites we suggested as guidelines. Moreover, it seldom wrote "issues" that were just reaffirmations of good design decisions. Finally, Gemini provided answers that did not follow the expected template, adding many issues in a single statement, and suggestions that did not match a specific issue. Furthermore, it focused on issues unrelated to the task in some cases and also provided many "issues" that were reaffirmations of good design choices.

A human evaluator read the output of each AI and, whenever needed, formatted it into a list of issues and their given recommendations. The human evaluator used the suggestions in the consolidation round, adding each suggestion to the consolidation sheet. Sometimes, the suggestions from the AI were placed in a different heuristic than other AIs or human evaluators suggested. In these cases, the human consolidator chose the best heuristic to add the suggestion, to avoid repeated entries. Also, the human evaluator was responsible to group similar suggestions according to their scope.

[12]Available at: `https://osf.io/qfj5v/?view_only=5d7f51364a8b4426802ebf83413557c1`

Here, we present a consolidation of the issues found in the format of a list. For each issue, we add a symbol to indicate their source (many symbols indicate a consensus on these issues and suggestions): (* - Person 1; # Person 2; % Person 3; † - ChatGPT; ‡ - DeepSeek; § - Claude; ¶ - Gemini). For this list, we ignored suggestions that did not relate to our specific task or did not belong to a heuristic evaluation scope. Furthermore, for brevity, the recommendations were omitted but can be found in our spreadsheet in our OSF repository.

1. Visibility of System Status:

- "Fechar/Close" is used with the intention to cancel creating a Task, but may incorrectly lead users to think it will close the application.*†‡
- "Salvar Tarefa/Save Task" is enabled even when the user has not completed filling the data. *†
- The "Adicione uma Tarefa/Add a Task" page does not provide feedback after a task is added. †‡
- The system status is unclear on the "Add Task" screen. #%
- The title from Screen 1 ("Add a Task") remains unchanged on Screen 2. #

2. Match Between the System and the Real World:

- The "Data/Date" field appears as a text field. *§
- Usually, a task does not have a name only, but also a description. *¶
- The field "Importante/Important" seems misleading: if the user wants to add a task to the app, it probably is important to them. *‡§
- Action buttons lack icons. %

3. User Control and Freedom:

- The system does not show a clear way to delete a created task. *
- The "Fechar" button (Close) is provided, but users may not know if their input will be saved or discarded when closing. †‡§
- No way to return to the previous screen from the "Add Task" screen. %

4. Consistency and Standards:

- The main task menu illustrates "Adicionar/Add" a Task. But in the task creation menu, the button to create the task uses the phrase "Salvar Tarefa/Save Task". *§
- The "Salvar Tarefa/Save Task" button is stretched through the screen, while the "Fechar/Close" button is not. There are other button styles and font inconsistencies. *‡
- On screen 2, the "Important" checkbox is placed too far away from its label. #
- The accessibility section uses a larger font than other sections. %

5. Error Prevention:

- When filling out the form to add a task, no feedback or validation is illustrated for wrong inputs. *%†‡§¶

- No feedback is given about what fields are required when adding a task. *%§¶
- Users might accidentally close the task addition page without saving. ‡
- On Screen 2, the "Close" button is overly prominent (centered and highlighted), increasing the risk of accidental clicks instead of saving. #

6. Recognition Rather than Recall:

- No pre-filled task examples for guidance. *†‡§
- Section meanings are unintuitive. %

7. Flexibility and Efficiency of Use:

- No known keyboard shortcuts or quick-add options to create tasks. *†‡§
- Tasks do not allow any customization settings. *¶
- No apparent way to create multiple tasks efficiently.§
- No shortcut for adding recurring tasks. †§¶

8. Aesthetic and Minimalist Design:

- When adding tasks, the user can still see buttons to change context and pending tasks. This is not necessary for the task creation context. *‡§
- The task addition form could be better structured and spaced for readability.†‡§
- The top menu has too many options, confusing users. %

9. Help Users Recognize, Diagnose, and Recover from Errors:

- No errors are shown when leaving blank fields when adding tasks or invalid inputs. *†‡§¶
- No undo options for accidental actions. †
- Users might not know how to correct mistakes in their input. ‡

10. Help and Documentation:

- No help information is given on how to create a task and what each field requires. *†‡§¶
- No onboarding tutorial or instructions for first-time users. †§
- It's unclear where users would go for assistance.%§

We may observe that 35 issues were consolidated across the 10 heuristics. 26 were identified by the human evaluators, and 25 by the AIs. The first person found 16 issues, the second found 4, and the third found 9. Claude was the AI that found the most issues: 16. DeepSeek and ChatGPT were very close, with 14 and 13 issues, respectively. The only model that fell short was Gemini, locating only 7 issues. Nonetheless, it is worth noting that all models, except for Gemini, contributed to locating at least one issue that the others did not find, while the same happened to human evaluators. Overall, 13 issues were found exclusively by 1 evaluator, showing that this diversity of evaluators helps finding issues that otherwise would not be discovered.

This means that using such models, at least for our use case, contributed to locating 10 issues that the human evaluators were unable to identify and also by recommending viable actions to solve said issues. Proportionally, the AI found more issues that humans did not find in heuristics 7 and 9, with half or more issues highlighted by AI alone. Finally, we call attention to the fact that the human evaluators are still a major component in such analysis, both by contributing to issues that models could not find and by also being able to consolidate the findings and define priorities and action courses.

(a) Main screen with the "Habits" tab selected.

(b) Screen with the field and button to add a habit.

(c) Main screen with the "Habits" tab selected after adding some habits.

3.3.6 Nielsen's Usability Heuristic Evaluation - Add Habit

To help generalize our findings, we conducted the heuristic evaluation on a second, but very similar task: to add a habit instead of a task. Figures 17a to 17c were shown to the evaluators (humans and AIs) and they were asked to provide another round of the heuristic evaluation (with the same prompt for the AIs). 40 issues were consolidated across the 10 heuristics, shown below. The recommendations are also present in our OSF repository [13].

We highlight that Claude and Gemini were unable to format the answer according to our input, showing, in most cases, a list of positive points instead of issues and suggesting recommendations that were not necessarily related to these "issues". However, we filtered the recommendations and added issues regarding then when relevant, making them contribute to the heuristic evaluation after this process. Also, the version of some AIs changed for this test: GPT-4o, DeepSeek-V3, Claude Sonnet 4, and Gemini 2.5 Flash.

1. Visibility of System Status:

[13] Available at: https://osf.io/qfj5v/?view_only=5d7f51364a8b4426802ebf83413557c1

- "Fechar/Close" is used with the intention to cancel creating a Task, but may incorrectly lead users to think it will close the application.*
- "Salvar Hábito/Save Habit" is enabled even when the user has not completed filling the data. *
- The "Habitos/Habits" menu button does not indicate the user is in the "Habits" menu. *#%§
- The title of the first screen [Add a Habit] remains the same on the second screen. #
- No visual feedback when clicking "Salvar Hábito/Save Habit" (e.g., loading indicator or confirmation). †‡§¶
- The transition between states (empty habit list → list with habits) is silent. †¶
- The "Hábitos" section does not clearly indicate which habits are active/completed. ‡

2. Match Between the System and the Real World:

- Usually, a habit may be daily, weekly, or have any other regular interval, and sometimes it is sporadic. The system does not allow this differentiation.*§¶
- Icons are missing on the main and "Add Habit" screens. %§
- On the "Added Habits" screen, the loop icon's function is non-intuitive, as well as the X. %†‡¶
- Only task items have icons on the "Added Habits" screen—action buttons lack them. %
- The term "Hábito/Habit" and others may not be fully clear without examples or descriptions. %†‡
- Unexplained elements: The meaning of the numbers "1" and "2" next to the habits, and the blue refresh-like icon (likely for marking as complete or resetting) and the red 'X' (for deletion), while somewhat standard, could benefit from tooltips or a brief explanation for users who might not infer their meaning easily. †¶

3. User Control and Freedom:

- The system does not show a clear way to edit or delete a created habit. *‡§¶
- "Fechar" button does not confirm if the user has unsaved data. As well as when deleting a habit. %†‡§¶
- No "Cancel" option on the "Adicione um Hábito" screen—users must click "Fechar" to exit.‡§

4. Consistency and Standards:

- The main task menu shows "Adicionar/Add" a Habit. But in the task creation menu, the button to create the Habit uses the phrase "Salvar Hábito/Save Habit". *‡
- The "Salvar Hábito/Save Habit" button is stretched through the screen, while the "Fechar/Close" button is not. *§
- Inconsistent font sizes and types between screens and sections.%§
- Layout spacing differs slightly from the "Add Task" section. †¶

5. Error Prevention:

- No feedback is given about what fields are required when adding a task. *§¶
- On screen 2, the "Close" button is prominently displayed at the center of the screen. This may cause inattentive users to accidentally click it instead of saving the task. #
- System allows submitting an empty habit name without any warning. *†‡§¶

6. Recognition Rather than Recall:

- The user does not know the last time the habit was updated. *
- No suggestions for habit names or categories. †‡
- Hard to search for habits. §¶

7. Flexibility and Efficiency of Use:

- No known keyboard shortcuts to create habits. *%‡§
- Habits do not allow any customization settings. *
- Too many options in the header menu, causing confusion. %
- Header customization is unavailable. %
- No support for batch input of habits. †§
- No quick-edit or add functionality on the list view. †‡§

8. Aesthetic and Minimalist Design:

- When adding habits, user can still see buttons to change context. This is not necessary for the habit creation context and makes the design cluttered. *%‡§
- Large buttons dominate the space. †

9. Help Users Recognize, Diagnose, and Recover from Errors:

- No error messages if habit creation fails or if user deletes a habit unintentionally. †§¶
- The system does not show a clear way to undo a mistakenly added streak to a habit. *†‡§¶
- No error messages for invalid inputs (e.g., duplicate habits). ‡§
- No color or icons to indicate error states. §

10. Help and Documentation:

- No help information is given on how to create a habit and what each field requires. *%†‡§¶
- No visual walkthroughs for newcomers. †‡§

From the 40 total issues, 23 were highlighted by humans and 26 by the AIs, showing how the latter may enhance the evaluation from the former. Person 1 found 13 issues, the second found 3, and the third person found 11. Claude found 17 issues, while ChatGPT and DeepSeek found 13, and Gemini found 12. 10 issues were reported by a single evaluator (mostly human ones). Heuristics 6, 8, 9 and 10 were proportionally more impacted by the aid of AI than the others, with half or more of the found issues being pointed out exclusively by AI evaluators.

Therefore, based on these two study cases, we bring evidence that using LLMs to aid in heuristic evaluation may help find early issues that human evaluators may not identify. This is especially useful as heuristic evaluations are usually conducted with a small sample of specialists. However, we reinforce that the human evaluation is an essential part

of said method, and the suggestions from the AI may not be applicable to the application, may be presented in an undesired format, and still need human validation for the consolidation step.

We used free models that may be readily available for users, and we recommend using a sample of different LLM models to achieve better results. However, Claude was the model that was able to find the most issues (although showing results in a wrong format), and ChatGPT was able to find slightly more new issues that humans did not find, with DeepSeek having similar results. So we recommend using at least one of these models to aid heuristic evaluations.

3.3.7 Evaluation Discussion

The evaluations carried out at this stage were supported by tools and carried out by the development team itself. Now, with approval from the ethics committee (protocol number 76853723.3.0000.5504), empirical tests are being planned to be carried out with target users from the partner institution to evaluate the effectiveness of the solution and the emotions aroused by its use. Moreover, the issues and recommendations gathered from the heuristic evaluation are being attended to and will help improve the system.

4 Accessibility Discussion

Developing and assessing a gamified solution for individuals with intellectual disabilities offers significant societal value in multiple ways. By integrating accessibility features, the solution provides entertainment and fosters digital inclusion, ensuring equal opportunities for participation.

The goal of this project was not only to provide entertainment but also to help individuals develop essential skills such as organization, time management, and task completion, all of which are crucial for autonomy and active participation in society. People with intellectual disabilities often face challenges in adapting to routines, which can negatively impact their lives. In this context, *Task Complete* enables these individuals to maintain consistent and structured routines, as the motivation provided by gamification encourages them to stay engaged.

By emphasizing the significance of accessibility, the project helps raise awareness about the needs of individuals with intellectual disabilities, fostering a culture of empathy and inclusion within society.

The evaluations conducted in the project include a range of metrics that emphasize the commitment to ensuring both accessibility and usability, with a specific focus on tests aligned with the WCAG guidelines. They revealed that there are still adjustments to be made to improve the user experience with the solution.

These evaluations ensure adherence to specific standards and guidelines while fostering a more inclusive user experience. By addressing factors like contrast, font size, and readability, the project meets regulatory requirements and offers a digital experience that respects and accommodates diverse user needs. This approach not only benefits individuals with intellectual disabilities but also helps create a more inclusive and mindful digital environment.

By detailing the methodological tests and accessibility practices, this work provides a comprehensive reference for developers aiming to design inclusive applications. In this way, the initiative has the potential to not only enhance the quality of life for individuals with intellectual disabilities but also contribute to the development of a more accessible, equitable, and inclusive digital environment that addresses the diverse needs of society.

This discussion is particularly relevant in the context of games and gamified solutions, as the industry continues to explore approaches for delivering more accessible solutions.

5 Final Remarks

During the development of the accessible gamified solution, critical factors were considered to ensure an inclusive and engaging experience for users with varied profiles. Guided by literature research, the requirements gathering process informed the creation of representative personas. The prototyping phase enabled the visualization and iterative refinement of the application's design, with a focus on elements such as color schemes, layout, and the spatial arrangement of components on the screen.

The Front-End implementation, employing technologies such as *React.js* was driven by the principles of efficiency, performance, and usability. The adoption of a modular architecture and structured code organization within *Visual Studio Code* reflected a commitment to the system's long-term maintainability. Additionally, the integration of unit tests throughout the development process contributed to ensuring the robustness and stability of the Front-End.

Regarding the Back-End, the selection of a *RESTful* architecture, supported by Spring Boot and *PostgreSQL*, offered a robust framework for developing scalable Web services. This approach facilitates the expansion of the solution, as the team plans to integrate additional functionalities in the future.

During the evaluation stage, several models were applied, including the PENS model, which focuses on competence, autonomy, relatedness, and intuitive controls. These factors were strategically integrated to optimize player satisfaction, achieve an appropriate balance between challenges and skills, promote autonomy in decision-making, foster emotional engagement, and ensure the usability of an intuitive interface.

Furthermore, the project incorporated WCAG guidelines to enhance accessibility, implementing alternative text for non-textual content and ensuring sufficient contrast. Tests conducted using the *Color Contrast Analyzer* tool confirmed compliance with WCAG AA and AAA standards, underscoring the project's dedication to both accessibility and visual quality. The evaluation addressed both technical considerations and user experience, with the goal of delivering an engaging and accessible experience for diverse user groups.

In conclusion, the accessible gamified solution *Task Complete* represents a concerted effort to eliminate barriers

and deliver entertainment and utility to an often neglected audience. The integration of accessibility features, alongside the use of robust technologies, significantly contributes to promoting digital inclusion. Continuous development and the active incorporation of community feedback are critical for the ongoing refinement and advancement of this accessible application.

Ongoing activities involve planning a longitudinal case study, which will be conducted in collaboration with educational professionals from the partner institution. In this study (which is awaiting approval from the educational coordinator), five individuals receiving support from the institution will be monitored for a period of one month as they use *Task Complete* to engage in activities of daily living and content related to the mathematics curriculum (treatment group). Throughout the intervention, professionals will assess the participants' performance. Upon completion, participants will be asked to complete the Game Experience Questionnaire [Poels *et al*., 2007; IJsselsteijn *et al*., 2013] and the Game Engagement Questionnaire [Brockmyer *et al*., 2009].

Simultaneously, another group of five individuals receiving support (control group) will work on the same skills using the methods currently implemented by the institution's professionals. At the end of the observation period, the professionals will assess the performance of participants in both groups to compare the different methodologies and evaluate the effectiveness of the proposed solution.

An accessibility assessment by inspection is also planned, using the same tools outlined in Section 3.3.5, and applying the Simplified Accessibility Assessment technique of de Santana *et al*. [2008], which is also based on heuristics.

Through professional observation, the collection of system interaction logs, and the assessment of experience and engagement via questionnaires, the objective is to empirically evaluate whether the gamified solution facilitated the development of skills related to daily routines and positive habits, as well as skills in the content of the associated disciplines, within the study group.

Declarations

Acknowledgements

The authors declare that the Google Translate tool was used to support the translation of the paper and the Grammarly tool for spelling adjustments. ChatGPT and DeepSeek were also used to rewrite some parts of the text.

Funding

The authors thank the Commission for Culture and Extension (CCEX) of the Institute of Mathematics and Computer Science (ICMC) at USP for financial support through extension project grants.

Authors' Contributions

Pedro Afonso F. Michalichem: software, methodology, visualization, data curation, writing-original draft, investigation. **Leonardo Tórtoro Pereira**: validation, data curation, formal analysis, investigation, methodology, visualization, writing-original draft, writing-review & editing, formal analysis. **Kamila Rios da Hora Rodrigues**: conceptualization, data curation, project administration, validation, visualization, investigation, writing-original draft, writing-review & editing, formal analysis, methodology, resources, supervision.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

All data generated and analyzed during this study are included in this paper, and supplementary materials can be found in our OSF's project[14].

Citation Diversity Statement

This study acknowledges the importance of diversity and inclusivity in academic citations. Our reference list includes works from a range of scholars with different backgrounds, institutions, and geographic locations, reflecting contributions from diverse perspectives in the fields of games.

By including this statement, we aim to raise awareness of citation bias and encourage equitable referencing practices in future research on serious games and inclusive education.

References

AAIDD (2021). Defining criteria for intellectual disability. Available at: `https://www.aaidd.org/intellectual-disability/definition`. Accessed on 21 September 2025.

Bai, S., Hew, K. F., and Huang, B. (2020). Does gamification improve student learning outcome? evidence from a meta-analysis and synthesis of qualitative data in educational contexts. *Educational Research Review*, 30:100322. DOI: https://doi.org/10.1016/j.edurev.2020.100322.

Brockmyer, J. H., Fox, C. M., Curtiss, K. A., McBroom, E., Burkhart, K. M., and Pidruzny, J. N. (2009). The development of the game engagement questionnaire: A measure of engagement in video game-playing. *Journal of experimental social psychology*, 45(4):624–634. DOI: https://doi.org/10.1016/j.jesp.2009.02.016.

Cooper, A. and Saffo, P. (1999). *The Inmates Are Running the Asylum*. Macmillan Publishing Co., Inc., USA. DOI: https://doi.org/10.1007/978-3-322-99786-9$_1$.

da Cruz Netto, O. L., Rodrigues, S. C. M., de Castro, M. V., da Silva, D. P., da Silva, R. R., de Souza, R. R. B., de Souza, A. A. F., and Bissaco, M. A. S. (2020). Memorization of daily routines by children with down syndrome assisted by a playful

[14]Available at: `https://osf.io/qfj5v/?view_only=5d7f51364a8b4426802ebf83413557c1`

virtual environment. *Scientific Reports*, 10(1):3144. DOI: https://doi.org/10.1038/s41598-020-60014-5.

de Santana, V. F., Almeida, L. D., and Baranauskas, M. C. C. (2008). Aprendendo sobre acessibilidade e construção de websites para todos. *Revista Brasileira de Informática na Educação*, 16(03). DOI: https://doi.org/10.5753/rbie.2008.16.03.

Deterding, S., Sicart, M., Nacke, L., O'Hara, K., and Dixon, D. (2011). Gamification. using game-design elements in non-gaming contexts. In *CHI'11 extended abstracts on human factors in computing systems*, pages 2425–2428. Association for Computing Machinery. DOI: https://doi.org/10.1145/1979742.1979575.

Domingos Filho, C. A. and Vale, M. E. B. d. (2017). Solução gamificada para o aprendizado de física no contexto do ensino de jovens e adultos. Technical report, Universidade de Brasília.

Duan, P., Warner, J., and Hartmann, B. (2023). Towards generating ui design feedback with llms. In *Adjunct Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*, UIST '23 Adjunct, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3586182.3615810.

Duan, P., Warner, J., Li, Y., and Hartmann, B. (2024). Generating automatic feedback on ui mockups with large language models. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*, CHI '24, New York, NY, USA. Association for Computing Machinery. DOI: https://doi.org/10.1145/3613904.3642782.

Giacobo, D. and de Souza, O. M. (2023). Gamifica- uma solução gamificada para engajar alunos de ensino médio de um curso técnico em informática. *Anais do Computer on the Beach*, 14:465–467. DOI: https://doi.org/10.14210/cotb.v14.p465-467.

Helen, S., Jenny, P., and Yvonne, R. (2019). Interaction design: beyond human-computer interaction.

Hosseini, C., Humlung, O., Fagerstrøm, A., and Haddara, M. (2022). An experimental study on the effects of gamification on task performance. *Procedia Computer Science*, 196:999–1006. DOI: https://doi.org/10.1016/j.procs.2021.12.102.

IJsselsteijn, W. A., De Kort, Y. A., and Poels, K. (2013). The game experience questionnaire. Technical report, Technische Universiteit Eindhoven.

Isasi, A. R., Basterretxea, A. L., Zorrilla, A. M., and Zapirain, B. G. (2013). Helping children with intellectual disability to understand healthy eating habits with an ipad based serious game. In *Proceedings of CGAMES'2013 USA*, pages 169–173. DOI: https://doi.org/10.1109/CGames.2013.6632628.

Jadán-Guerrero, J., Avilés-Castillo, F., Buele, J., and Palacios-Navarro, G. (2023). Gamification in inclusive education for children with disabilities: Global trends and approaches - a bibliometric review. In Gervasi, O., Murgante, B., Rocha, A. M. A. C., Garau, C., Scorza, F., Karaca, Y., and Torre, C. M., editors, *Computational Science and Its Applications – ICCSA 2023 Workshops*, pages 461–477, Cham. Springer Nature Switzerland. DOI: https://doi.org/10.1007/978-3-031-37105-9_31.

Lewis, Z. H., Swartz, M. C., and Lyons, E. J. (2016). What's the point?: a review of reward systems implemented in gamification interventions. *Games for health journal*, 5(2):93–99. DOI: https://doi.org/10.1089/g4h.2015.0078.

Marczewski, A. (2015). Even ninja monkeys like to play.

Martins, T., Carvalho, V., Soares, F., and Moreira, M. F. (2011). Serious game as a tool to intellectual disabilities therapy: Total challenge. In *2011 IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH)*, pages 1–7. DOI: https://doi.org/10.1109/SeGAH.2011.6165444.

Meinecke, A., Heidrich, D., Dworatzyk, K., and Theis, S. (2025). A comparative heuristic evaluation of kadi4mat through human evaluators and gpt-4. In Marcus, A., Rosenzweig, E., Soares, M. M., Rau, P.-L. P., and Moallem, A., editors, *HCI International 2024 – Late Breaking Papers*, pages 91–108, Cham. Springer Nature Switzerland. DOI: https://doi.org/10.1007/978-3-031-76821-7_8.

Michalichem, P., Pereira, L., and Rodrigues, K. (2024). Task complete: A gamified solution to exercise positive habits in players with intellectual disabilities. In *Anais do XXIII Simpósio Brasileiro de Jogos e Entretenimento Digital*, pages 474–495, Porto Alegre, RS, Brasil. SBC. DOI: https://doi.org/10.5753/sbgames.2024.241189.

Mora, A., Riera, D., González, C., and Arnedo-Moreno, J. (2017). Gamification: a systematic review of design frameworks. *Journal of Computing in Higher Education*, 29:516–548. DOI: https://doi.org/10.1007/s12528-017-9150-4.

Mori, N. N. R., dos Passos Santos, J. P., Shimazaki, E. M., Goffi, L. C. D., and Auada, V. G. C. (2017). Jogos e brincadeiras no desenvolvimento da atenção e da memória em alunos com deficiência intelectual. *Práxis Educativa*, 12(2):551–569. DOI: https://doi.org/10.5212/PraxEduc.v.12i2.0015.

Neves, L. A. and Kanda, J. Y. (2016). Desenvolvimento e avaliação de jogos educativos para deficientes intelectuais. In *Congreso Internacional de Informática Educativa (Conferência Internacional sobre Informática na Educação-TISE)*, page 612. Available at = `https://www.tise.cl/volumen12/TISE2016/612-617.pdf`, Access on 22 September 2025.

Nielsen, J. and Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 249–256. DOI: https://doi.org/10.1145/97243.97281.

Organization, W. H. *et al*. (1992). The icd-10 classification of mental and behavioral disorders. Technical report, WHO. Available at: `https://www.who.int/publications/i/item/9241544228`. Accessed on 21 September 2025.

Pagulayan, R. J., Keeker, K., Wixon, D., Romero, R. L., and Fuller, T. (2002). *User-centered design in games*, pages 915–938. CRC Press.

Pedreira, O., García, F., Brisaboa, N., and Piattini, M. (2015). Gamification in software engineering–a systematic

mapping. *Information and software technology*, 57:157–168. DOI: https://doi.org/10.1016/j.infsof.2014.08.007.

Pichiliani, T. C. P. B. (2020). *Gaia: Um Guia de Recomendações Sobre Design Digital Inclusivo para Pessoas com Autismo.* Appris, 1ª edição edition.

Poels, K., de Kort, Y. A., and IJsselsteijn, W. A. (2007). D3. 3: Game experience questionnaire: development of a self-report measure to assess the psychological impact of digital games. Technical report, Technische Universiteit Eindhoven.

Pruitt, J. and Grudin, J. (2003). Personas: practice and theory. In *Proceedings of the 2003 conference on Designing for user experiences*, pages 1–15. DOI: https://doi.org/10.1145/997078.997089.

Ryan, R. M., Rigby, C. S., and Przybylski, A. (2006). The motivational pull of video games: A self-determination theory approach. *Motivation and emotion*, 30:344–360. DOI: https://doi.org/10.1007/s11031-006-9051-8.

Schell, J. (2008). *The Art of Game Design: A book of lenses*. CRC press.

Schmidt, A., Elagroudy, P., Draxler, F., Kreuter, F., and Welsch, R. (2024). Simulating the human in hcd with chatgpt: Redesigning interaction design with ai. *Interactions*, 31(1):24–31. DOI: https://doi.org/10.1145/3637436.

Simões-Silva, V., Marques, A., Pinho, C., Pereira, D. G., Oliveira, J. F., and Barros, M. L. (2022). Gamification applied to autism spectrum disorder. In *Digital Therapies in Psychosocial Rehabilitation and Mental Health*, pages 163–186. IGI Global. DOI: http://doi.org/10.4018/978-1-7998-8634-1.ch008.

Valle, J. W. and Connor, D. J. (2014). *Ressignificando a deficiência: da abordagem social às práticas inclusivas na escola*. AMGH Editora.

Venturelli, S. and Ferraz, D. P. d. A. (2019). A visão do professor sobre jogos digitais no ensino da matemática para alunos com deficiência intelectual: Estado da arte. *Educação Matemática Pesquisa*, 21(1):180–196. DOI: https://doi.org/10.23925/1983-3156.2019v21i1p180-196.

Viveiros, C. S. d., Figueiredo, S. M. d. S., Mori, R. M. S. C., Nascimento, M. H. M., Sousa, J. M. A., and Teixeira, E. (2023). Jogo da alimentação saudável: tecnologia de prática educativa para crianças com síndrome de down. *Saude e pesqui.(Impr.)*, pages 11776–11776. DOI: https://doi.org/10.17765/2176-9206.2023v16n4.e11776.