This article has been accepted for publication in IEEE Revista Iberoamericana de Tecnologias del Aprendizaje. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/RITA.2025.3636607

Design and Validation of Guidelines for Creating Mathematical Applets for Students with Autism

Steven Van Vaerenbergh, Irene Polo-Blanco, Álvaro García Gómez

***Abstract*—Mathematical applets have become essential tools in the teaching and learning of mathematics, offering interactive and dynamic representations of mathematical concepts. For students with autism spectrum disorder (ASD), however, these applets must be designed to address their specific cognitive and learning needs to be truly effective. Despite increased efforts to support students with ASD in digital learning environments, no validated set of guidelines exist for designing mathematical applets tailored to these students. This study fills this gap by developing and validating a set of guidelines, operationalized for dynamic mathematical environments such as GeoGebra. In particular, the study advances beyond prior work by following a rigorous three-phase validation methodology involving two rounds of expert review with quantitative agreement analysis, and by explicitly grounding each guideline in the cognitive traits characteristic of ASD and their documented impact on mathematical learning. The multiphase process starts with a literature review that incorporates cognitive traits of ASD, their impact on mathematical learning, and interface design principles. The guidelines are then refined through two rounds of expert validation, first by specialists in ASD and mathematics education, then by experts in GeoGebra applet development. Expert feedback is systematically integrated, ensuring both theoretical soundness and practical feasibility. The final guidelines provide a structured approach for designing mathematical applets tailored to the needs of students with ASD, while also serving as a foundation for further research in this field.**

***Index Terms*—mathematical applets, mathematics education, autism spectrum disorder, inclusive learning environments, expert validation**

I. INTRODUCTION

In recent decades, mathematical applets have established themselves as essential tools for teaching and learning mathematics. The term "applet" refers to a small software application that runs within a web page and is designed for a specific interactive function. Applets are typically used to dynamically visualize complex algebraic and geometric relationships, offering the possibility to interact directly with mathematical elements and to observe the results of their manipulations in real time [1], [2]. This interactivity not only enhances the understanding of abstract and complex concepts but also increases students' motivation to delve deeper into the study of mathematics [3]. A wide variety of tools exist for building such applets, and they are being applied in increasingly diverse learning contexts [4].

For mathematical applets to be effective in an inclusive classroom, they must be designed with the specific challenges of students with special educational needs (SEN) in mind [5]. This consideration is especially important for students with ASD, who represent a significant group within this cohort. Although individuals with ASD exhibit a wide range of cognitive, verbal, and functional abilities, they often experience more learning difficulties than their typically developing peers [6]. Learning may be hindered by challenges related to attention, working memory, communication, and abstract thinking [6], [7]. The growing prevalence of ASD in educational systems [8] has highlighted that existing mathematical applets remain insufficiently adapted for this population [9], [10]. Although recent research has outlined general principles for creating accessible digital environments (e.g. [11], [12]), there remains a significant gap in specific, validated guidelines tailored to supporting their mathematics learning that provides a more thorough consideration of their cognitive traits. Notably, this issue aligns with broader efforts to promote educational inclusion through digital technologies [13].

This study makes two key contributions to the field: (i) it provides the first rigorously validated set of guidelines for mathematical applets designed for students with ASD, employing a structured three-phase methodology that includes two rounds of expert validation with different expert panels and quantitative agreement analysis; and (ii), unlike previous work that presents general interface recommendations, the proposed guidelines are explicitly grounded in the cognitive traits characteristic of ASD (executive function, sensory processing, visual strengths, working memory) and their documented effects on mathematical learning. This explicit linking between guidelines and cognitive profiles provides a principled foundation that distinguishes this work from general accessibility recommendations. Importantly, while the guidelines presented in this study are operationalized and validated within GeoGebra due to its widespread adoption in mathematics education and open-source accessibility, they are grounded in cognitive principles that transfer to other interactive mathematical learning environments.

The remainder of this article is structured as follows. Section II presents relevant literature on the characteristics of ASD, its impact on mathematical learning, and principles of human-computer interaction and interface design. Based on these findings, Section III details the methodology used to develop the guidelines, structured in three phases: the initial design based on the literature (Section III-A), validation by experts in ASD and mathematics education (Section III-B), and further validation by experts in GeoGebra applet develop-

Steven Van Vaerenbergh (e-mail: steven.vanvaerenbergh@unican.es) and Irene Polo-Blanco are with the Departamento de Matemáticas, Estadística y Computación, Universidad de Cantabria, Santander (Spain). Álvaro García Gómez is with IES Diego de Siloé, Burgos (Spain).

This research was funded by Ministerio de Ciencia e Innovación and Agencia Estatal de Investigación under grant PID2022-136246NB-I00.

This article has been accepted for publication in IEEE Revista Iberoamericana de Tecnologias del Aprendizaje. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/RITA.2025.3636607

ment (Section III-C). Section IV presents the results of these validation rounds, including an analysis of expert feedback and the corresponding modifications made. Finally, Section V discusses the findings, and the article concludes with key takeaways and future directions.

II. Literature review

A. Autism spectrum disorder

ASD is a neurobiological developmental disorder that manifests from childhood and persists throughout an individual's life. According to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [14], ASD is characterized by difficulties in two main areas: social communication and repetitive behavior patterns. These challenges generally include the following traits to varying degrees: (1) difficulties in understanding verbal and nonverbal language, (2) challenges with socialization, (3) preference for strict routines, which can result in resistance to changes, and (4) varied sensory sensitivities, ranging from hypersensitivity to hyposensitivity to visual, auditory, or tactile stimuli.

Recent studies indicate that up to 70% of individuals with ASD have comorbid conditions such as Attention Deficit Hyperactivity Disorder (ADHD) or Anxiety Disorders, which can further complicate their educational environment [15]. Approximately 33% of children with ASD also have severe intellectual disabilities, while about 42% demonstrate average or above-average intellectual capabilities [16].

In addition to the diagnostic criteria, research has shown that individuals with ASD often experience difficulties with executive function compared to their same-age peers with similar IQ levels. Executive function encompasses a range of cognitive skills, including planning, working memory, impulse control, inhibition, task-switching, and initiating or monitoring actions [17]. Several of these functions have been found to be challenged in individuals with autism, such as response inhibition [18], cognitive flexibility [19], and working memory [20].

Additionally, studies suggest that individuals with ASD frequently face early challenges in the area of attention [21]. On the other hand, reading abilities in children with ASD vary widely and are closely linked to language skills, with many struggling with comprehension (e.g., [22], [23]). These challenges people with ASD often face can create significant obstacles in the classroom, making it difficult for ASD students to engage with their learning [10].

On the other hand, the common strengths found in individuals with ASD should also be emphasized. For example, Grandin [24] highlights their abilities to respond to visual stimuli and their natural ability to think and reason through images and visual systems. Students with ASD also often show meticulous attention to detail [25] and adaptability to routines [6]. When designing instructional materials, it is important to consider both the challenges and strengths that students with ASD may exhibit in order to improve engagement and support conceptual understanding. Teachers should adapt the environment and teaching methods to fit both the challenges and strengths of students with ASD [6], [26].

B. ASD and mathematics

The acquisition of logical-mathematical skills in individuals with ASD is a relevant aspect for their educational development as well as their social and occupational inclusion. According to Oswald et al. [27], the ability to solve applied mathematical problems is associated with academic performance, everyday problem-solving skills, and employment outcomes, highlighting the importance of fostering these competencies from a comprehensive educational perspective.

However, due to some of the characteristics mentioned earlier, students with autism often face specific challenges in developing logical-mathematical skills [28], [29]. In relation to this aspect, various studies such as the one published by Sambade et al. [30] indicate that difficulties in executive functioning and language directly impact ASD students' performance in mathematics, affecting the generalization of learning, as well as planning and organization, attention, and problem-solving. Additionally, as noted by Alderson-Day [31], a poor understanding of language can be an additional obstacle in grasping mathematical terms and statements. In the same line, recent research has identified visuo-spatial and linguistic abilities to be the most predictive factor for early mathematical performance [29]. Other research, such as that by [32], has linked the difficulties students with autism face in solving mathematical problems to a low inhibition, verbal comprehension, or theory of mind.

As previously mentioned, other characteristics typical of the disorder, such as strong visual processing, can be key to enhancing the mathematical competence of students with ASD. This further emphasizes the importance of a deep understanding of their needs and potential in order to design effective pedagogical strategies that promote their mathematical learning. In this regard, various studies focused on the mathematical learning of these students propose adaptations that include visual supports, task structuring, or using the student's areas of interest to engage them in problem-solving [33]–[36].

C. Human-computer interaction

Human-Computer Interaction (HCI) is an interdisciplinary field concerned with the design, evaluation, and implementation of interactive computing systems for human use [37]. Within educational contexts, HCI research examines how learners engage with digital tools and environments, with particular attention to usability, cognitive demands, and the alignment between interface design and pedagogical goals [38]–[41]. As educational technologies have become increasingly prevalent in mathematics instruction, understanding the principles that govern effective human-computer interaction has become essential for creating tools that support rather than hinder learning.

A central concern in HCI, particularly relevant to educational applications, is the management of cognitive load. Cognitive Load Theory distinguishes between intrinsic load (inherent to the learning material), extraneous load (imposed by poor instructional design), and germane load (devoted to schema construction and learning) [42], [43]. Effective

This article has been accepted for publication in IEEE Revista Iberoamericana de Tecnologias del Aprendizaje. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/RITA.2025.3636607

interface design minimizes extraneous cognitive load, allowing learners to direct their mental resources toward understanding mathematical concepts rather than navigating the interface itself [44]. This principle becomes particularly critical in dynamic mathematical environments, where students must simultaneously process mathematical relationships and interact with interface controls. Research has demonstrated that poorly designed interfaces can significantly increase cognitive demands, leading to split-attention effects and reduced learning outcomes [45]. Also, it is important to note that a balance must be struck between HCI principles. Some studies have cautioned that an overly rigid adherence to simplification can be counterproductive. In particular, in order to reduce cognitive load, designers might inadvertently remove the *cognitive struggle* that benefits deep learning [46].

Visual design principles constitute another fundamental aspect of HCI that directly impacts educational software. The organization and presentation of information on screen must support users' perceptual and attention-related capabilities [47]. Gestalt principles of visual grouping, such as proximity and similarity, guide designers in creating coherent visual structures that help users understand relationships between interface elements [48]. In mathematical learning environments, visual clarity becomes crucial for representing abstract concepts and supporting conceptual understanding. The strategic use of visual hierarchy, consistent layout patterns, and appropriate use of color and contrast all contribute to reducing the perceptual effort required to interpret on-screen information [49].

A third pillar of HCI, relevant to educational technology, consists in interaction design, which is concerned with how users manipulate and receive feedback from digital systems. Norman's principles of interaction design emphasize the importance of visibility (making functions discoverable), feedback (confirming user actions), and consistency (maintaining predictable behavior) [50]. These principles take on particular significance in mathematical environments, where students may manipulate mathematical objects directly and observe the consequences of their actions in real time. This form of direct manipulation, when well-implemented, can provide powerful support for mathematical exploration and hypothesis testing [51], [52]. The immediacy of feedback in such environments allows students to develop intuitions about mathematical relationships through active experimentation, supporting constructivist approaches to learning [53]. However, the design of these interactive features must carefully balance responsiveness with stability, ensuring that manipulations feel natural while maintaining mathematical precision [54].

In the context of mathematics education specifically, HCI research has examined how different representational formats and interaction modalities support or constrain mathematical reasoning. Studies have investigated the cognitive benefits and challenges of multiple linked representations, where algebraic, geometric, and numeric representations are simultaneously displayed and dynamically connected [55]. While such environments offer rich opportunities for exploring mathematical connections, they also introduce design challenges related to visual complexity and divided attention [56]. The field continues to evolve toward more nuanced understanding of how interface design decisions impact not only usability but also the nature and quality of mathematical thinking that digital environments enable.

D. Interface design considerations for students with ASD

Building on the discussed HCI principles, several authors have proposed specific design guidelines for educational technologies intended for students with ASD, grounded in their specific cognitive and sensory profiles. These guidelines emerge from literature reviews, surveys, and practical application testing with users.

A fundamental principle in designing educational technology for students with ASD is minimizing sensory overload and cognitive strain. Santos et al. [11] and Groba et al. [12] emphasize the importance of reducing distractions and simplifying visual elements to aid focus, such as using muted colors, consistent layouts, and avoiding visually intensive multimedia elements.

Regarding any textual instructions included in the digital environment, Pavlov [57] suggests that sentences should be as short as possible, ideally placing one sentence per line to avoid confusion. Furthermore, the use of a clear, legible font such as Arial, at a size of 14 points or larger, is advised to ensure that text is easily readable. To implement emphasis, Freyhoff et al. [58] recommend using bold text or underlining.

Several authors point out that pictograms should accompany the text to enhance comprehension for individuals with ASD. For instance, Groba et al. [12] recommend using images to display information paired directly with written words to reinforce meaning. This guideline is supported by Smith et al. [59], who found that pictograms help clarify textual instructions and facilitate quicker understanding. Pavlov [57] further supports this by advocating for the use of clear, large buttons with both icons and text, ensuring that action controls have descriptive labels such as "Attach file" instead of ambiguous commands like "Click here".

The design of interactive elements should prioritize predictability and consistency. Groba et al. [12] highlight the importance of consistent button placement to foster a predictable interface layout, which can help students with ASD navigate the technology more effectively. Each interactive element should be clearly labeled and logically organized, maintaining a straightforward structure that aligns with the user's cognitive expectations. Further emphasizing structure, Rasche [60] recommends using step-by-step instructional designs to clarify tasks and reduce cognitive load for students with ASD. Supporting this, Abidoğlu et al. [61] note that structured approaches help children with ASD learn more effectively by allowing them to control their pace and engagement with the material.

In terms of applet dimensions, no universal guideline exists. Useful recommendations can be found in the MatesGG project, which focused on the design of GeoGebra applets and their use in the classroom [62]. There, a width-to-height ratio of 1.618, or the golden ratio, was adopted, primarily because it is close to a 16:10 aspect ratio and fits well on most modern

This article has been accepted for publication in IEEE Revista Iberoamericana de Tecnologias del Aprendizaje. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/RITA.2025.3636607

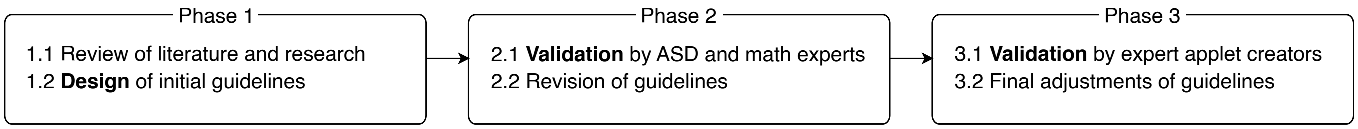

Fig. 1. Phases of the design and validation of the guidelines.

screens (including tablets, phones and monitors), leaving some space for interface elements.

Finally, the management of audio features in educational technologies requires careful consideration. Groba et al. [12] note that while synthesized speech is recommended to facilitate communication, auditory hyper-reactivity can lead to discomfort or negative reactions in some students with ASD. For this reason, it is recommended to have an option to disable audio features. In the case of platforms like GeoGebra, it is not always straightforward and may involve specific technical hurdles[1].

Although the literature offers several recommendations for designing digital interfaces accessible to students with ASD, these are often presented as general principles without validation or specific attention to mathematical learning. In particular, we found a lack of structured, validated guidelines that take into account the cognitive traits characteristic of ASD within the context of interactive mathematical applets. With this in mind, the present study aims to fill this gap by proposing a set of validated guidelines specifically designed for creating mathematical applets tailored to students with ASD.

E. Expert validation

The validation of instruments is fundamental across the social sciences to ensure that scientifically valid knowledge is generated. Kim [63] emphasizes that without rigorous validation, the reliability of research findings and their generalizability are compromised, particularly due to the subjective nature of many instruments used in data collection. While the issue of generalizability often pertains to questionnaires [64], the principles of validation are equally applicable to other types of instruments, such as checklists [65]. Instrument validation techniques vary widely, ranging from numerical methods like factor analysis and Cronbach's Alpha for assessing reliability and construct validity [64], to expert judgment and content analysis used to ensure content validity and practical relevance [63].

This work focuses on content validation through expert judgment, a methodology that involves assessment by individuals who are recognized in their field for their experience and specific knowledge and can provide information, evidence, judgments, and assessments [66]. In particular, experts should provide detailed insights on the clarity, coherence, relevance, and sufficiency of the instrument [67].

Expert judgment offers significant advantages, such as obtaining detailed and high-quality information. However, its implementation poses challenges, particularly in selecting experts and determining an adequate number. Clear and consistent criteria for expert selection are essential, considering factors such as education, experience, and academic recognition. Furthermore, it is essential to carefully determine the appropriate number of experts to ensure a diverse and representative validation panel. Diverse panels help mitigate the impact of individual biases that can skew assessments, promoting a more balanced and comprehensive evaluation [66]. In this regard, prior research on content validation suggests that panels of five to ten experts are generally sufficient, provided that participants are appropriately qualified and selected based on relevant expertise [67]–[69].

To ensure the effectiveness of the validation process, it is important to follow a series of organized and well-defined steps. This includes creating detailed templates which guide experts in evaluating each guideline against criteria such as clarity, relevance, and pertinence [65], [66], [70]. Initially, the process involves defining the objectives and scope of the validation effort, clearly outlining which aspects of the instrument are being validated and specifying the target population and context. This is followed by the selection of experts, where individuals are chosen based on their academic qualifications, experience, and recognition to ensure a balance of perspectives. Once the experts are selected, they are provided with comprehensive details about the instrument and its goals to guarantee informed evaluations. The process also includes facilitating active dialogue among the experts, allowing for a thorough exchange of insights. Finally, the responses from the experts are systematically reviewed using a structured response template, which helps in refining the instrument based on their insights [71].

III. Methodology

The design and validation of the guidelines were structured into three phases, as illustrated in Figure 1. Phase 1 focused on the initial design, translating insights from the literature into practical guidelines. Phase 2 involved validation by experts in ASD and mathematics education, who reviewed and evaluated the initially proposed guidelines and assessed whether they effectively address the needs of the students with ASD. Phase 3 consisted of a second validation round, conducted by experts in developing applets for the GeoGebra platform, focusing on refining the guidelines from technical and functional perspectives. The goal of the second validation round was to ensure the applets are both user-friendly and effective in educational settings. In both validation rounds, an analysis of expert agreement was conducted to assess the consistency of evaluations. In the following, each phase is described in detail.

[1] https://geogebra.github.io/docs/manual/en/commands/PlaySound/ GeoGebra's documentation for including sounds in applets.

TABLE I
OVERVIEW OF THE PRELIMINARY GUIDELINES, THEIR RATIONALE AND REFERENCE STUDIES.

Guideline	Rationale	References
G1. Short and clear instructions at the top.	The simplicity of the instructions can help with attention deficits and language comprehension difficulties.	[6], [14], [42], [43], [57]
G2. Use capital letters and highlight key words in boldface, such as verbs or concepts.	This can help with possible reading comprehension difficulties and take advantage of a frequent good instant recognition of words.	[22], [23], [58]
G3. Provide instructions with pictograms and a button to hide or show them according to each student's needs.	This may support the use of visuospatial strategies and reading comprehension.	[12], [57], [59], [72]
G4. Accompany the standard applet interface buttons with a pictogram (always the same) and place them in the same position for all applets.	This may provide consistent visual cues, which can improve understanding of the standard applet interface. The use of a consistent pictogram in the same position across applets aids in creating a predictable and supportive learning environment.	[12], [14], [24], [50]
G5. Provide a structured task, distinguishing each step clearly.	This may help with possible executive function deficits, in particular with planning, and reduce extraneous cognitive load.	[6], [17], [40], [42], [60], [61]
G6. Adopt visual simplicity by avoiding decorative elements and unnecessary colors.	This may support attention deficits by reducing distractions. A simpler visual design allows individuals with ASD to focus on the task at hand, particularly beneficial for those with challenges in inhibition. It also considers the strong visual processing abilities often characteristic of individuals with ASD.	[11], [12], [18], [21], [24], [47], [48]
G7. Avoid auditory content.	This may help with hyper- or hyporeactivity to sensory stimuli. Minimizing auditory distractions may reduce sensory overload in the task	[12], [14], [73], [74]
G8. Incorporate positive reinforcement in the form of images (pictograms) to encourage student motivation.	This can be effective for motivation by providing clear and visual cues that reinforce desirable behaviors, helping students engage and focus.	[6]
G9. Provide the opportunity to engage in new activities with different data through buttons.	Offering new problems or examples helps students practice flexibility and adapt to different situations, while using buttons provides a clear and predictable way to introduce changes.	[6], [17], [50], [51]
G10. Maintain a similar format throughout the applets.	A specific aspect ratio and dimension ensure a visually appealing and harmonious interface that is suitable for common device screen sizes.	[50], [62]

A. Design of initial guidelines

The initial design of the guidelines was based on the prior research from Section II on cognitive aspects, mathematical learning, HCI principles for educational software, and interface design considerations for students with ASD. Table I presents the preliminary set of guidelines, which is based on Gómez Casanueva's Master's thesis [75]. Each guideline is accompanied by a rationale that explains its basis and importance, taking into account the cognitive profile and interface design considerations for students with ASD, as well as the reference studies that support it.

B. Validation of the guidelines by experts in ASD and mathematics education

The first round of validation focused on evaluations conducted by a multidisciplinary team of experts in ASD and mathematics education (see Phase 2.1, Fig. 1). Selected experts were required to fulfill at least one of the following two criteria: i) demonstrated experience in research in mathematics education for students with SEN and/or ASD, and ii) ongoing academic or applied work in inclusive or special education. We used a convenience sampling strategy to recruit participants. Experts were contacted through existing research collaborations, professional networks, and educational institutions.

This process yielded a five-member panel combining theoretical and applied perspectives across different institutional contexts, career stages, and geographic regions within Spain. In particular, the panel consisted of: i) one university lecturer in mathematics education with 35 years of research experience and more than 10 peer-reviewed publications on inclusive mathematics education; ii) two doctoral researchers (both in their fourth year) specializing in inclusive mathematics education for students with ASD, each with five published journal articles on teaching and learning mathematics in students with ASD; iii) one postdoctoral researcher (one year postdoctoral experience) in mathematics education for students with ASD, with a doctorate in mathematics education and a master's degree in special education, focused on inclusive mathematics education and teacher training; and iv) one support specialist with specialized training in therapeutic pedagogy and ASD, practical experience in special education centers, and advanced training in educational psychology and attention to diversity.

All experts were affiliated with Spanish universities or educational support institutions across different regions of the country. This composition ensured evaluation from multiple professional perspectives: a senior researcher with an extensive publication record, emerging researchers actively investigating ASD-specific mathematics education, and a practitioner with frontline special education experience. While the geographic concentration reflects the national context of the research project, the diversity of expertise across career stages and professional contexts provided comprehensive assessment of both the theoretical soundness and practical applicability of the guidelines.

Initially, an email was sent to the experts outlining the purpose of the research. Two documents were attached to this message: the preliminary design of the guidelines and the evaluation document. The latter requested that experts assess the clarity, relevance/interest, and pertinence of each guideline using a five-point Likert scale (1 – Very Low, 2 – Low, 3 – Intermediate, 4 – High, 5 – Very High). Their assessments considered the specific needs and characteristics of individuals with ASD in the context of mathematics education. This

This article has been accepted for publication in IEEE Revista Iberoamericana de Tecnologias del Aprendizaje. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/RITA.2025.3636607

evaluation instrument was adapted from [67]. In addition to these quantitative evaluations, the instrument included an open-ended section for the experts to suggest improvements of each guideline. A final general question invited feedback on overarching aspects of the guidelines, such as potential omissions. All responses were collected within one week after the initial request.

Once all expert responses were collected, an individual analysis of each was conducted by one of the authors to examine the degree of agreement among experts and to identify significant patterns or discrepancies that required attention. Based on this analysis and the additional suggestions from the experts, the guidelines were subsequently revised and adjusted. Modifications were made to the original statements to ensure they accurately reflected the expert suggestions (see Phase 2.2, Fig. 1).

C. Validation of the guidelines by experts in mathematical applet creation

Expert selection for the second validation round followed the same strategy used in Phase 2. Candidates were required to have demonstrated expertise in designing mathematical applets and an active record of educational innovation with GeoGebra.

This process yielded a five-member panel of specialists in GeoGebra applet design, consisting of: i) Two secondary-school mathematics teachers with extensive GeoGebra experience, each having created over 1,000 publicly available applets, authored multiple publications on GeoGebra for educators, delivered teacher training workshops, and received national recognition for educational innovation; ii) two secondary-school mathematics teachers with advanced GeoGebra skills, having created 25 and 200 publicly available applets, respectively, delivered teacher training workshops, and organized regional conferences on technology in mathematics education; iii) one university professor in mathematics education with 35 years of research experience in computer-aided geometry and over 30 peer-reviewed publications on ICT-enhanced mathematics teaching.

All four secondary teachers were active members of the MatesGG project working group [62]. All experts were affiliated with Spanish institutions (secondary schools and universities) across different regions of the country. This composition balanced highly experienced GeoGebra innovators with accomplished practitioners and an established researcher in technology-enhanced mathematics education, ensuring comprehensive evaluation of both the technical feasibility and pedagogical effectiveness of the guidelines.

The evaluation process mirrored the approach used for the previous validation round. An email was sent explaining the purpose of the research and attaching the draft of the guidelines along with the evaluation document. This document prompted the experts to assess the clarity, relevance/interest, and pertinence of each guideline, and to provide suggestions for improvement. As in the previous validation round, the evaluation utilized a Likert scale from 1 to 5 to measure perceptions for each guideline, and each item included an open-ended question where experts could provide qualitative suggestions for improvement. Additionally, a specific request for new item proposals and useful command additions was included to capitalize on the experts' technical proficiency with GeoGebra. All responses were collected within two weeks after the initial request.

D. Analysis of the degree of agreement among experts

Once the responses from the experts across both validation phases were collected, an analysis was conducted on the responses to measure the degree of agreement among the evaluators. For this purpose, the Free-Marginal Multirater Kappa (FMK) index was calculated (in Phases 2.1 and 3.1, see Fig. 1). FMK was selected as the primary agreement measure because it accommodates variable marginal distributions across raters [76]. Unlike Fleiss' fixed-marginal kappa [77], which assumes raters distribute their ratings uniformly across all items, FMK permits experts to legitimately vary in their evaluation patterns, which is an appropriate assumption when some guidelines may warrant more critical scrutiny than others [78]. This flexibility is particularly important in validation studies where items vary in quality or applicability. As emphasized by Jakobsson and Westergren [79], chance-corrected agreement measures such as FMK provide more rigorous assessment than simple percentage agreement by distinguishing genuine consensus from coincidental agreement. FMK has been successfully employed in similar validation studies across education and medical fields, achieving values ranging from 0.52 to 1.00 across different assessment dimensions [80]–[82].

The FMK was computed separately for each evaluation dimension (clarity, relevance/interest, and pertinence) by aggregating responses across all guidelines. To interpret the FMK values, we used the categorization framework established by Landis and Koch [83], which classifies the strength of agreement based on the Kappa statistic as: Poor ($<$0); Slight (0 to 0.20); Fair (0.21 to 0.40); Moderate (0.41 to 0.60); Substantial (0.61 to 0.80); Almost perfect (0.81 to 1). This interpretation framework is widely adopted in instrument validation research and provides established benchmarks for evaluating inter-rater reliability [80], [81].

IV. Results

In this section, we present the evaluations and feedback received by the experts in ASD and mathematics education (Phase 2.2, Fig. 1), as well as by the experts in GeoGebra applet creation (Phase 3.2, Fig. 1). The section concludes with the final version of the validated guidelines along with a practical example showcasing their use.

A. Validation by experts in ASD and mathematics education

Table II summarizes the results of the validation conducted by the experts in ASD and mathematics education. The table includes the average Likert scores assigned to each guideline for clarity, relevance/interest and pertinence, as well as a summary of the expert suggestions, and the corresponding changes introduced in the guidelines.

The Likert scores, reflecting the evaluations of the original guidelines, showed overall positive feedback across the three

TABLE II
RESULTS FROM THE VALIDATION BY EXPERTS IN ASD AND MATHEMATICS EDUCATION.

G	C	R	P	Expert suggestions	Introduced changes
G1	4.4	5	5	Use clear and straightforward instructions with basic sentence structures for improved comprehension. Shift the focus from "attention deficits" to addressing sensory sensitivities and enhancing engagement, since attention deficits are not a standard diagnostic criterion for ASD.	The guideline was updated to emphasize the use of straightforward sentence structures for clarity. The rationale was also revised to better address the sensory sensitivities and engagement challenges associated with ASD, rather than generic attention deficits.
G2	4	4.4	3.8	Include an option allowing educators to select between uppercase and lowercase letters. Use a simple, clear font for ease of reading.	The rationale was expanded to provide a more comprehensive explanation. To maintain simplicity and avoid cluttering the applet interface, the option to switch to a lowercase font was not included.
G3	4.2	4.8	5	Make pictograms optional, as they may distract students who do not require them. Include additional explanations regarding their appropriate use.	Pictogram optionality was emphasized, and a more detailed explanation was added regarding where and how to use the pictograms to clarify the instructions.
G4	4.6	5	5	Clarify that the pictograms are fixed representations for specific actions and should correspond to buttons in the interface. Include specific details about their format and style to ensure consistency.	Details were added to specify the style, format, and functioning of the pictograms associated with buttons to enhance clarity and usability.
G5	3.8	5	5	Simplify the rationale by removing redundant phrasing to improve clarity. Optionally specify the total number of steps involved in the task to aid comprehension.	The guideline was updated to recommend optionally including the total number of task steps. The rationale was revised to remove redundant phrasing, ensuring a concise explanation of how structured tasks assist with executive function deficits.
G6	3.6	5	5	Include suggestions to ensure visual simplicity by avoiding distractions, emphasizing contrast, and limiting content to essential information. Adjust phrasing to better align with attentional focus.	Explicit instructions were added to the guideline to avoid unnecessary visual complexity and ensure a high-contrast design. Ambiguous terminology was replaced with a clearer phrasing that focuses on maintaining task engagement by minimizing attentional shifts.
G7	4.8	4.8	4.8	Provide the option to enable or disable auditory content, as some individuals with ASD may benefit from audio to enhance comprehension, while others may find it distracting.	The optionality of auditory content was added to the guideline.
G8	4.6	5	5	Incorporate text alongside pictograms (e.g., "Great job!") for positive reinforcement. Add a separate guideline focused on providing immediate, clear feedback indicating whether the answer is correct or incorrect.	A new guideline was introduced preceding positive reinforcement to specify that applets should provide immediate feedback on correctness. The original guideline, afterwards designated as G9, was refined to focus solely on motivational reinforcement, emphasizing the inclusion of positive phrases with the images.
G9	4.8	4.4	4.4	Avoid potential distractions, ensure buttons introducing new activities are not overly prominent.	A note was added to ensure clear instructions on the use of buttons and to avoid overuse. This guideline was then designated as G10.
G10	5	4.8	4.8	Maintain a similar format throughout the applets.	Since the original G10 addressed a technical aspect rather than a recommended guideline, it was removed and reserved for consideration in the future development of a template for applets. In its place, a new guideline, G8, focusing on correction and feedback, was introduced.

Note: G = Guideline, C = Clarity, R = Relevance/interest, P = Pertinence.

dimensions. Most items received an average score above 4 for *clarity*, with general agreement among the experts. Nevertheless, some variability was observed in specific items where suggestions for improvement were provided. Regarding *relevance/interest*, the majority of items were rated as "high" or "very high", with only two items receiving an "intermediate" rating. Similarly, for *pertinence*, most items were rated "high" or "very high", with only one item receiving a "low" rating, accompanied by a constructive suggestion for refinement.

The FMK index was calculated across all items for each dimension to assess the degree of agreement among the experts. For *clarity*, the FMK value of 0.36 indicated "fair" agreement, reflecting slight variations in evaluations for certain items. In contrast, *relevance/interest* achieved an FMK value of 0.68, denoting "substantial" agreement among the experts, with minimal differences in responses. Similarly, *pertinence* achieved an FMK value of 0.71, also indicating "substantial" agreement, despite a few outlier evaluations, including one "low" rating associated with a specific suggestion for improvement. These values are consistent with similar validation studies reporting FMK ranges of 0.52-1.00 [80]–[82]. The lower agreement for *clarity* (0.36) compared to content-focused dimensions reflects the inherently subjective nature of linguistic evaluation. This pattern is consistent with validation research showing that stylistic judgments typically exhibit greater variability than assessments of content relevance [81].

In addition, the experts provided qualitative suggestions for improvement, collected through the open-ended questions in the instrument. These included recommendations to enhance the clarity of specific items, refine language for better alignment with ASD characteristics, and address potential omissions. Based on this input, a thorough review of the guidelines was conducted, leading to modifications in aspects such as the inclusion of pictograms and task feedback. Additionally, one guideline (originally G10) was removed for addressing an overly technical aspect that is more effectively handled through standardized applet templates. Based on the expert suggestions, a new guideline was added. It was placed before the original G8, causing G8 and subsequent guidelines to be renumbered. The updated set of guidelines, not reproduced here explicitly, was then used for the second round of expert validation (Phase 3, Fig. 1), detailed in Section IV-B.

B. Validation by experts in mathematical applet creation

The results of the validation conducted by experts in mathematical applet creation are summarized in Table III.

TABLE III
RESULTS FROM THE VALIDATION BY EXPERTS IN MATHEMATICAL APPLET CREATION.

G	C	R	P	Expert suggestions	Introduced changes
G1	4.6	5	5	Specify the use of imperative forms in instructions to ensure consistency. Maintain uniformity and enhance readability by centering text consistently. Include illustrative examples to clarify instructional content.	The guideline was updated to specify imperative sentence structures and standardize text alignment. An illustrative example was added for clarity at the end of the guidelines document.
G2	4.8	4.6	5	Use a large sans-serif font (18pt or 'Large' setting in GeoGebra). Use LaTeX formatting to highlight key words in bold, ensuring that essential information stands out.	The use of the sans-serif font was included, together with font size requirements.
G3	4.8	5	5	Query as to whether the pictogram sequence should always be framed in a rectangle, indicating that flexibility in visual formatting might be beneficial.	Updated the guideline to state that the use of a rectangular frame for pictograms is optional.
G4	4.6	5	5	Include visual examples of buttons with their associated pictograms.	Example images of buttons with fixed pictograms were included.
G5	4.8	5	5	Include an indicator to show the current step in the task.	A recommendation was added to include a step indicator to help users track their progress.
G6	5	5	5	No suggestions were received.	No changes were introduced.
G7	4.4	4	4	Remove auditory content entirely, as its implementation is uncommon and can lead to compatibility issues across platforms.	The guideline was revised to explicitly recommend against including audio in the applets, as this feature is rarely used in platforms such as GeoGebra and can create unnecessary complexities.
G8	4.8	5	5	Distinguish between evaluative and exploratory activities. Solutions should always be shown for evaluative tasks. Specify the types of pictograms to indicate correct or incorrect responses.	The guideline was revised to distinguish between evaluative and exploratory tasks, specifying that solutions should be shown for the former. A recommendation to include standardized pictograms for feedback was added.
G9	5	5	5	Use a uniform format for positive reinforcement (e.g., consistent imagery or wording) to maintain coherence and support structured learning for students with ASD.	The wording was refined to mandate uniform reinforcement styles, using identical images or text to create a cohesive and predictable learning environment. Example pictograms with text in a standardized format were included to provide clear guidance for implementation.
G10	4.8	4.8	5	Specify the location of the button to start a new activity.	The fixed location of the 'New activity' button was specified as the top right corner for consistency and ease of use.

Note: G = Guideline, C = Clarity, R = Relevance/interest, P = Pertinence.

This table presents the average Likert scores for *clarity*, *relevance/interest*, and *pertinence*, as well as a summary of expert suggestions and the corresponding changes introduced to the guidelines.

The Likert scores for this phase were consistently high, with all guidelines receiving an average score of at least 4.4 across all dimensions. Most guidelines were rated as highly clear, relevant, and pertinent for the creation of mathematical applets adapted for students with ASD. A notable exception arose in the guideline related to the inclusion of sound, where one expert provided a lower rating. However, this score reflected a more specific consideration related to integrating sound within embedded applets, which did not impact the overall guidelines.

Suggestions from the experts in this round primarily focused on guidelines addressing technical aspects of applet creation, with fewer comments on non-technical items. Guideline G6, which pertains to visual simplicity, did not receive any suggestions for improvement, indicating it was well-aligned with the experts' expectations.

The lowest FMK value was observed for *clarity*, at 0.54, indicating "moderate" agreement. This can be attributed to a slight variation in responses, with two intermediate ratings and some variability between high and very high scores. In contrast, the evaluation of *relevance/interest* and *pertinence* achieved "almost perfect" agreement, with FMK values of 0.81 and 0.91, respectively, demonstrating strong consensus in these dimensions. The improvement in FMK values between rounds (clarity: from 0.36 to 0.54; relevance/interest: from 0.68 to 0.81; pertinence: from 0.71 to 0.91) demonstrates effective iterative refinement, with *pertinence* advancing from

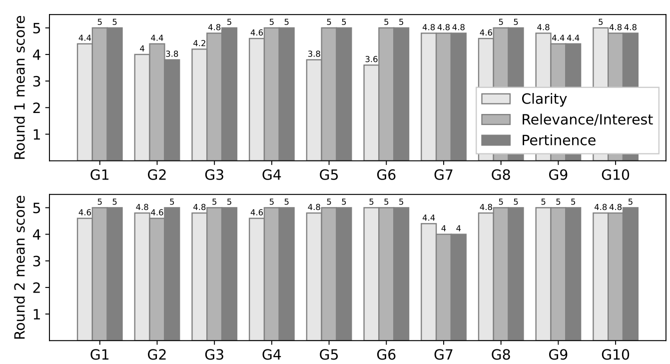

Fig. 2. Mean expert validation scores for the ten guidelines (G1-G10) across clarity, relevance/interest, and pertinence, on a 5-point Likert scale.

"substantial" to "almost perfect" agreement. This pattern of strengthening consensus through revision aligns with precedent from similar validation studies [80], [82], supporting the robustness of the multi-phase validation methodology. The validation scores obtained in both rounds of expert validation are summarized in Fig. 2.

C. Final guidelines

The final set of guidelines is presented in Table IV, incorporating all changes from the two rounds of expert validation to ensure clarity, relevance, and pertinence. To demonstrate their application, Fig. 3 presents an annotated example of a GeoGebra applet that follows the validated guidelines. The

This article has been accepted for publication in IEEE Revista Iberoamericana de Tecnologias del Aprendizaje. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/RITA.2025.3636607

TABLE IV
THE FINAL VALIDATED SET OF GUIDELINES.

Guideline	Rationale
G1. Short and clear instructions, always centered at the top.	Using straightforward, direct instructions enhances reading comprehension and reduces cognitive load, supporting individuals who may face challenges with sensory processing and complex language structures.
G2. Use capital letters in a large sans-serif font and highlight key words, such as verbs or concepts, in boldface.	The use of capital letters for instructions can help students with ASD quickly identify key directives. A large (size 18pt), rounded, sans-serif font is recommended to enhance readability. Highlighting key words in bold face, such as verbs or concepts that indicate the task's objective, further supports comprehension and focusses attention.
G3. Provide instructions with an optional pictogram sequence, which may be framed by a rectangle, and should include a button to toggle their visibility.	Using pictograms helps students with ASD by using visuospatial strategies and supporting reading comprehension. This feature is optional and particularly recommended for Early Childhood and Primary Education students.
G4. Replace standard applet interface buttons with fixed pictograms that correspond to specific actions, ensuring they are consistently styled and positioned in all applets.	Fixed pictograms for interface buttons provide consistent visual cues, creating a predictable and supportive environment for users. These pictograms will always represent the same action, be placed consistently across applets, and, when associated with interface-specific actions, will be framed and function as clickable buttons to execute the indicated action. (Example images are shown in Fig. 4.)
G5. Provide a structured task, clearly distinguishing each step, and optionally specifying the current step and total number of steps. Use a 'Next Step' button to advance through the task.	Structured tasks assist with executive function deficits by enhancing planning and organization. Specifying the total number of steps can improve comprehension, while a step indicator or a 'Next Step' button helps users track and progress through tasks. Sliders are to be avoided for this purpose, as they can complicate task navigation.
G6. Adopt visual simplicity by avoiding decorative elements, ensuring high-contrast colors, and limiting content to essential information.	Simplifying visual elements reduces distractions and supports individuals with ASD by enhancing focus. High-contrast colors improve visibility, while minimal content takes advantage of their visual processing abilities.
G7. Avoid auditory content.	Many people with ASD exhibit hyper- or hyporeactivity to sensory stimuli. By avoiding auditory content, the task becomes more accessible, allowing users to focus better. Additionally, audio is rarely implemented in GeoGebra and similar platforms, and its exclusion reduces potential compatibility issues when embedding applets.
G8. Provide immediate feedback on correctness using standardized pictograms. Include correct solutions for evaluative tasks and offer flexible feedback for exploratory ones.	Clear and immediate feedback helps students understand their performance and make necessary adjustments. Differentiating between evaluative and exploratory tasks ensures that feedback supports the learning goals of each activity. Standardized visual indicators, like pictograms, improve accessibility and consistency. (Example feedback images are shown in Fig. 5.)
G9. Incorporate positive reinforcement using standardized images with accompanying text to encourage motivation.	Consistent use of positive reinforcement, such as standardized pictograms with text, fosters motivation by providing clear and visual cues that reinforce desirable behaviors. This approach helps students with ASD engage and focus while offering a structured learning experience and reducing cognitive strain.
G10. Include a button in the top right corner to introduce new activities.	Offering new problems or examples helps students with ASD practice flexibility and adapt to different situations, while using buttons provides a clear and predictable way to introduce changes.

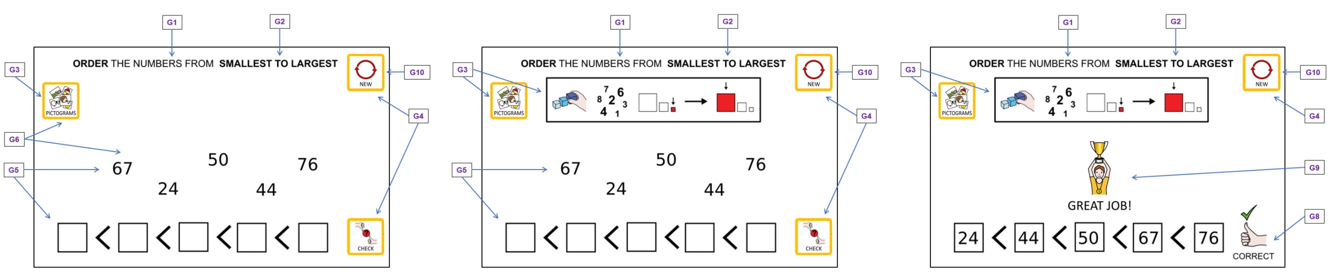

Fig. 3. Annotated examples of a GeoGebra applet following the validated guidelines, with annotations indicating the application of specific guidelines. Left: Initial screen. Middle: Screen displayed after clicking 'PICTOGRAMS'. Right: Final screen after correctly placing the numbers and pressing 'CHECK'. More examples of adapted applets can be found at https://matematicasyautismo.unican.es/english.html. Pictograms created by the Aragonese Center for Augmentative and Alternative Communication (ARASAAC), https://arasaac.org/. Used with permission.

Fig. 4. Example images of buttons associated with interface actions.

Fig. 5. Examples of positive feedback images with accompanying text.

screenshots highlight specific applet features, illustrating how each guideline was effectively implemented.

V. Discussion and Conclusions

In this study, we presented a validated set of guidelines for the creation of mathematical applets tailored to the needs of students with ASD. These guidelines respond to the demand for digital learning materials that are inclusive and adapted to the specific challenges faced by these students. Each guideline is supported by evidence from cognitive research, characteristics of the disorder that affect mathematical learning, and careful considerations of interface design. The guidelines are based on established HCI principles (particularly cognitive load management, visual design clarity, and consistent interaction patterns), adapted to the specific cognitive profile and learning needs of students with ASD.

Although validated within GeoGebra, the guidelines are fundamentally platform-independent, rooted in the cognitive characteristics of students with ASD rather than software-specific features. The applied design principles apply broadly to web-based applets, mobile applications, and other digital mathematical tools. While implementation details may vary, the underlying cognitive rationale remains constant. Future research examining implementation in other platforms would help identify any platform-specific considerations that complement these foundational guidelines.

Unlike previous work in which recommendations are provided based primarily on literature synthesis or pilot studies [11], [12], [57], our approach ensures both theoretical validity and practical feasibility through a rigorous three-phase validation process involving distinct expert panels and systematic agreement analysis. Furthermore, we explicitly ground each guideline in the cognitive profile of students with ASD (executive function challenges, sensory sensitivities, visual processing strengths, working memory limitations) and their documented impact on mathematical learning [6], [17], [28]–[32]. This theoretical anchoring distinguishes this work from general interface design principles by providing educators and developers with clear rationales rooted in the cognitive characteristics of the disorder. For instance, while reducing cognitive load is a universal HCI principle [42]–[44], the proposed guidelines specify *how* and *why* this matters for students with ASD in mathematical contexts: connecting visual simplicity (G6) to attention deficits [21] and inhibition challenges [18], structured tasks (G5) to executive function deficits [17], and pictogram use (G3, G4) to visual processing strengths [24] and language comprehension difficulties [31]. The resulting guidelines emphasize structured interfaces, clear feedback, visual supports, and step-by-step instructional design as essential components of educational applets, consistent with established principles [60], [61] while providing ASD-specific justification for each recommendation.

The guidelines are intended for teachers and content creators working in Early Childhood, Primary, and Secondary Education, with applicability that extends beyond any single platform. A notable advantage of these guidelines is their development with GeoGebra, a free and open-source software, which enhances accessibility and facilitates immediate application by educators and content creators. The proposed recommendations are platform-independent yet validated within GeoGebra, a widely used system. This makes them readily applicable in classrooms and adaptable to other digital learning environments.

A limitation of this study lies in the absence of direct validation involving students with ASD. By design, the present work focuses on establishing content validity and practical feasibility through expert review prior to classroom use. Empirical testing with students represents a crucial next phase. Conducting such validation requires a separate, longer research timeline, particularly due to challenges related to recruiting participants with ASD and the establishment of school partnerships and ethically grounded procedures for recruiting participants with ASD and coordinating with educators, families, and support or clinical centers. Future research will therefore focus on implementing and evaluating these guidelines in authentic classroom contexts (both inclusive and specialized) to examine their impact on (i) usability and accessibility (for instance, task completion, error patterns, time on task, prompts required), (ii) student engagement and sensory load, and (iii) learning indicators aligned with targeted mathematical competencies. This phased approach is intended to yield robust, generalizable evidence to refine and strengthen the guidelines.

Finally, although the guidelines in this study are developed for students with ASD, several principles such as visual clarity, reduction of extraneous cognitive load, predictable interaction patterns, and consistent scaffolding are plausibly relevant to other groups of students with SEN. Investigating transferability to additional profiles and identifying any profile-specific adaptations is a promising direction for future research.

References

[1] W. Daher, "Preservice teachers' perceptions of applets for solving mathematical problems: Need, difficulties and functions," *Journal of Educational Technology & Society*, vol. 12, no. 4, pp. 383–395, 2009.

[2] G. D. Heath, "Using applets in teaching mathematics," *Mathematics and Computer Education*, vol. 36, no. 1, pp. 43–52, 2002.

[3] S. Radović, M. Radojičić, K. Veljković, and M. Marić, "Examining the effects of Geogebra applets on mathematics learning using interactive mathematics textbook," *Interactive Learning Environments*, vol. 28, no. 1, pp. 32–49, 2020.

[4] L. Echeverría, R. Cobos, and M. Morales, "Improving the students computational thinking skills with collaborative learning techniques," *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, vol. 14, no. 4, pp. 196–206, 2019.

[5] C. d. P. Gallardo-Montes, M. J. Caurcel Cara, and A. Rodríguez Fuentes, "Technologies in the education of children and teenagers with autism: evaluation and classification of apps by work areas," *Education and Information Technologies*, vol. 27, no. 3, pp. 4087–4115, 2022.

[6] S. Ozonoff and P. L. Schetter, "Executive dysfunction in autism spectrum disorders," in *Executive function in education: From theory to practice*, L. Meltzer, Ed. New York: Guilford Press, 2007, pp. 133–160.

[7] S. Baron-Cohen, H. Tager-Flusberg, and M. Lombardo, *Understanding other minds: Perspectives from developmental social neuroscience*, 2nd ed. Oxford University Press, 2000.

[8] C. A. McDonald, J. P. Donnelly, A. L. Feldman-Alguire, J. D. Rodgers, C. Lopata, and M. L. Thomeer, "Special education service use by children with autism spectrum disorder," *Journal of autism and developmental disorders*, vol. 49, no. 6, pp. 2437–2446, 2019.

[9] M. García-Moya and R. Blanco, "Applications with mathematical content for users with autism," *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, vol. 19, pp. 111–119, 2024.

[10] C. Mallory and B. Keehn, "Implications of sensory processing and attentional differences associated with autism in academic settings: An integrative review," *Frontiers in Psychiatry*, vol. 12, p. 695825, 2021.

[11] M. I. Santos, A. Breda, and A. M. Almeida, "Design approach of mathematics learning activities in a digital environment for children with autism spectrum disorders," *Educational Technology Research and Development*, vol. 65, pp. 1305–1323, 2017.

[12] B. Groba, L. Nieto-Riveiro, N. Canosa, P. Concheiro-Moscoso, M. d. C. Miranda-Duro, and J. Pereira, "Stakeholder perspectives to support graphical user interface design for children with autism spectrum disorder: A qualitative study," *International Journal of Environmental Research and Public Health*, vol. 18, no. 9, p. 4631, 2021.

[13] C. S. González-González, J. Muñoz-Arteaga, and C. A. Collazos, "Educational inclusion through ICT," *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje*, vol. 16, no. 4, pp. 352–354, 2022.

[14] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders (DSM-5)*. American Psychiatric Association Publishing, 2013, vol. 5.

[15] M.-C. Lai, M. V. Lombardo, and S. Baron-Cohen, "Autism," *The Lancet*, vol. 383, no. 9920, pp. 896–910, 2014.

[16] M. Maenner, K. Shaw, J. Baio, A. Washington, M. Patrick, M. DiRienzo, D. Christensen, L. Wiggins, S. Pettygrove, J. Andrews, M. Lopez, A. Hudson, T. Baroud, Y. Schwenk, T. White, C. Rosenberg, L. Lee, R. Harrington, M. Huston, A. Hewitt, A. Esler, J. Hall-Lande, J. Poynter, L. Hallas-Muchow, J. Constantino, R. Fitzgerald, W. Zahorodny, J. Shenouda, J. Daniels, Z. Warren, A. Vehorn, A. Salinas, M. Durkin, and P. Dietz, "Prevalence of autism spectrum disorder among children aged 8 years-autism and developmental disabilities monitoring network, 11 sites, United States, 2016," *MMWR Surveillance Summaries*, vol. 69, no. 4, pp. 1–12, 2020.

[17] E. L. Hill, "Evaluating the theory of executive dysfunction in autism," *Developmental review*, vol. 24, no. 2, pp. 189–233, 2004.

[18] C. Sanderson and M. L. Allen, "The specificity of inhibitory impairments in autism and their relation to ADHD-type symptoms," *Journal of autism and developmental disorders*, vol. 43, no. 5, pp. 1065–1079, 2013.

[19] Y. Yasuda, R. Hashimoto, K. Ohi, H. Yamamori, M. Fujimoto, S. Umeda-Yano, H. Fujino, and M. Takeda, "Cognitive inflexibility in japanese adolescents and adults with autism spectrum disorders," *World journal of psychiatry*, vol. 4, no. 2, p. 42, 2014.

[20] L. Bennetto, B. F. Pennington, and S. J. Rogers, "Intact and impaired memory functions in autism," *Child development*, vol. 67, no. 4, pp. 1816–1835, 1996.

[21] B. Keehn, R.-A. Müller, and J. Townsend, "Atypical attentional networks and the emergence of autism," *Neuroscience & Biobehavioral Reviews*, vol. 37, no. 2, pp. 164–183, 2013.

[22] J. Åsberg Johnels, E. Carlsson, C. Norbury, C. Gillberg, and C. Miniscalco, "Current profiles and early predictors of reading skills in school-age children with autism spectrum disorders: A longitudinal, retrospective population study," *Autism*, vol. 23, no. 6, pp. 1449–1459, 2019.

[23] C. R. Jones, F. Happé, H. Golden, A. J. Marsden, J. Tregay, E. Simonoff, A. Pickles, G. Baird, and T. Charman, "Reading and arithmetic in adolescents with autism spectrum disorders: peaks and dips in attainment." *Neuropsychology*, vol. 23, no. 6, pp. 718–728, 2009.

[24] T. Grandin, *Thinking in Pictures*. New York, NY, USA: Vintage Press Edition, 1996.

[25] F. Happé and U. Frith, "The weak coherence account: detail-focused cognitive style in autism spectrum disorders," *Journal of autism and developmental disorders*, vol. 36, pp. 5–25, 2006.

[26] G. B. Mesibov and V. Shea, "The TEACCH program in the era of evidence-based practice," *Journal of autism and developmental disorders*, vol. 40, pp. 570–579, 2010.

[27] T. M. Oswald, J. S. Beck, A.-M. Iosif, J. B. McCauley, L. J. Gilhooly, J. C. Matter, and M. Solomon, "Clinical and cognitive characteristics associated with mathematics problem solving in adolescents with autism spectrum disorder," *Autism Research*, vol. 9, no. 4, pp. 480–490, 2016.

[28] J. C. Bullen, M. C. Zajic, N. McIntyre, E. Solari, and P. Mundy, "Patterns of math and reading achievement in children and adolescents with autism spectrum disorder," *Research in Autism Spectrum Disorders*, vol. 92, p. 101933, 2022.

[29] R. Fernández-Cobos, I. Polo-Blanco, E. Castroviejo, M. Juncal-Ruiz, and A. Vicente, "What predicts early math in autism? a study of cognitive and linguistic factors," *Journal of Autism and Developmental Disorders*, in press.

[30] L. Sambade, B. Fraga González, and B. López, "Aprendizaje lógico-matemático en tea y problemas de atención," *Revista de Estudios e Investigación en Psicología y Educación*, no. 11, pp. 184–187, 2017.

[31] B. Alderson-Day, "Verbal problem-solving difficulties in autism spectrum disorders and atypical language development," *Autism Research*, vol. 7, no. 6, pp. 720–730, 2014.

[32] I. Polo-Blanco, P. Suárez-Pinilla, J. Goñi-Cervera, M. Suárez-Pinilla, and B. Payá, "Comparison of mathematics problem-solving abilities in autistic and non-autistic children: the influence of cognitive profile," *Journal of Autism and Developmental Disorders*, vol. 54, pp. 353–365, 2024.

[33] A. Bruno, I. Polo-Blanco, S. Van Vaerenbergh, R. Fernandez-Cobos, and M. J. González López, "Strategies for solving multiplicative problems using a conceptual model-based problem-solving approach. a case study with a student with autism spectrum disorder," *ZDM – Mathematics Education*, vol. 56, no. 6, pp. 1239–1256, 2024.

[34] R. Fernández-Cobos and I. Polo-Blanco, "Using modified schema-based instruction to teach problem-solving to students with autism and language impairments: A single case design study," *International Journal of Developmental Disabilities*, in press.

[35] I. Polo Blanco, M. J. González López, and A. Bruno, "Influencia del contexto en problemas de multiplicación y división: estudio de caso de un alumno con autismo," *Siglo Cero*, vol. 52, no. 1, pp. 59–78, 2021.

[36] J. R. Root, A. Saunders, S. K. Cox, D. Gilley, and A. Clausen, "Teaching word problem solving to students with autism and intellectual disability," *TEACHING Exceptional Children*, vol. 57, no. 1, pp. 44–55, 2024.

[37] J. M. Carroll, "Human-computer interaction: Psychology as a science of design," *Annual Review of Psychology*, vol. 48, no. 4, pp. 61–83, 1997.

[38] A. Dix, J. Finlay, G. D. Abowd, and R. Beale, *Human–Computer Interaction*, 3rd ed. Harlow, England: Pearson Education, 2004.

[39] T. C. Reeves, "The impact of media and technology in schools," *Journal of The Journal of Art and Design Education*, vol. 2, pp. 58–63, 1998.

[40] N. Hollender, C. Hofmann, M. Deneke, and B. Schmitz, "Integrating cognitive load theory and concepts of human–computer interaction," *Computers in human behavior*, vol. 26, no. 6, pp. 1278–1288, 2010.

[41] J. Nielsen, *Usability engineering*. Morgan Kaufmann, 1994.

[42] J. Sweller, "Cognitive load during problem solving: Effects on learning," *Cognitive science*, vol. 12, no. 2, pp. 257–285, 1988.

[43] J. Sweller, J. J. Van Merrienboer, and F. G. Paas, "Cognitive architecture and instructional design," *Educational psychology review*, vol. 10, no. 3, pp. 251–296, 1998.

[44] P. Chandler and J. Sweller, "Cognitive load theory and the format of instruction," *Cognition and instruction*, vol. 8, no. 4, pp. 293–332, 1991.

[45] P. Ayres and J. Sweller, "The split-attention principle in multimedia learning," *The Cambridge handbook of multimedia learning*, vol. 2, pp. 135–146, 2005.

[46] E. Gkintoni, H. Antonopoulou, A. Sortwell, and C. Halkiopoulos, "Challenging cognitive load theory: The role of educational neuroscience and artificial intelligence in redefining learning efficacy," *Brain Sciences*, vol. 15, no. 2, p. 203, 2025.

[47] C. Ware, *Information visualization: perception for design*. Morgan Kaufmann, 2000.

[48] D. Chang, L. Dooley, and J. E. Tuovinen, "Gestalt theory in visual screen design–a new look at an old subject," in *Selected Papers from the 7th World Conference on Computers in Education (WCCE'01), Copenhagen, Computers in Education 2001: Australian Topics, Volume 8*. Melbourne: Australian Computer Society, 2002, pp. 5–12.

[49] E. R. Tufte, *Envisioning Information*. Cheshire, CT: Graphics Press, 1990.

[50] D. A. Norman, *The Design of Everyday Things*, revised and expanded edition ed. New York: Basic Books, 2013.

[51] B. Shneiderman, "Direct manipulation: A step beyond programming languages," *Computer*, vol. 16, no. 08, pp. 57–69, 1983.

[52] C. Laborde, "Integration of technology in the design of geometry tasks with cabri-geometry," *International Journal of Computers for Mathematical Learning*, vol. 6, no. 3, pp. 283–317, 2002.

[53] S. Papert, *Mindstorms: Children, Computers, and Powerful Ideas*. Basic Books, 1980, reprint 2020.

[54] S. J. Hegedus and L. Moreno-Armella, "Intersecting representation and communication infrastructures," *ZDM*, vol. 41, no. 4, pp. 399–412, 2009.

[55] L. Moreno-Armella, S. J. Hegedus, and J. J. Kaput, "From static to dynamic mathematics: Historical and representational perspectives," *Educational Studies in Mathematics*, vol. 68, no. 2, pp. 99–111, 2008.

[56] S. Ainsworth, "Deft: A conceptual framework for considering learning with multiple representations," *Learning and instruction*, vol. 16, no. 3, pp. 183–198, 2006.

[57] N. Pavlov, "User interface for people with autism spectrum disorders," *Journal of Software Engineering and Applications*, vol. 7, pp. 128–134, 2014.

[58] G. Freyhoff, G. Hess, L. Kerr, B. Tronbacke, and K. Van Der Veken, *Make it Simple: European Guidelines for the Production of Easy-to-Read Information for People with Learning Disability for Authors, Ed-*

itors, Information Providers, Translators and Other Interested Persons. ILSMH European Association, 1998.

[59] B. R. Smith, F. Spooner, and C. L. Wood, "Using embedded computer-assisted explicit instruction to teach science to students with autism spectrum disorder," *Research in Autism Spectrum Disorders*, vol. 7, no. 3, pp. 433–443, 2013.

[60] N. Rasche, "Design strategy for the development of applications for autism instruction," Master's thesis, Purdue University, 2013.

[61] Ü. P. Abidoğlu, O. Ertuğruloğlu, and N. Büyükeğilmez, "Importance of computer-aided education for children with autism spectrum disorder (asd)," *Eurasia Journal of Mathematics, Science and Technology Education*, vol. 13, no. 8, pp. 4957–4964, 2017.

[62] S. Van Vaerenbergh, J. A. Reyes Delgado, and J. A. Armesto, "El proyecto MatesGG: Matemáticas con GeoGebra," *Suma*, no. 108, pp. 103–113, 2025.

[63] Y.-M. Kim, "Validation of psychometric research instruments: The case of information science," *Journal of the American Society for Information Science and Technology*, vol. 60, no. 6, pp. 1178–1191, 2009.

[64] D. Collingridge. (2014, September 23) Validating a questionnaire. Retrieved November 25, 2023. [Online]. Available: https://researchmethodscommunity.sagepub.com/blog/validating-a-questionnaire

[65] A. R. C. S. Thomé, T. H. L. Bernardo, P. d. A. Sarmento, J. A. P. d. M. Coelho, and E. M. M. Fedocci, "Checklist validation for use in safe heart surgery," *Revista Gaúcha de Enfermagem*, vol. 43, p. e20220025, 2022.

[66] R. Skjong and B. H. Wentworth, "Expert judgment and risk perception," in *Proceedings of the Eleventh International Offshore and Polar Engineering Conference*. ISOPE, June 2001, pp. 537–544.

[67] J. Escobar-Pérez and Á. Cuervo-Martínez, "Validez de contenido y juicio de expertos: una aproximación a su utilización," *Avances en medición*, vol. 6, no. 1, pp. 27–36, 2008.

[68] M. R. Lynn, "Determination and quantification of content validity," *Nursing research*, vol. 35, no. 6, pp. 382–386, 1986.

[69] D. F. Polit and C. T. Beck, "The content validity index: are you sure you know what's being reported? critique and recommendations," *Research in nursing & health*, vol. 29, no. 5, pp. 489–497, 2006.

[70] C. R. D. Nora, E. Zoboli, and M. M. Vieira, "Validation by experts: importance in translation and adaptation of instruments," *Revista Gaúcha de Enfermagem*, vol. 38, 2018.

[71] N. Elangovan and E. Sundaravel, "Method of preparing a document for survey instrument validation by experts," *MethodsX*, vol. 8, p. 101326, 2021.

[72] L. Schreibman and A. C. Stahmer, "A randomized trial comparison of the effects of verbal and pictorial naturalistic communication strategies on spoken language for young children with autism," *Journal of autism and developmental disorders*, vol. 44, pp. 1244–1251, 2014.

[73] G. T. Baranek, F. J. David, M. D. Poe, W. L. Stone, and L. R. Watson, "Sensory experiences questionnaire: discriminating sensory features in young children with autism, developmental delays, and typical development," *Journal of child Psychology and Psychiatry*, vol. 47, no. 6, pp. 591–601, 2006.

[74] A. Ben-Sasson, E. Gal, R. Fluss, N. Katz-Zetler, and S. A. Cermak, "Update of a meta-analysis of sensory symptoms in asd: A new decade of research," *Journal of autism and developmental disorders*, vol. 49, pp. 4974–4996, 2019.

[75] C. Gómez Casanueva, "Adaptación de materiales del proyecto Mates GG para alumnado con TEA," Master's thesis, Universidad de Cantabria, 2022, [Adaptation of materials from the Mates GG project for students with ASD].

[76] J. J. Randolph, "Free-marginal multirater kappa (multirater k [free]): An alternative to Fleiss' fixed-marginal multirater kappa," in *Learning and Instruction on Multiple Contexts and Settings III: Proceedings of the Fifth Joensuu Symposium on Learning and Instruction*, 2005.

[77] J. L. Fleiss, "Measuring nominal scale agreement among many raters." *Psychological bulletin*, vol. 76, no. 5, p. 378, 1971.

[78] M. J. Warrens, "Inequalities between multi-rater kappas," *Advances in data analysis and classification*, vol. 4, no. 4, pp. 271–286, 2010.

[79] U. Jakobsson and A. Westergren, "Statistical methods for assessing agreement for ordinal data," *Scandinavian journal of caring sciences*, vol. 19, no. 4, pp. 427–431, 2005.

[80] W. S. Ng, P. Watts, Z. Lawson, A. Kemp, and S. Maguire, "Development and validation of a standardized tool for reporting retinal findings in abusive head trauma," *American journal of ophthalmology*, vol. 154, no. 2, pp. 333–339, 2012.

[81] J. M. García-Ceberino, A. Antúnez, S. J. Ibáñez, and S. Feu, "Design and validation of the instrument for the measurement of learning and performance in football," *International Journal of Environmental Research and Public Health*, vol. 17, no. 13, p. 4629, 2020.

[82] C. Pinya-Medina, V. Morcillo-Loro, M. Ferrer-Ribot, and M. d. M. Oliver Barceló, "What competencies should early childhood educators possess for the future? an expert judgement approach," in *Frontiers in Education*, vol. 9. Frontiers Media SA, 2024, p. 1422950.

[83] J. R. Landis and G. G. Koch, "An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers," *Biometrics*, vol. 33, no. 2, pp. 363–374, 1977.