Using Implicit Measures to Assess User Experience in Children: A Case Study on the Application of the Implicit Association Test (IAT)

Eleonora Aida Beccaluva
eleonora.beccaluva@polimi.it
e.beccaluva@campus.unimib.it
Politecnico di Milano, Department of Electronics, Information and Bioengineering
Milan, Italy
University Milano-Bicocca, Department of Psychology
Milan, Italy

Marta Curreri
m.curreri2@campus.unimib.it
University Milano-Bicocca, Department of Psychology
Milan, Italy

Giulia Da Lisca
g.dalisca@campus.unimib.it
University Milano-Bicocca, Department of Psychology
Milan, Italy

Pietro Crovari
pietro.crovari@polimi.it
Politecnico di Milano, Department of Electronics, Information and Bioengineering
Milan, Italy

ABSTRACT

The number of multimodal agents aimed at children with or without neurodevelopmental disorders (NDD) has increased tremendously during the last decade. As this expands, so does research into methods, tools, and metrics that can reliably assess their impact. Traditionally, the majority of UX Research tools have been produced for an adult audience, with fewer tools developed for a younger population. Furthermore, most of these tools use a "direct" method, in which detailed questions are asked directly to the individuals. However, when assessing youngsters, and mainly when direct inquiries are posed, the literature identifies several challenges and pitfalls not usually faced when testing adults. If overlooked, they might lead to biased judgments. This paper proposes a novel approach to UX Evaluation using implicit metrics, which offers the obvious advantage of avoiding direct questions. We investigated the application of the Implicit Association Test (IAT) - one of the most acknowledged tests in psychology to reveal unconscious attitudes, automatic preferences, and hidden biases - to determine whether 60 school-aged children enjoyed a multimodal interface dedicated to language assessment. The results, although preliminary, disclose discrepancies between what children state directly and what the test detects. With our work, we want to offer two contributions. The first, technical, describes both the logic as well as the tool we used to develop the IAT. The second, methodological, offers preliminary but exciting evidence to support the usefulness of implicit measures, and the IAT, in this field.

ICMI '23 Companion, October 09–13, 2023, Paris, France

ACM ISBN 979-8-4007-0321-8/23/10.
https://doi.org/10.1145/3610661.3617513

CCS CONCEPTS

• **Human-centered computing** → **Human computer interaction (HCI)**; **HCI theory, concepts and models**; **Empirical studies in HCI**.

KEYWORDS

User Experience, Implicit Measures, Children, IAT, Language Assessment

ACM Reference Format:
Eleonora Aida Beccaluva, Marta Curreri, Giulia Da Lisca, and Pietro Crovari. 2023. Using Implicit Measures to Assess User Experience in Children: A Case Study on the Application of the Implicit Association Test (IAT). In *INTERNATIONAL CONFERENCE ON MULTIMODAL INTERACTION (ICMI '23 Companion), October 09–13, 2023, Paris, France.* ACM, New York, NY, USA, 10 pages. https://doi.org/10.1145/3610661.3617513

1 INTRODUCTION

The past decade has seen a remarkable transformation in technologies dedicated to children and able to address the broadest range of purposes (education, entertainment, therapy). Not only the technologies are variegate (e.g., from virtual reality to tablet applications to tangibles interfaces, etc.), but also the modalities of interaction between the users and the technology itself are constantly evolving, making it even more complex to correctly and, more importantly, realistically assess its actual impact.

One of the most interesting and promising examples of new technology in this field is Conversational Agents (CA), aka software that allows access to information and services by the use of natural language communication [18] and can provide a more accessible User Experience (UX) to subjects. This "naturalization" of the interaction and, as a consequence, of the UX marks a significant step forward in the Human-Computer Interaction research arena and paves the way for a closer analysis of the resulting perks to younger generations.

The benefits of CA are even more evident when considering kids who do not exhibit typical development but, in contrast, an atypical one, such as children with Neurodevelopmental Disorders (NDD). The DSM V - the Diagnostic and Statistical Manual of Mental Disorders: V edition - describes NDD as a category of developmental disorders defined by deficiencies in cognitive, social, and communication domains [4] that embrace intellectual developmental disorder, global developmental delay, communication disorders, autism spectrum disorder, attention-deficit/hyperactivity disorder (ADHD), neurodevelopmental motor disorders, and specific learning disorders[4]. Unsurprisingly, a considerable part of the research interest and literature has been directed toward a better understanding of the role of these agents as a *support tool* for such clinical cohort [14].

CAs, especially those that also present with multimodal interactions (namely the variety of input and output prompts beyond the voice that can be associated with the specific interaction [27]), can adjust the difficulty level, content, information access, and pace of activities to provide a better-customized and accessible experience [19, 42, 45] that could be an invaluable asset. However, the broad spectrum of special needs to be addressed and the very multidisciplinary aspects to consider both when designing and evaluating this target population represent an *unprecedented and very intricate challenge* for the scientific community. It is imperative that advances in how to engineer these technologies be accompanied by equal advances in evaluating and understanding them and vice versa to fuel a virtuous circle and ensure real progress in the field.

However, how do we estimate the real impact of this technology on different populations? In adults, UX research is the subject of in-depth studies, publications, and guidelines [24], but designing and evaluating agents for children, and more specifically for typical and atypical children, is an open topic that needs further investigations [44].

We believe that a realistic and reliable UX assessment is fundamental and salient for three main reasons: (i) to prompt the design and the development of multimodal agents, (ii) to advance research in this field, and (iii) to provide a better understanding of the role of these agents in typical and atypical populations [31].

Today, there is still limited reliable knowledge about the usability, effectiveness, and perception of conversational agents in children and, even more crucial, in children with NDD [14]. Children's experience should be carefully examined differently from what is typically the practice for adults[35], and assessing children with or without NDD poses even more challenges and requires distinct adjustments according to age groups, necessities, skills, and interaction with technology. Most UX Research tools have traditionally been built for an adult audience, with fewer tools developed for younger people.

Furthermore, most of these tools employ a "direct" approach, in which comprehensive queries are directed directly at the participant's [51]. However, when assessing children, mainly when direct questions are presented, the research outlines various obstacles and hazards that are only sometimes encountered when testing adults. If ignored, they may result in biased conclusions. In our opinion, a necessary step that needs to be taken is the expansion of the tools available to perform UX testing in children that consider this population's psychological characteristics. This deserves, on the one hand, some reflections on the limitations of the methodologies currently in use and reported in the literature and, on the other, a systematic search for new methods or tools to improve and enhance them [3]. Also outside the proper UX domain.

To fill this research vacuum, the present paper reports a case study and suggests an unusual approach to UX Evaluation based on implicit measures, which has the obvious advantage of avoiding direct inquiries. We investigated whether school-aged children enjoyed a multimodal interface dedicated to language assessment by using the Implicit Association Test (IAT) [26], one of the most widely used tests in psychology for revealing unconscious attitudes, automatic preferences, and hidden biases [22, 28, 38]. The study involved 60 children aged 6 to 7 y.o. who interacted with a conversational agent in two modalities, one multimodal and one non-multimodal, desktop-based. One researcher later explicitly asked them to state which one they preferred. Subsequently, all the subjects underwent an IAT test. Although early, the findings of this case study show differences between what children say directly and what the test discovers. With our work, we want to offer two contributions. The first, technical, describes both the logic and the tool we used to develop the IAT. The second, methodological, offers preliminary but exciting evidence to support the usefulness of implicit measures and the IAT in this field. The rest of the paper is organized as follows: Section 2 provides an overview of the theoretical background, with sub-sections on UX assessment in children and the main pitfalls and risks. Section 3 describes the Implicit Measures and Implicit Association Test (IAT) and its applications to children. Section 4 presents the case study of MARS, the UX, and the tasks. Section 5 focuses on the empirical case study. Section 6 presents the preliminary results. Section 7 reports the discussion and the limitations of the work. Section 8 draws conclusions and sheds light on the directions of our future work.

2 THEORETICAL BACKGROUND

2.1 UX Experience Assessment in Children

There are many different techniques and tools that have been developed to study User Experience and interaction: quantitative ones, like user surveys and A/B testing; qualitative ones, such as user interviews and focus groups; and mixed ones, such as usability testing [53]. Trying out a product or an app is a complex experience that can be broken down into different constructs such as desirability, aesthetic appeal, accessibility, or usability, to name a few [40]. When it comes to applying these techniques to children, the number of generic publications on what methodologies to use is reduced. Among the first to cover this issue were Sherwin and Nielsen. In their usability studies (2001, 2010, 2018)[48], the authors tested 125 children aged 3 to 12 years old by asking them to perform a range of tasks they had prepared while applying the 'think out loud' paradigm. From this series of experiments, Sherwin and Nielsen defined a series of guidelines that should be applied when designing for a young audience. They described some skills children have developed as using tablets and phones has become more widespread. For example, kids have higher expectations of interactivity and are more willing to troubleshoot than adults if something is not working. Despite the differences, many of the User Interface conventions defined for adults were automatically extended also for children. They had been employed when designing technology, from web apps or Conversational Agents. Other studies navigated children's acceptance and enjoyment of different applications through questionnaires such as the User Experience Questionnaires (UEQ)[17, 32, 33], natural observation and interviews [29]. UX research is not only applicable to the digital world.

In recent years, there has been a return to tangible manipulation and physical interactions as a new way to approach digital data in a younger population. This field is labeled as 'Tangible Embedded Interaction' (TEI) and focuses on mechanical devices to allow for a more natural and less cognition-driven approach. Vanden Abeele, Zaman, and De Groff (2012)[52] utilize laddering in their case study to evaluate the interaction with a 'cuddly toy interface' with preschoolers. UX Laddering is an adapted interview method and data analysis process for investigating the User Experience. The goal of laddering is to identify and understand the linkages between crucial perceptual elements across the range of attributes, consequences, and values. The target audience should include children aged 5 and older, as younger preschoolers may have more difficulty with such tasks. The studies presented up to now focus on discoverability and learnability of the technology at the first and novel interaction - aka the first time for the first time. The experiment by Vissers, De Bot, and Zaman (2013)[54] explores long-term User Experience by adapting the UX Curve. This method relies on the user to report retrospectively why their experience with a specific product changed within a timespan. All the methodologies mentioned above rely on explicit attitudes: the users, in these cases children, are directly asked what their beliefs are regarding the products, and their overt behaviors are observed and recorded. As we will further explore in the following paragraphs, these methodologies may not always be the optimal way of investigating the children's true feelings because of this population's greater tendency to give in to social desirability and be more susceptible than older people.

2.2 Testing children: main pitfalls and risks

When testing children, it is imperative to carefully tailor the research protocol around their peculiar characteristics. Children behave differently from adults in many critical cognitive skills, such as attention, self-regulation, and effortful control, which are crucial during testing sessions. Moreover, children tend to engage in different social dynamics [5]: they are more motivated to please adults, and their ability to adjust to their surroundings and new people may change in the course of the session [49]. They are more inclined to tell "white lies", aka small, harmless lies or untruths that are often told to avoid hurting someone's feelings, maintain social harmony, or prevent unnecessary conflict. These lies are generally well-intentioned and not intended to cause harm [12], but their impact on an evaluation might significantly alternate the outcomes.

Part of the research involving children deals with their opinion on the experiences they were subjected to or their recollection of life events. As mentioned before, children could adjust their opinions to please adults [49]. The ability to recollect events has been a topic of interest for the past three decades, and most research has been produced in the legal context, given the importance of children's testimonies when it comes to certain crimes. However, we can reasonably assume that the dynamics emerging from a legal interview can affect any situation where a child is asked to report on something they did or felt. For this reason, experimental protocols involving an interview with the subjects to collect data must proceed with caution.

In general, it has been proved that young children are most susceptible to misleading suggestions [1, 9, 10]. Often these suggestions do not carry any malevolent intent and are simply the result of interviewer bias. Like all individuals, researchers tend to exhibit biases towards information that confirms their beliefs and reject contradictory evidence[20]. Biased interviewers are more likely to ask specific questions instead of more "open-ended" ones. However, because children give less accurate answers to specific questions [11], this strategy soon becomes problematic.

Testing children requires consideration not only of their cognitive, emotional, and social abilities but also how our biases influence our communication with them. For example, the question *"Was the game you just played nice?"* already includes the adjective *"nice"*, which might lead the child to assume that the person who asked the question thinks it was nice and, not wanting to contradict an adult, prompts the child to answer more likely positively than negatively. A less biased way to ask this question would have been, *"What did you think of the game you just played?"*. There are no triggers for judgment in this formulation. However, being a much broader question, it requires considerably more time to narrow the focus of the answer to what is of interest, as well as specific training of the researcher on how to conduct such unbiased questions. However, is it possible to think of any other solution that avoids the root of the problem? The answer to this question comes from studies in psychology that have delved into the use of *implicit measures* to probe people's opinions and biases without ever explicitly asking.

3 IMPLICIT MEASURES AND IMPLICIT ASSOCIATION TEST (IAT)

In psychology, implicit measures refer to assessment methods that aim to capture unconscious or automatic cognitive processes and attitudes without requiring direct or conscious self-report. These measures are designed to reveal implicit biases, attitudes, or associations that individuals might not be aware of or may be hesitant to acknowledge in explicit or conscious measures [21].

Even though we have seen a surge in interest in the implicit-explicit dichotomy in the last decade, its origins can be traced back to the start of the 20th century. One of the most notable examples of this distinction can be found in the "Treatise on Persuasion" by Hovland, Janis, and Kelley (1953) [25]. In their work, the authors defined *attitudes* as "implicit responses," in contrast with opinions, which were "verbal answers that one covertly expresses to (oneself)." When applied to measures, the term "implicit" most commonly refers to the fact that people are unaware of what the measure is assessing, contrary to the case of an explicit measure, where people are fully aware of what is being requested. Indirect measures seek to overcome the limitations of the self-report methodological techniques [16, 36, 55] which could be grouped in (i) self-deception (e.g., false belief and a contradictory unconscious genuine belief) (ii) other deception, such as social desirability and (iii) linguistic bias [2]. Included in this category are a wide variety of methods such as the Thematic Apperception Test [34], the Implicit Relational

Assessment Procedure (IRAP), various reaction time-based tasks designed to assess implicit attitudes and stereotypes[43] and the physiological measures.

The concept of implicit measures gained even more scientific consensus and significant attention with the development of the Implicit Association Test (IAT) introduced by Anthony G. Greenwald, Debbie E. McGhee, and Jordan L. K. Schwartz in 1998. The Implicit Association Test [21] is the most widely used measure of implicit cognition [41]. The test measures the strength of a person's automatic association between two concepts and evaluation attributes by instructing participants to sort examples from two pairs of concepts using just two response options. The idea behind the test is that the sorting task should be easier (and thus, faster) when two concepts are strongly associated. Stimulus items can be presented as words, pictures, sounds, or a combination of the three. The first field of application of the IAT was the implicit attitudes towards race [21]. Since then, it has been used in several disciplines, including social and cognitive psychology, clinical psychology, developmental psychology, neuroscience, market research, and health psychology; see Schnabel et al. for a review on the topic [47]. In the field of social science, in particular, it has been used to investigate implicit attitudes when it comes to gender identity, race, and self-esteem.

Given its nature as a measure of implicit constructs, the IAT results can often differ from the self-reports. Regarding the latter, the individual is probably unaware of what is being measured or could be aware of it. However, they could answer differently if the association does not conform to their belief system or if concerns regarding the acceptability of such a response arise [39]. The IAT neutralizes these processes, giving us what could be considered the "real" answer.

3.1 IAT with Children

As discussed before, measuring children's attitudes is a difficult task as children show a tendency to respond in socially desirable ways [49] or follow self-preservation biases [23] when tested with an explicit measure. The Implicit Association Test represents an advantage in this scenario, as it permits an assessment of spontaneous social and non-social attitudes [50].

Although IAT is widespread among adults, little has been done regarding its application to children. Scientific publications are, to the best of our knowledge, still limited. On the one hand, it is understandable how an instrument requiring sustained attention, working memory and inhibitory skills, such as the IAT, should be carefully evaluated for very young children (under 5 years of age). On the other hand, however, it could be extremely interesting for older children. The test has been adapted for children as young as six years old [6], with promising results for children around the age of three [50]. Thanks to its adaptability, the IAT has been used to investigate children's attitudes in different domains regarding the world around them and their inner selves.

Although studies on children can usually take some liberties in the design of the procedure, such as the number of trials, modality of stimulus presentation, scoring algorithm, and response type, the mechanism underlying the test stays the same.

3.2 IAT In UX Research

When we talk about product evaluation, we have to consider functionality and usability; however, with the increasing use of technology in our field of research, we cannot underestimate the concept of "experience" [37]. With this term, we describe the level of engaging interaction with a device [2].

Our cognitive evaluations and implicit attitudes are triggered when we interact with an object. This means that self-reports and other questionnaires may only partially capture the experience's effect on the individual. As we have discussed, these issues have been long studied in the psychological literature, and implicit measures such as the IAT are well-established solutions. However, the User Experience is still generally evaluated through self-report questionnaires. By asking to report the interaction's feelings, emotions, affective states, and aesthetical experience, it is taken for granted that the user is perfectly aware of their feelings and emotion and can identify what triggers them. Another erroneous assumption is based on the belief that our behavior is determined by implicit assumption only when we interact with other humans or human-like devices, or the interaction has a high degree of social or emotional valence [2].

The IAT has already been used in marketing, with a variation called Brand Association Reaction Time Test (BARTT) [30], suggesting that multiple kinds of stimuli and situations can trigger implicit assumptions. Another exciting factor that points to the IAT as a good tool for measuring the impact of UX is how easy its protocol is: the usual experimental paradigms are not appealing for the industrial market (the foremost field in which UX plays an important role), given they often take a long time to perform [2]. Since most of the research regarding UX comes from the private market, establishing the IAT as a valid test in this field would benefit this area of study, even beyond the needs of marketing and product development. Even with all of its benefits, the use of IAT in UX is still a relatively new application, as the first pioneering studies have been emerging only in the last couple of years [2, 30, 56]

4 MARS: A CASE STUDY

In this paper, we propose the application of the Implicit Association Test (IAT) in the context of User Experience (UX) Research to a sample of n=60 schooled 6/7 y.o. children.

Our study aimed to determine whether using a multimodal or non-multimodal agent to administer repetitive and unappealing tasks to children is optimal. We present MARS, a multimodal agent designed to distinguish children with Neurodevelopmental Disorders (NDD) and language disorders from typically developed ones through rhythmic tasks.

In our protocol, we proposed that children perform some tasks through the multimodal version of MARS and others through the non-multimodal version. Through explicit and implicit tests, we assessed their preference for one version over the other. In this chapter, we will introduce MARS, provide an overview of the IAT and the tool used to create it and discuss the study's explorative results.

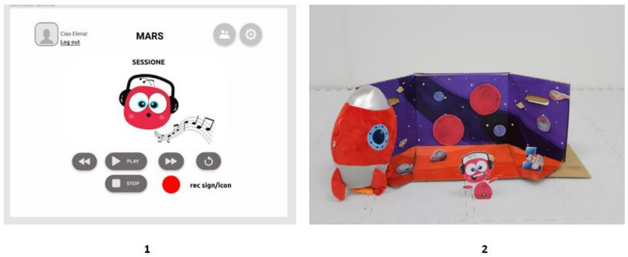

Figure 1: MARS. Image 1 is the "running activity screenshot" where the play/pause/stop/continue buttons are placed, the logging id, some written description of all functionalities, and the face of a funny red Alien that is the MARS character. Image 2 is the MARS landscape.

4.1 Multimodal Conversational Experience with MARS

MARS is a multimodal conversational agent created to deliver rhythmic tasks and automatically record children's vocal productions [8]. We inserted gamification elements in the User Experience because MARS administers repetitious and unappealing activities to a child's eyes.

The application was developed for researchers as a Progressive Web App (PWA) that can be easily used on laptops and mobile devices (smartphones, tablets, etc.), acting as a powerful audio collection and storage device. As a general premise, we assumed that the children could feel uncomfortable repeating rhythmic pairs of "ba-ba" and "bu-bu" in front of the investigator. This discomfort could significantly impact children's vocal production (inducing them to speak in a lower or trembling voice) and, thus, the tool's effectiveness.

We developed a Multimodal Interface (Conversational and Tangible) and a tale around a fun, charming, and friendly character to avoid such risk. The story follows MARS, a little alien journeying through space as he makes his way back home to his planet. However, an unexpected twist occurs when his rocket malfunctions, causing him to crash-land on an unfamiliar planet. To engage the children, we immerse them in the narrative, allowing them to assume the role of the captain of another spaceship, with the researcher serving as their deputy pilot. Their mission is to assist MARS by providing him with the necessary objects to repair his rocket. Each completed task earns the child one of the elements essential for fixing the rocket. We use picture cards and a plush rocket to depict these items, adding a tactile and visual dimension to the experience. For audio communication, we incorporated a space-themed metaphor, wherein both MARS and the astronauts constantly interact through headphones. This setup also requires the children to wear headphones with microphones, allowing seamless communication during the session.

The conversational experience is designed to onboard and engage users, telling them the story in which the session is set. By combining storytelling, tangible elements, and agent-guided interactions, we have created an engaging and interactive experience that sparks the children's imagination and involvement in MARS' interstellar adventure. We built a spatial setting with several components that the kid may control to immerse them in the story better: a plush toy rocket, a space-themed background, the MARS figure, and meteorites or moveable satellites, see figure 1. MARS and the Tangible Interfaces were designed using a set of general principles derived from various sources, including theories underlying gamification and TUIs (Tangible User Interfaces)for children with NDD [7] and concepts emerging from the study of clinical approaches to therapy.

4.2 IAT: Development and Tools

We will now describe some basic concepts of how the IAT works to provide the reader with a better understanding of the logic behind the test. As stated previously, the IAT measures the strength of the association between concepts based on the latency in reaction times. The IAT typically presents participants with a series of trials where they must rapidly categorize stimuli (e.g., words or images) into different categories on a computer screen by pressing specific keys (e.g., "e" for the left category and "i" for the correct category). For example, participants may be asked to categorize words related to either positive or negative concepts (e.g., "joy," "love" for positive, and "pain," "hate" for negative) and associate them with specific groups or social categories (e.g., black and white faces).

The test measures the speed and accuracy of participants' responses, capturing the strength of automatic associations between the categories. The underlying assumption is that individuals with implicit solid biases will respond more quickly when positive words are paired with the favored social group and when negative words are paired with the non-favored group. Conversely, individuals with weaker biases may struggle more when positive words are paired with the non-favored group and when negative words are paired with the favored group.

Usually, the IAT consists of five main parts or blocks:

In each block, the subjects are exposed to a random combination of stimuli belonging to categories. Usually, from 4 to 6 stimuli for each category. The IAT score is based on the average time participants took to sort words in the critical blocks, the third compared to the fifth. The test measures implicit preferences by observing how quickly participants associate concepts with specific evaluations. The IAT presented in this study follows Carpenter's model (2019) [13] that suggested a 7-block version, in which blocks 3, 4, 6, and 7 are the critical blocks.

In our version, we favored images over writing, except for category labels. We used images of the MARS multimodal version, such as images of the landscape toy rocket or character and images of a computer, keyboard, or headphones for the non-multimodal category, see figure 2.

Since our goal was to see which of the two versions of the application (multimodal or MARS vs. non-multimodal or computer) the children liked best, we used different images for the two categories. For the Multimodal category, the images of the landscape toy rocket or character, while for the non-multimodal category, the images of a computer, keyboard, or headphones. While investigating the enjoyment of MARS, we used the labels "Happy" and "Sad" and images of emoticons that prototypically represented the two emotions. We used 4 images for each of the 4 categories (happy/sad and

Multimodal/non-multimodal) category for a total of 16 images. See figure 2 for all images we used in the IAT.

Our IAT consists of seven blocks:

- **Block 1**: in the first block, participants sort images related to the concept (e.g., character, headphones) into categories based on their position (left or right).
- **Block 2**: the second block involves sorting emoticons related to feeling (e.g., happy, sad) into the same categories as in the first part.
- **Block 3**: in the third block, concept and feelings images are combined, and participants sort them accordingly. The categories on the left may be Computer/Happy, and on the right, MARS/Sad.
- **Block 4**: in the fourth, the categories on the left may be MARS/Happy, and on the right, Computer/Sad.
- **Block 5**: the fifth block switches the placement of the concepts from left to right or vice versa, with more trials to minimize the effects of the practice. So if the category "MARS" was on the left, now it is on the right.
- **Block 6**: in the sixth block, concept and feelings images are again combined, but the categories on the feeling are swapped, so now the left may be Computer/sad, and on the right, MARS/happy.
- **Block 7**: the final block, also called critical, combines the categories in a way opposite to the previous combination.

Data collection occurred during blocks 3, 4, 6, and 7. The children were asked to categorize Mars-related and emotion-related images during these blocks.

On the screen, feeling categories (happy/ sad) and concepts (Mars/computer) were randomly associated in the first two blocks and oppositely in the next two, as shown in figure 3. For example, if in blocks 3 and 4 the associations were Mars/ sad and computer/happy, in blocks 6 and 7, the child had to categorize by Mars/happy and computer/ sad.

To develop our IAT, we use a specific tool available for the online survey software Qualtrics (2005) [15]. Qualtrics is a powerful and widely used web-based survey software platform that allows researchers, businesses, and organizations to design, distribute, and analyze online surveys, polls, and questionnaires. It offers a comprehensive suite of data collection and analysis tools, making it a popular choice for academic research, market research, customer experience evaluation, and employee feedback assessments.

We used an extension called iatgen.org [13]. The iatgen tool serves as an adjunct to the methods paper and streamlines the entire procedure without the need for manual code editing. With iatgen, we effortlessly customize our IAT through a web application or a downloadable R package. The software then generates a ready-to-run Qualtrics Survey File (QSF) incorporating the IAT. It is also possible to further tailor the survey to their preferences, such as adding explicit measures, randomization, or multiple IATs.

Additionally, iatgen offers the flexibility to customize existing HTML/JavaScript code, which can be seamlessly integrated into a Qualtrics survey or template. Once the research is completed, iatgen provides a comprehensive set of analysis tools, conducts data reduction, and offers diagnostics, including assessments of internal consistency and error rates. Furthermore, the software allows for exporting clean data for further analysis.

4.3 MARS: Study goals and Ethical Considerations

Given the complexity of implementing MARS with all the multimodal components described above (agent, digital and physical), we questioned the added value these could bring to the User Experience and, thus, whether they were necessary. On the one hand, multimodality could be a key element in keeping children engaged in the task and reducing anxiety and stress levels. Conversely, it makes MARS hardly scalable and replicable, especially in its tangible part. We hypothesize that the multimodal elements make the activity much more appealing to the children's eyes, improving their performance and reducing potential misleading in the data collected of children's productions.

Partecipants			
Gender	n	Age	
		Mean	SD
Total	60	6.25	0.47
Male	33	6.37	0.39
Female	27	6.23	0.51

Table 1: Partecipants Demographical Information

The study design aimed to gather data on the children's preferences and attitudes toward the two versions of MARS, both explicitly through direct questioning and implicitly through the IAT test. The research questions can be abstracted in the following:

- Do children really like MARS multimodality? Or do they prefer a non-multimodal interaction?
- Can we compare implicit and explicit measures?

All research participants and their families were previously advised and recruited voluntarily. Before starting the study, parents or legal tutors provided informed permission, including information about the study procedures, aims, and data treatment. The Ethical Committees of the University of Milano-Bicocca approved the study protocol and authorized its execution, protocol number 588/2021. Security, privacy, and confidentiality problems were managed in accordance with Article 13 of the GDPR (General Data Protection Regulation), EU Regulation No. 679/2016 of April 27, 2016, and the European Data Protection Supervisor's instructions (EDPS).

4.4 Participants

Sixty (N= 60) children were included in the study, with mean age = 6.25, SD = 0.47. Thirty-three kids were male (mean age = 6.37, SD = 0.39), and twenty-seven were female (mean age = 6.23, SD = 0.51), see table 1. All participants were attending first grade and were tested in the second four-month period. Moreover, we previously tested the children for IQ using Raven's Progressive Matrices [46] to ensure that cognitive profiles were comparable and the only differences were in language proficiency. Participant recruitment was conducted at a Milan (Italy) metropolitan public school. Study subjects had age-appropriate nonverbal IQ, no auditory or visual deficiencies, and normal or appropriately corrected eyesight.

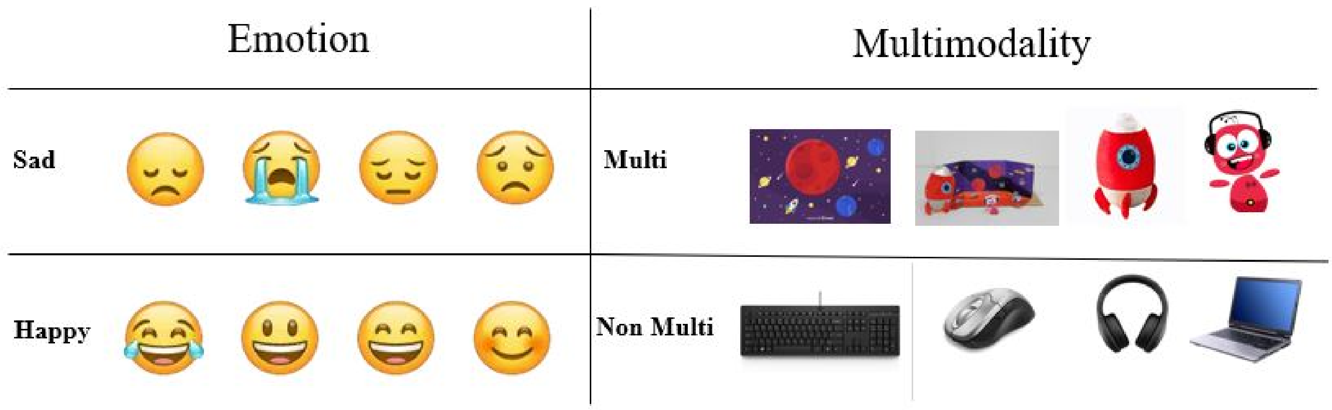

Figure 2: Images used for each category of the IAT

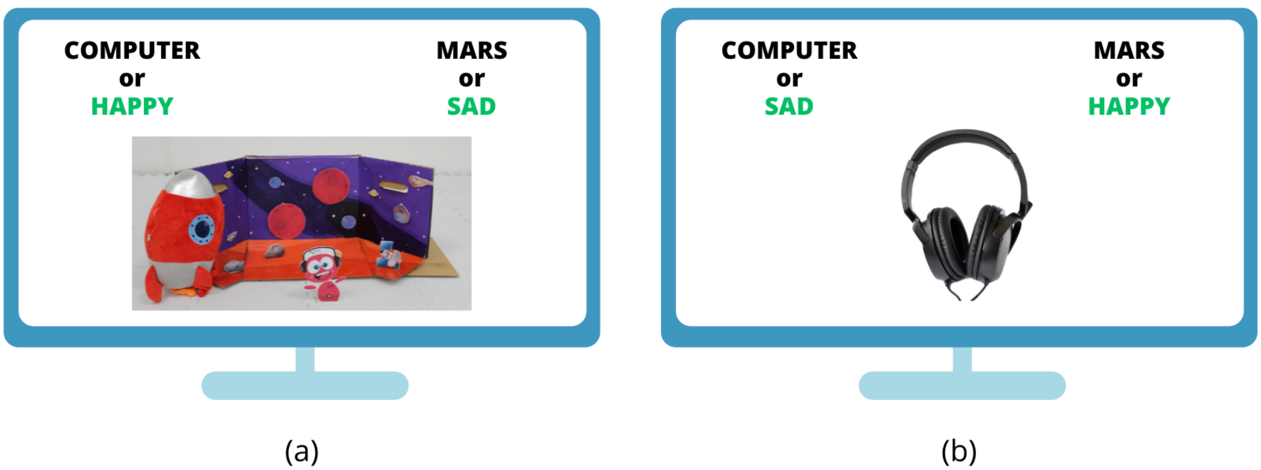

Figure 3: Example of the IAT screen presenting an image linked with the gamified version of MARS.

4.5 Setting

The study single-session was conducted in a designated space at the school. We particularly requested a quiet space to reduce ambient noise during the session. To prevent distractions, the furnishing was kept to a minimum: the room contained a few seats, a table where we set MARS, and a laptop with a wireless mouse at a distance. The children sat in front of the app landscape setting, close to one of the researchers, with the computer in front of them. The kids were asked to put on a headset with a microphone and interact with the MARS agent. The session lasted around 15 minutes, and at the end, each participant received a toy golden medal and a certificate of achievement for their efforts.

4.6 Procedure

In the study, each child participated in both versions of MARS, and the order of exposure to the two versions was randomized to avoid any potential biases. Between the trials of MARS, the children were given a categorization task involving sound stimuli on the computer. This task served as a short break between the two MARS experiences. The session lasted about 20 min. Following completing both versions of MARS, the children were asked a direct question with a facial prompt, inquiring which of the two modes they enjoyed the most, *"Do you prefer MARS or the PC-only version?"*. After listening to and noting the children's initial responses, the researcher probed further by asking them again if they were absolutely sure about their choice, *"Are you sure?"*.

Subsequently, the children underwent the Implicit Association Test (IAT) to assess implicit biases or attitudes. The IAT was conducted after the MARS tasks and preference questioning to examine potential associations between the different modes and any implicit preferences that might be present. While doing the IAT, the children were sitting in front of the PC and using the response keys "e" and "i" on the keyboard according to the task. The test lasted approximately 15 min.

4.7 Results

4.7.1 Explicit measures: direct questions. The survey responses showed that 43 children (71.67% of the participants) initially preferred MARS in the multimodal version, while 17 children (23.33%) preferred the non-multimodal version. However, after the researcher's follow-up question asking if they were sure about their choice, the number of children who preferred the multimodal version increased to 49 (81.67%), and the number of those who preferred the non-multimodal experience decreased to 11 (18.33%), see table 2. Interestingly, six children (10% of the participants) changed their statements immediately after the simple question, *"Are you sure?"*.

	Explicit Measure		
Implicit Measure	0	1	Total
0	2	8	10
1	9	43	50
Total	11	49	60

Table 2: Number of implicit and explicit answers to the question "Do you prefer MARS or the Computer version? In the table, 0 = "Computer" and 1 = "MARS".

4.7.2 Implicit Measures: IAT. The interpretation of the Implicit Association Test (IAT) score involves understanding the strength of implicit biases or associations measured by the test. The IAT score is typically based on the difference in response times between two critical blocks of the test, and the following 'block score' is calculated: Mean Response Time (Mean RT) of target 1 with attribute 1 is calculated. Mean RT of target 1 with attribute 2 is calculated. The difference in mean RTs of both blocks is divided by the inclusive standard deviation (SD). A D-score expresses the results.

The descriptive statistical analysis indicates a D-score mean = 0.08, SD = 0.5, see table 3. The D-score indicates the difference in the implicit preference between the two stimuli (Multimodal vs. non Multimodal), with a positive score indicating a preference for multimodality (a negative score would have indicated a preference toward non-multimodality) This result suggests a slight preference for the gamified condition. Nonetheless, the t-test t(60)= 1.19 p= 0.24 returns a non-significant value. Cohen's d also indicates a small effect size (d= 0.15), and reliability is very high, reaching 90%. So our data can not be interpretable as conclusive.

4.7.3 Implicit vs Explicit Measures. The second aim of our study was to compare the explicit and implicit evaluations of the children between the two conditions. In the contingency table 2, a representation of the number of children that changed their minds with the second question is depicted.

Most children (N=49) stated in the second part of the explicit evaluation that the gamified version was the better one, and within this group, 43 children were consistent with the answer given for question 1. Of the 11 children that preferred the PC-only version, only 2 were consistent between the two questions.

	Implicit Preference	
D-score	mean = 0.08	s.d. = 0.5
T-test	t(60) = 1.19	p = 0.24
Effect Size	d = 0.15	

Table 3: Difference in the implicit preference between the two stimuli (MARS vs Computer). A positive score indicates a preference for MARS.

To verify if the two explicit answers predicted the performance on the IAT, two independent sample t-tests were performed. The results were non-significant in both conditions with t(question1) = -1.14, p=0.26, and t(question2) = -1.45, p=0.15. This indicates that the response to the explicit questions was non-predicting the implicit preference evaluated with the IAT and highlights the importance of evaluating children with indirect measures to obtain reliable data. To further confirm the results obtained, two linear regressions were performed. For the first question, the t = -0.29, p = 0.77, R2 = 0.02; for the second question t = -0.78, p = 0.44, R2 = 0.04, see table 4. The p-values and the R2 values in the linear regression confirmed what was found with the Student's t-tests: the explicit questions are not predictive and explain very little of the performance on the IAT test, which reflects the implicit preference of the children.

	Predicting the implicit answer from the explicit one	
	Question 1	Question 2
Independent sample t-test	t = -1.14, p = 0.26	t = -1.45, p = 0.15
Linear regression	t = - 0.29, p = 0.77, R2 = 0.02	t = -0.78, p = 0.44, R2 = 0.04

Table 4: Statistical analysis to verify whether the explicit response to the two questions predicts the performance on the IAT. Question 1 is "Do you prefer MARS or the Computer?". Question 2 is "Are you sure?".

5 DISCUSSION

This paper presented a case study on assessing User Experience (UX) in children, mainly focusing on a multimodal agent called MARS, designed to identify children with Neurodevelopmental Disorders (NDD) and language disorders. The study aimed to compare explicit and implicit measures to gain insights into children's preferences and attitudes toward multimodal interaction. We utilized classical explicit measures and a more sophisticated and novel implicit measure known as the Implicit Association Test (IAT) [21].

The IAT is a well-known test in psychology that helps explore unconscious biases and attitudes between concepts. We presented both the theoretical framework and the practical tools for implementing an IAT applied to a multimodal agent called MARS, dedicated to identifying children with Neurodevelopmental Disorders (NDD) and language disorders. As we argued, assessing children, mainly when using direct questions, presents various obstacles and challenges that are not often encountered when testing adults. If ignored, they may result in biased conclusions.

Based on this hypothesis, our experimental study aimed to answer the following research questions: Do children like MARS multimodality? Or do they prefer a non-multimodal interaction? Can we compare implicit and explicit measures? The study found that using direct questions as explicit measures could lead to biased conclusions, especially when dealing with children as reported by literature [5, 12, 49].

A significant percentage (10%) of children changed their responses when given a simple direct question, highlighting the importance of careful probing and question design. When given an explicit measurement by means of a very simple direct question, "Are you sure?", six children changed their version immediately. This supports our initial hypotheses and highlights the risks of using direct, explicit measurement without adequate preparation to avoid potentially biased questions. Furthermore, the data from the explicit questions were neither correlated nor predictive of those from the implicit measurements. We can make a bold assumption and state that if the researcher had asked the question in a more aggressive tone by also adding a reinforcement such as "Are you really, really sure?" and asked the question a third time, the number of children who would have changed their minds might be even higher. These observations confirm the importance of using measures that can replace or supplement explicit ones.

The IAT results supported the findings from explicit measures, indicating that children favored the multimodal interaction with MARS over a non-multimodal one. The IAT provided additional insights into the strength and direction of implicit associations between concepts, further confirming the preferences observed in explicit measures. However, we acknowledged that interpreting IAT scores requires caution due to individual variability and potential measurement limitations. Therefore, we recommended using the IAT in combination with other methods and measures to obtain a more comprehensive understanding of implicit biases and attitudes.

Notwithstanding the inherent limitations delineated above, we assert that the application of this methodology holds substantial significance within the domain. Notably, this tool circumvents the prerequisite for reading proficiency, emphasizing instead the imperative of visual cognitive acumen. This attribute renders it enjoyable to individuals in their earlier developmental stages and those with reading impediments. Moreover, the mode of response hinges upon motor dexterity, entailing the pressing of specific key pairs. Conventionally, this involves the "i" and "e" keys situated on the standard keyboard. It is, however, imperative to clarify that this convention merely serves as a reference, obviating any demand for the participants, especially children, to discern these characters on the keyboard or comprehend them textually. Their task merely entails the pressure of the designated keys, which can be easily substituted by alternative inputs or a specialized button interface. Adopting a motor-driven modality also extends the applicability of the Implicit Association Test (IAT) to encompass individuals grappling with Language Disorders, as evinced in our study, and those encountering challenges in phono articulatory functions. However, our firm conviction rests upon the necessity for a broadened scholarly discourse within this domain. The existing body of research concerning the pediatric cohort of IAT application remains circumscribed. In light of the inherent attributes characterizing the administered test, particularly its active engagement of sustained attention, working memory functions, and the capacity for inhibitory control, a methodical investigation into its compatibility with considerably severe phenotypes of Neurodevelopmental Disorders (NDD) is undeniably needed.

6 CONCLUSION

Although early, the findings of this case study show differences between what children say directly and what the test discovers. With our work, we want to offer two contributions. The first, technical, describes both the logic as well as the tool we used to develop the IAT. Using iatgen simplifies and automates the IAT implementation process, providing researchers with an efficient and user-friendly solution for conducting and analyzing Implicit Association Tests. The second, methodological, offers preliminary but exciting evidence to support the usefulness of implicit measures and the IAT, in this field.

In summary, the study highlights the significance of employing implicit measures when assessing UX in children and underscores the importance of careful questioning to avoid potential biases in research findings. The paper contributes to the existing literature, reinforcing the importance of careful assessment techniques when studying children's preferences and attitudes. By incorporating implicit measures like the IAT, researchers can gain valuable insights beyond what explicit measures alone may offer.

ACKNOWLEDGMENTS

This research has been partially funded by Ico Falck Foundation – Project "Agenti multimodali intelligenti per interventi valutativi e terapeutici nei bambini con disturbi evolutivi del linguaggio" ("Multimodal intelligent agents in evaluation and therapeutic interventions for children with developmental language disorders")

REFERENCES

[1] Jennifer K Ackil and Maria S Zaragoza. 1995. Developmental differences in eyewitness suggestibility and memory for source. *Journal of experimental child psychology* 60, 1 (1995), 57–83.

[2] Rossana Actis-Grosso, Roberta Capellini, Francesco Ghedin, and Francesca Tassistro. 2021. Implicit Measures as a Useful Tool for Evaluating User Experience. In *International Conference on Human-Computer Interaction*. Springer, 3–20.

[3] AH Allam and Halina Mohamed Dahlan. 2013. User experience: challenges and opportunities. *Journal of Information Systems Research and Innovation* 3, 1 (2013), 28–36.

[4] D American Psychiatric Association, American Psychiatric Association, et al. 2013. *Diagnostic and statistical manual of mental disorders: DSM-5*. Vol. 5. American psychiatric association Washington, DC.

[5] Scarvia Anderson and Samuel Messick. 1974. Social competency in young children. *Developmental psychology* 10, 2 (1974), 282.

[6] Andrew Scott Baron and Mahzarin R Banaji. 2006. The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological science* 17, 1 (2006), 53–58.

[7] Eleonora Beccaluva, Fabiano Riccardi, Mattia Gianotti, Jessica Barbieri, and Franca Garzotto. 2022. VIC—A Tangible User Interface to train memory skills in children with Intellectual Disability. *International Journal of Child-Computer Interaction* 32 (2022), 100376.

[8] Eleonora Aida Beccaluva, Fabio Catania, Fabrizio Arosio, and Franca Garzotto. 2023. Predicting developmental language disorders using artificial intelligence and a speech data analysis tool. *Human–Computer Interaction* (2023), 1–35.

[9] David F Bjorklund, William S Cassel, Barbara R Bjorklund, Rhonda Douglas Brown, Cynthia L Park, Kim Ernst, and Felicia A Owen. 2000. Social demand characteristics in children's and adults' eyewitness memory and suggestibility: The effect of different interviewers on free recall and recognition. *Applied*

Cognitive Psychology: The Official Journal of the Society for Applied Research in Memory and Cognition 14, 5 (2000), 421–433.
- [10] Maggie Bruck and Stephen J Ceci. 1999. The suggestibility of children's memory. *Annual review of psychology* 50, 1 (1999), 419–439.
- [11] Maggie Bruck, Stephen J Ceci, and Helene Hembrooke. 2002. The nature of children's true and false narratives. *Developmental Review* 22, 3 (2002), 520–554.
- [12] Erin M Bryant. 2008. Real lies, white lies and gray lies: Towards a typology of deception. *Kaleidoscope: A Graduate Journal of Qualitative Communication Research* 7 (2008), 23.
- [13] Thomas P Carpenter, Ruth Pogacar, Chris Pullig, Michal Kouril, Stephen Aguilar, Jordan LaBouff, Naomi Isenberg, and Alek Chakroff. 2019. Survey-software implicit association tests: A methodological and empirical analysis. *Behavior research methods* 51 (2019), 2194–2208.
- [14] Fabio Catania, Micol Spitale, and Franca Garzotto. 2023. Conversational agents in therapeutic interventions for neurodevelopmental disorders: a survey. *Comput. Surveys* 55, 10 (2023), 1–34.
- [15] Don A Dillman, Jolene D Smyth, and Leah Melani Christian. 2014. *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons.
- [16] Claudiu V Dimofte. 2010. Implicit measures of consumer cognition: A review. *Psychology & Marketing* 27, 10 (2010), 921–937.
- [17] Allison Druin. 2002. The role of children in the design of new technology. *Behaviour and information technology* 21, 1 (2002), 1–25.
- [18] Asbjørn Følstad, Theo Araujo, Effie Lai-Chong Law, Petter Bae Brandtzaeg, Symeon Papadopoulos, Lea Reis, Marcos Baez, Guy Laban, Patrick McAllister, Carolin Ischen, et al. 2021. Future directions for chatbot research: an interdisciplinary research agenda. *Computing* 103, 12 (2021), 2915–2942.
- [19] Radhika Garg, Hua Cui, Spencer Seligson, Bo Zhang, Martin Porcheron, Leigh Clark, Benjamin R Cowan, and Erin Beneteau. 2022. The last decade of HCI research on children and voice-based conversational agents. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. 1–19.
- [20] Gail S Goodman and Annika Melinder. 2007. Child witness research and forensic interviews of young children: A review. *Legal and Criminological Psychology* 12, 1 (2007), 1–19.
- [21] Anthony G Greenwald, Debbie E McGhee, and Jordan LK Schwartz. 1998. Measuring individual differences in implicit cognition: the implicit association test. *Journal of personality and social psychology* 74, 6 (1998), 1464.
- [22] Anthony G Greenwald, T Andrew Poehlman, Eric Luis Uhlmann, and Mahzarin R Banaji. 2009. Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of personality and social psychology* 97, 1 (2009), 17.
- [23] Tara L Gruenewald, Margaret E Kemeny, Najib Aziz, and John L Fahey. 2004. Acute threat to the social self: Shame, social self-esteem, and cortisol activity. *Psychosomatic medicine* 66, 6 (2004), 915–924.
- [24] Rex Hartson and Pardha S Pyla. 2012. *The UX Book: Process and guidelines for ensuring a quality user experience*. Elsevier.
- [25] Carl Iver Hovland, Irving Lester Janis, and Harold H Kelley. 1953. Communication and persuasion. (1953).
- [26] Andrew Karpinski and James L Hilton. 2001. Attitudes and the implicit association test. *Journal of personality and social psychology* 81, 5 (2001), 774.
- [27] Fakhreddine Karray, Milad Alemzadeh, Jamil Abou Saleh, and Mo Nours Arab. 2008. Human-computer interaction: Overview on state of the art. *International journal on smart sensing and intelligent systems* 1, 1 (2008), 137–159.
- [28] Do-Yeong Kim. 2003. Voluntary controllability of the implicit association test (IAT). *Social Psychology Quarterly* (2003), 83–96.
- [29] Panos Markopoulos, Janet C Read, Stuart MacFarlane, and Johanna Hoysniemi. 2008. *Evaluating children's interactive products: principles and practices for interaction designers*. Elsevier.
- [30] Maurizio Mauri, Gaia Rancati, Andrea Gaggioli, and Giuseppe Riva. 2021. Applying Implicit Association Test Techniques and Facial Expression Analyses in the Comparative Evaluation of Website User Experience. *Frontiers in Psychology* (2021), 4392.
- [31] Guy Merchant. 2012. Mobile practices in everyday life: Popular digital technologies and schooling revisited. *British Journal of Educational Technology* 43, 5 (2012), 770–782.
- [32] Sevi Merter and Deniz Hasırcı. 2016. Children with Autism providing a knowledge base for educational and therapeutic product design. *Proceedings of the EDULEARN16 Proceedings; IATED: Burjassot, Spain* 1 (2016), 1771–1781.
- [33] Khairunnisa Mispa, Evi Indriasari Mansor, and Azrina Kamaruddin. 2019. Evaluating children's user experience (UX) towards mobile application: The fantasy land prototype. In *Proceedings of the 5th International ACM In-Cooperation HCI and UX Conference*. 46–54.
- [34] Henry Alexander Murray. 1943. Thematic apperception test. (1943).
- [35] Kalpana Nand, Nilufar Baghaei, John Casey, Bashar Barmada, Farhad Mehdipour, and Hai-Ning Liang. 2019. Engaging children with educational content via Gamification. *Smart Learning Environments* 6 (2019), 1–15.
- [36] Jeffrey S Nevid. 2010. Introduction to the special issue: Implicit measures of consumer response—The search for the holy grail of marketing research. *Psychology & Marketing* 27, 10 (2010), 913–920.
- [37] Jakob Nielsen and Rolf Molich. 1990. Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 249–256.
- [38] Brian A Nosek, Anthony G Greenwald, and Mahzarin R Banaji. 2005. Understanding and using the Implicit Association Test: II. Method variables and construct validity. *Personality and Social Psychology Bulletin* 31, 2 (2005), 166–180.
- [39] Brian A Nosek, Anthony G Greenwald, and Mahzarin R Banaji. 2007. The Implicit Association Test at age 7: A methodological and conceptual review. (2007).
- [40] Marianna Obrist, Virpi Roto, and Kaisa Väänänen-Vainio-Mattila. 2009. User experience evaluation: do you know which method to use? In *CHI'09 Extended Abstracts on Human Factors in Computing Systems*. 2763–2766.
- [41] B Keith Payne and Bertram Gawronski. 2010. A history of implicit social cognition: Where is it coming from? Where is it now? Where is it going. *Handbook of implicit social cognition: Measurement, theory, and applications* 1 (2010), 1–15.
- [42] Lydia Plowman and Joanna McPake. 2013. Seven myths about young children and technology. *Childhood Education* 89, 1 (2013), 27–33.
- [43] Patricia Power, Dermot Barnes-Holmes, Yvonne Barnes-Holmes, and Ian Stewart. 2009. The Implicit Relational Assessment Procedure (IRAP) as a measure of implicit relative preferences: A first study. *The Psychological Record* 59, 4 (2009), 621–640.
- [44] Andrew K Przybylski. 2019. Digital screen time and pediatric sleep: evidence from a preregistered cohort study. *The Journal of pediatrics* 205 (2019), 218–223.
- [45] Jairo Quintero, Silvia Baldiris, Rainer Rubira, Jhoni Cerón, and Gloria Velez. 2019. Augmented reality in educational inclusion. A systematic review on the last decade. *Frontiers in psychology* 10 (2019), 1835.
- [46] John C Raven and John Hugh Court. 1998. *Raven's progressive matrices and vocabulary scales*. Oxford Psychologists Press Oxford.
- [47] Konrad Schnabel, Jens B Asendorpf, and Anthony G Greenwald. 2008. Assessment of individual differences in implicit cognition: A review of IAT measures. *European Journal of Psychological Assessment* 24, 4 (2008), 210–217.
- [48] K Sherwin and Jakob Nielsen. 2019. Children's UX: Usability issues in designing for young people. *Nielsen Norman Group* 13 (2019).
- [49] Victoria Talwar, Susan M Murphy, and Kang Lee. 2007. White lie-telling in children for politeness purposes. *International journal of behavioral development* 31, 1 (2007), 1–11.
- [50] SR Thomas, R Burton Smith, and PJ Ball. 2007. Implicit attitudes in very young children: An adaptation of the IAT. *Current Research in Social Psychology* 13, 7 (2007), 75–85.
- [51] Katherine Valencia, Cristian Rusu, and Federico Botella. 2021. User experience factors for people with autism spectrum disorder. *Applied Sciences* 11, 21 (2021), 10469.
- [52] Vero Vanden Abeele, Bieke Zaman, and Dirk De Grooff. 2012. User eXperience Laddering with preschoolers: unveiling attributes and benefits of cuddly toy interfaces. *Personal and Ubiquitous Computing* 16, 4 (2012), 451–465.
- [53] Arnold POS Vermeeren, Effie Lai-Chong Law, Virpi Roto, Marianna Obrist, Jettie Hoonhout, and Kaisa Väänänen-Vainio-Mattila. 2010. User experience evaluation methods: current state and development needs. In *Proceedings of the 6th Nordic conference on human-computer interaction: Extending boundaries*. 521–530.
- [54] Jorick Vissers, Lode De Bot, and Bieke Zaman. 2013. MemoLine: evaluating long-term UX with children. In *Proceedings of the 12th International Conference on Interaction Design and Children*. 285–288.
- [55] Bernd Wittenbrink and Norbert Schwarz. 2007. *Implicit measures of attitudes*. Guilford Press.
- [56] Qian Yi, Shiquan Xiong, Biao Wang, and Shuping Yi. 2020. Identification of trusted interactive behavior based on mouse behavior considering web User's emotions. *International Journal of Industrial Ergonomics* 76 (2020), 102903.