



Phyigital interfaces for people with intellectual disability: an exploratory study at a social care center

Mirko Gelsomini¹ · Micol Spitale¹ · Franca Garzotto¹

Received: 15 April 2020 / Revised: 28 May 2021 / Accepted: 15 June 2021 /

Published online: 24 August 2021

© The Author(s) 2021

Abstract

Phyigital interaction is a form of tangible interaction where digital and physical contents are combined in such a way that the locus of multimedia information is detached from the physical material(s) manipulated by the user. The use of phyigital interaction is supported by several theoretical approaches that emphasize the development of cognitive skills dependent upon embodied interactions with the physical environment. Several studies demonstrate the potential of using phyigital technologies for supporting people with intellectual disabilities (ID) in the development of cognitive, sensorimotor, social and behavioral skills. Our research aims at exploring the potential of phyigital interaction for (young) adults with ID in a real setting, using a research platform called Reflex as a case study. For this purpose, we ran an empirical study involving 17 participants with ID and 8 specialists, and compared Reflex with approaches making use of only digital contents or paper-based materials. Our findings highlighted the potentials of phyigital approaches to perform interventions with people with ID, enhancing their performances with an appreciated interaction method. In addition, the post-study interviews with specialists favoured the adoption of phyigital technologies in a social care context.

Keywords Intellectual disability · Tangible interfaces · Phyigital interaction · Autism spectrum disorder · Comparative study

1 Introduction

Intellectual Disability (ID), also defined as Disorder of Intellectual Development by the International Classification of Diseases (ICD-11 [27], code: 6A00), is a disorder arising during the developmental period characterized by often co-occurring challenges in the

Mirko Gelsomini and Micol Spitale contributed equally to this research.

✉ Mirko Gelsomini
phd.mirko.gelsomini@gmail.com

¹ Politecnico di Milano, Piazza Leonardo da Vinci, 32, Milan, Italy

cognitive, social, communicative, motor, behavioral and emotional spheres [1]. The disorder includes deficits in intellectual functioning (e.g., reasoning, problem solving, planning, abstract thinking, judgment, academic learning) and adaptive behavior (communication, social participation), mining autonomy in everyday life [41].

Several studies investigated the use of interactive technology [32] to support people with ID in enhancing their cognitive, behavioral, social, and sensorimotor capabilities. Some of them focused on Tangible User Interfaces (TUIs) [3, 14, 20, 56], which are characterized by the combination of digital and physical contents and the use of manipulative interaction. Past research suggested that TUIs are a promising approach to help people with ID improve cognitive and sensorimotor skills, and they could complement current treatments for this target group. This potential is supported by some (mostly preliminary) empirical studies and is grounded on theory. Theoretical approaches posit learning as both an intellectual and a physical process, and emphasize the formative role of embodiment - the way an organism's sensorimotor capacities enable it to interact with the physical environment successfully - in the development of cognitive skills, such as mental imagery, memory, reasoning and problem solving [17, 63].

A special kind of TUIs are Phyigital Interfaces, combining digital and physical contents in such a way that: i) there is a clear separation between the locus where digital information are (dis)played (e.g., on screen) and the place where physical materials are manipulated by the user; ii) physical objects have both an interaction purpose (being instrumental to control the behavior of digital elements [4]) and a representational role, having a direct semantic mapping with the digital contents.

Our research focuses on the unexplored field of Phyigital Interfaces for people with ID. The paper discusses a preliminary exploratory study that was performed on the field at a care center, and involved 17 young adults with ID (among whom 14 completed the study) and their 8 therapists, for a period of 6 weeks. The research aimed at investigating Phyigital Interfaces from the perspectives of both persons with ID and caregivers, to investigate: how persons with ID perform with Phyigital Interfaces; how they like them; the factors promoting, or preventing, the adoption of Phyigital Interfaces at care centers.

For the purpose of our research, we defined a set of game-based tasks that could be performed using physical materials only (paper cards, typical tools in ID interventions), digital materials only (on tablet, also frequently used at ID care centers), or a phyigital system (called Reflex [21, 22, 55]). Reflex is inspired to a commercial system named Osmo [59]. It provides an extensible set of game-based learning activities - designed in cooperation with ID specialists (academic researchers and practitioners) - that involve multimedia contents displayed on a tablet (images, animations, sounds) and physical items (e.g., cards showing images, sentences, letters, or numbers) that the user manipulates to interact with digital contents. All participants with ID performed all assigned tasks using (in a randomized way) each of these tools. Therapists were involved in a group interview at the end of the study.

Our work is original for a number of aspects. As highlighted in Section 3, existing research mainly explores TUIs for children with and without disability and only in limited cases [20] specifically considers Phyigital Interfaces. Our study involves a target group - young adults with Intellectual Disability, who - to our knowledge - were seldom considered in past studies on TUIs. Furthermore, most research focuses on the learning benefits of TUIs or Phyigital Interfaces.

Our work explores factors related to both learning (performance and likeability) and adoption, and our comparative approach enables us to dig more deeply on the characteristics of phyigital interfaces that might affect all these factors. We contribute to a better understanding of Phyigital Interfaces not just from the learning point of view but also from the

practical perspective of adoption in real life settings. Finally, there are a number of studies in the TUI arena that compare TUIs with other interaction paradigms, but there is no comparative study specifically focusing on Phygital Interfaces.

The rest of the article is organized as follows:

Section 2 provides some background information about our target population, the current care approaches for them, and the contexts where they are performed. Section 3 summarizes the state of the art in tangible and phygital interfaces, with particular regard to people with (Intellectual) Disability, and report about comparative studies in the field.

Section 4 describes the methodology we exploited context, and Section 5 execution of the empirical study. Section 6 reports the approach for data analysis, and Section 7 presents the main findings. 8 discuss them from a critical perspective, also highlighting the limitations of our work. Section 10 draws the conclusions and portrays the directions of future work.

2 Background

The Diagnostic and Statistical Manual of Mental Disorders, 5th Edition (DSM-5 [1]), defines Intellectual Disabilities (ID) as neurodevelopmental disorders that begin in childhood and are characterized by symptoms and severity, resulting in different degrees of mental, emotional, physical, and economic consequences for individuals. These are conceptualized by the International Classification of Functioning a “functioning levels” which stand for persons level of functioning as a dynamic interaction between their health conditions, environmental factors, and personal factors, as proposed by the biopsychosocial model of disability [33]. For intellectual disabilities, functioning is categorized in four levels: “mild,” “moderate,” “severe,” and “profound”.

The American Association on Intellectual and Developmental Disabilities (AAIDD) publishes a framework for evaluating the severity of ID, the Supports Intensity Scale (SIS), which focuses on the types and intensities of supports needed to enable an individual to lead a normal and independent life, rather than defining severity in terms of deficits [49]. The SIS evaluates the support needs of an individual across 49 life activities, divided into six categories: home living, community living, life-long learning, employment, health and safety, and social activities.

Treatments for ID generally fall into three main categories:

- treatments that address or mitigate any underlying cause of ID such as particular diets;
- treatments of comorbid physical and mental disorders with the aims of improving the patient’s functioning and life skills such as targeted pharmacologic treatments [24]
- early behavioral and cognitive interventions, special education, re-habilitation, and psycho-social support [12]

Social care centers address the third category by offering services that reinforce the development of fundamental occupational skills such as the physical, cognitive, behavioral and social abilities needed to engage in daily life occupations. These centers foster people’s occupational needs by performing several activities that address:

- Employment activities and skills;
- Self-care (e.g., grooming, dressing, feeding, bathing);
- Leisure activities (e.g., knitting, playing games);
- Meaningful and purposeful activities;
- Cognitive stimulation.

These interventions are generally personalized to address each person's peculiar sensory processing and help them decoding "chaotic misinformation" that may lead to maladaptive or dangerous behaviors [18, 29].

In many cases, activities - individual or group-based - involve physical materials, most frequently paper-based. In the last 2 decades, these traditional tools have been complemented with digital applications on PCs, smartphones and tablets, acknowledged to stimulate motivation and engagement [32].

2.1 Previous works from the authors

In preliminary publications [21, 22, 55], the authors described the design rationale and requirements of the system and provided details about a first exploratory study aimed at deriving the right target group from a pool of 27 people with neuro-developmental disorder.

This paper describes a new empirical study which investigates the use of Reflex among those persons with ID whose profile was in line with the previous exploratory study, in comparisons to digital-only and paper-only activities. Our aim was to explore the potential of phygital approaches during interventions in a real setting.

3 Related works

3.1 Tangible user interfaces

While Graphical User Interfaces (GUIs) involve only user actions on visual elements in digital displays [35], by means of touch or control devices (e.g., mouse), with Tangible User Interfaces (TUIs) users interact with digital information by manipulating physical materials, and both physical and digital items are objects of interest for the user and fundamental semantic ingredients of the user experience (unlike a mouse, for example, which acts as a generic controller and a transient intermediary only) [26]). Compared to GUIs, TUIs are thought to be more natural, enjoyable and engaging, and a number of studies suggest that they are potentially more effective for learning, particularly among children [60].

Tangible technologies, especially those related to educational and ludic contexts, take their roots on the studies undertaken by Maria Montessori and Friedrich Fröbel [66]. Their work both encouraged users to explore and construct objects and artefacts freely with their hands, positing the manipulative experience at the center of cognitive development (Fig. 1).



Fig. 1 Left: Montessori inspired toys. Right: Fröbel gifts



Fig. 2 Left: necklace of Digital Beads. Center: a BitBall. Right: Dancing creature with communicating cricket. Source: [45]

Hundred years later, the work conducted by Mitch Resnick laid the groundwork for the future of tangible learning objects with the introduction of Digital Manipulatives (e.g., Programmable Brick, Digital Beads, BitBall, and Cricket as depicted in Fig. 2), a new breed of manipulative materials with integrated computational power designed “to expand the range of concepts that children can explore through direct manipulation”, enabling them to learn concepts that were previously considered too advanced [46, 47, 66, 67].

TUIs are thought to enhance skills related to physical manipulation, physical-digital mappings, and multisensory exploration, providing richer sensory and learning experiences through the interweaving of computation and physical materials [2, 3, 14]. They extend the intellectual and emotional potential of interactive artifacts and integrate compelling and expressive aspects of traditional educational technologies with creative and valuable properties of physical objects [13].

Some empirical studies highlighted the benefits of TUIs also for children with Intellectual Disability. Tangible interaction is claimed to be more natural and familiar than other types of interfaces [38], lowering the threshold of participation; enhancing e-accessibility and inclusion; and fostering independent exploratory, assistive and collaborative learning [48, 64].

Garzotto et al. [20] focused on supporting social and cognitive development of low functioning children with ID in school contexts by designing and evaluating the Talking Paper (Fig. 3). The tool enabled teachers to pair conventional paper based elements (e.g. cards, drawing, pictures) with multimedia resources (e.g. videos, sounds). The authors discovered enhanced participation and sense of community in the overall classroom.

Starcic et al. [56] studied the efficacy of TUIs to improve teaching and learning over a 2-year period with 39 teachers and 145 students aged 1 to 15, including 30 students with



Fig. 3 The use of Talking Paper. Left: a child interacting with the little red cap. Right: class work. Source: [20]



Fig. 4 From left to right: Augmented Objects, LightTable, Drum Machine, Sifteo cubes. Source: [15]

low fine-motor skills or learning difficulties. The authors found TUIs supported inclusion in the classroom and concept development with physical and virtual representations based on dynamic geometry. Similar results were also obtained in the studies performed by Dandashi et al. [11] and by Zajc et al. [64]. The former study evaluated a TUI-based system aimed at enhancing the learning process with 77 children with ID. Results showed positive effects of children's physical activation and motivation levels. Particularly, those with medium functioning levels achieved the best results in terms of scores and coordination. The latter study questioned the usefulness of TUI-based games (Raindrop Catcher) in inclusive classrooms for students with severe to mild ID. Their exploitation allowed all students to be equally engaged in the game-based learning process, thus supporting collaborative learning and overcoming those limitations noticed in computer-based activities only.

Falcao [15] studied children with ID playing with different tangible artifacts (Augmented Objects, LightTable, Drum Machine, Sifteo cubes depicted in Fig. 4) and points out the effectiveness of tangible interaction for exploratory learning purposes by suggesting the most efficient gaming paradigm as the one with a clear mapping between specific physical objects and their meanings.

A similar mapping approach was also proposed by Tam et al. with Polipo [57]. The authors investigated the learning benefit of a 3D printed smart toy co-designed with therapists that provides various manipulatory affordances and offers feedbacks and rewards by means of lights, sounds, and music integrated in its body. An exploratory study highlighted some benefits of Polipo for children with neuro-developmental disorder with very severe impairments in the motor and cognitive area, improving fine motor skills and encouraging children's communication with their therapist Fig. 5. With Polipo, authors identified several guidelines to design low-cost TUIs for children with ID.

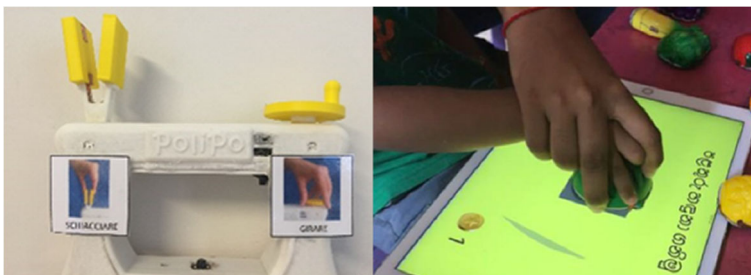


Fig. 5 Left: Polipo mapping between required interaction and real-life one (e.g. turn action on a 3d printed pommel mapped to rotating a door knob) [57]. Right: A child and a therapist playing with POMA [36]

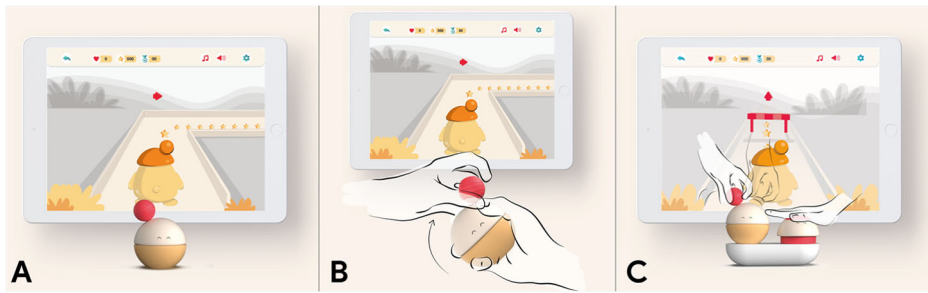


Fig. 6 a Yogo. One b or more c users interact with physical tools and the result is shown on the screen

3.2 Phygital interfaces

Phygital Interfaces can be regarded as a subcategory of the broader class of Tangible User Interfaces (TUIs). They are characterized by the combination of digital and physical contents in such a way that: i) there is a clear separation between the locus where digital information are (dis)played (e.g., on screen) and the place where the physical material(s) is manipulated by the user; ii) physical objects have both an interaction purpose (being instrumental to control the behavior of digital elements [4]) and a representational role, having a direct semantic mapping with the digital contents.

To clarify these characteristics, we can compare the use of Phygital Interfaces such as Yogo [8] (Fig. 6) with two systems cited in the previous subsection - Reactable [30] and Polipo [57]. Differently from Yogo, Polipo and Reactable are tangible but not phygital, since the visual, light, and sound effects are “on” (in Polipo) or “closely aside” (in Reactable) the physical objects manipulated by the user.

An example of Phygital Interface can be found in NIKVision [9]. The system, making use of a tabletop computer and tangible interaction, increased collaborative learning in children with ID (Fig. 7). Another example can be found in Poma [36], a phygital system to improve social and cognitive skills of Sri-Lankan children with autism spectrum disorder (ASD) (Fig. 5).

The Phygital Interface used for the purpose of our exploratory study is Reflex [21, 22, 55], Reflex was co-designed with ID specialists and developed at our lab, taking inspiration from a commercial system named Osmo [59]. Reflex provides a number of game-based



Fig. 7 NIKVision farm game prototype. Left: the platform. Center: the GUI. Right: Children during the study. Source: [37]

educational activities for persons with ID that involve multimedia contents (images, animations, sounds) and physical items (e.g., cards displaying images, sentences, letters, or numbers) that are manipulated by the user to interact with digital contents. Reflex is discussed in detail in the following section.

3.3 Comparative studies

There are few examples of works that compare tangible interaction against other interaction paradigms. In [43], the authors studied how different interaction modes in a museum exhibition impacted on visitors' mobility and engagement. Researchers evaluated the use of paper based tools with near-field communication (NFC) tags inserted against digital ones (smartphone) and tangible ones (smartphone + NFC tags) and found that tangible interaction was the most liked, while the digital version favoured mobility to the detriment of engagement with the exhibition.

Many comparative studies in the TUI arena involve children with disability, but in no cases consider Phygital Interfaces. Sitdhisanguan et al. [52, 53] investigated the effects of TUI-based approaches against Graphical User Interfaces (GUI)-based ones making use of a standard computer mouse as a pointing device (Fig. 8). The empirical study, performed with 20 children aged 3 to 7 on the Autism Spectrum with comparable learning abilities, shown that TUI were more usable and potentially more effective in learning than mouse-based ones when learning geometry.

Chipman et al. [10] compared the use of a paper-only system with a tablet-based system to discover new dynamics in children's drawing experiences. The study identified several collaborative advantages of using the technology-based system, including: increased awareness, more shared experiences, and longer time participating in activities.

Finally, Schneider et al. [50] and Song et al. [54] discussed the role of tangibility in cognitive tasks by comparing tangible interfaces with multi-touch and screen-based respectively, providing evidence in favor of tangibles in terms of motivation, performance, collaboration and learning.

4 Methodology

4.1 Context

The study was carried out at Fraternità & Amicizia (F&A), a local social care center in Milan (Italy). F&A is an accredited private association that provides numerous services

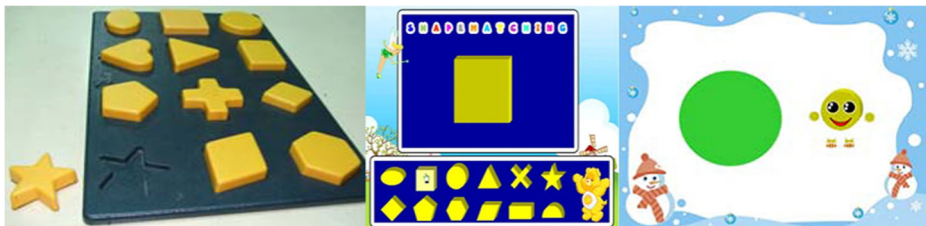


Fig. 8 Learning basic geometry shapes. Left: conventional training materials. Center: GUI for the mouse interaction. Right: GUI for the tangible interaction. Source: [52, 53]

to persons with ID, with the intent to promote their occupational skills, integration, autonomy, and well-being: from purely assistance interventions to job-training classroom, from operative laboratories (e.g., arts and crafts) to individual sessions and psychological support. At F&A, as well as in most public and private social care institutions in Italy, guests are an heterogeneous population with different cognitive diagnoses and functioning levels. Despite these differences, the common factor is the presence of an Intellectual Disability (severe to mild) eventually deriving from genetic conditions and/or associated with other pathologies. Guests are usually organized in groups, composed by people with comparable functioning level, regardless their diverse diagnosis. This approach helps ensuring the maximum level of homogeneity and allowing caregivers to propose activities that are not under/over stimulating or frustrating/boring.

4.2 Tasks

For the purpose of the study, we asked participants to play a Tangram activity. Participants performed the Tangram task using 3 tools with related interaction modality: a phygital mode (Reflex, hereinafter RM), a paper-based mode (hereinafter PM), and a digital one (with tablet, hereinafter TM). Each modality shows a figure composed by essential geometric items (e.g. triangle, square, trapeze) and participants have to match the figure by putting every item in its exact position and rotation. According to the literature [55] and the practice we observed in several care centers during past research, Tangram is commonly used in interventions for persons with ID by being usable by all participants of the center regardless their functioning levels.

As Tangram tasks, we selected a set of twelve figures with the same complexity (i.e., 7 items each) and difficulty representing animals as shown in Fig. 9.

4.3 Tools

4.3.1 Reflex

For the purpose of the research we exploited a phygital system called Reflex, co-designed with specialists, developed at our lab [21, 22, 55]. Reflex is an application on iOS and Android devices inspired by Osmo, an educational app developed by Tangible Play [58, 59].

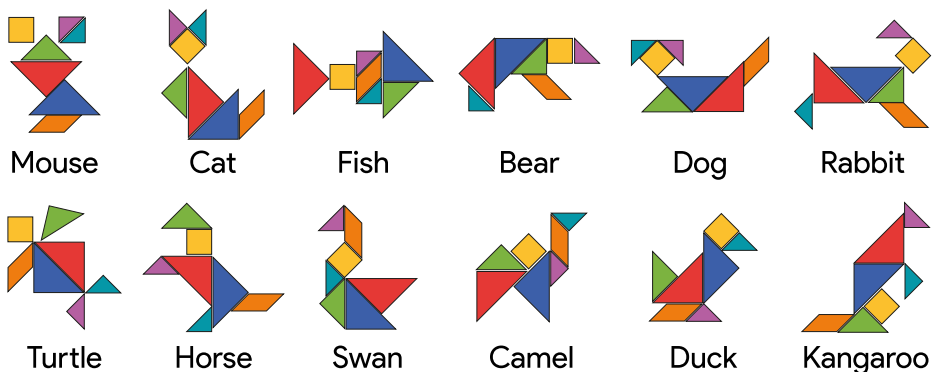


Fig. 9 Tangram figures

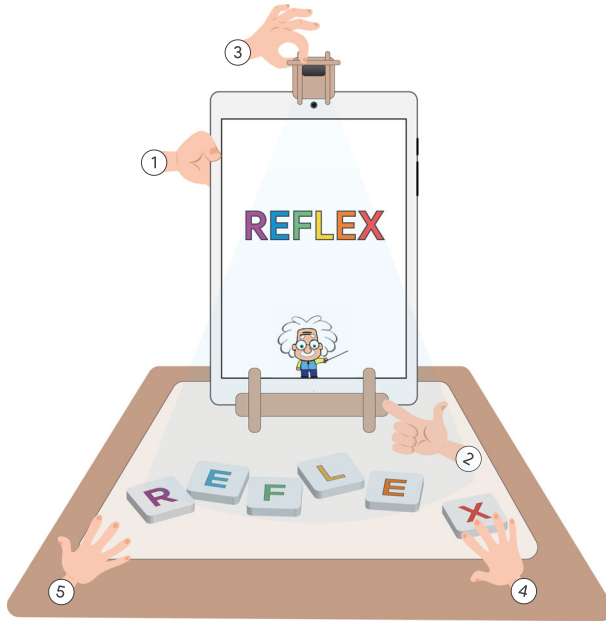


Fig. 10 Reflex components: generic tablet (1), device stand (2), adjustable camera attachment (3), cards (4) and specifically designed software, playground mat with solid black borders (5). Source: [55]

The application tracks and recognizes physical items (such as cards, Fig. 10.4) placed on a bordered (solid black) mat (“Play Zone”) as depicted Fig. 10.5) via a bottom-looking mirror position on the device camera, and controls the behaviour of multimedia elements on the device screen (images, video, animations, sounds, and music) according to the game logic of the ongoing user activity. Reflex equipment includes a (wooden or plastic) device stand (Fig. 10.2) and an adjustable camera attachment (Fig. 10.3).

Technically speaking, the camera attachment adapts a camera integrated in a device and includes a housing including a slot on a first side. The slot is configured to receive and retain an edge of a body of the computing device. The housing is configured to cover at least a portion of the field of view of the camera of the device. A reflective element is recessed at an angle into the first side of the housing to redirect the camera field of view toward an activity surface located proximate the device [51]. The app processes the reflected video stream; identifies contours of each object placed on the play zone and infers an object shape by mapping the coordinates of vertices, orientation, position, colors. The virtual information on the screen is then generated based on the recognized objects.

The main steps to play Tangram on Reflex are depicted in Fig. 11 and described as follows. First, a random shape is highlighted to suggest the user to focus on the shape and to manually reproduce its position and rotation (Fig. 11.1 - red shape). Second, the next shapes are freely chosen by the user (Fig. 11.b, e.g. yellow triangle). The entire animal figure starts appearing more colored and, for each item placed in the right place, a virtual character, resembling a little Einstein, congratulates the user (Fig. 11.c). Last, after having positioned all shapes correctly (Fig. 11.d), the animal figure is entirely colored and the little Einstein shows and tells his happiness through its motion as depicted in Fig. 12.

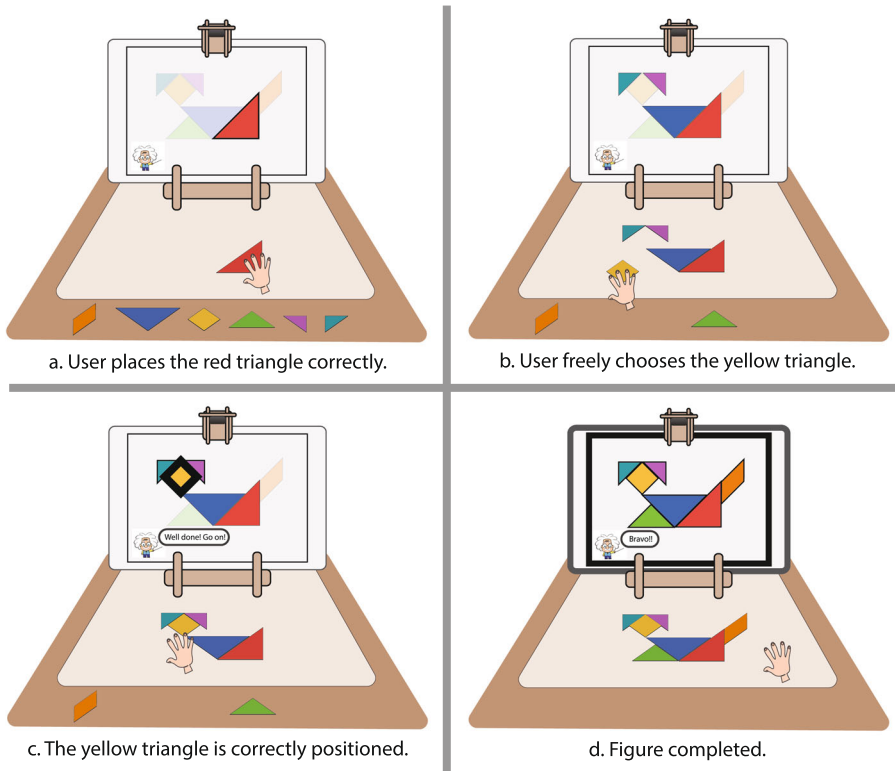


Fig. 11 Tangram activity steps in reflex



Fig. 12 Little Einstein - happy animation sprites

4.3.2 Paper and tablet modalities

Participants have to complete the Tangram tasks similarly in paper and tablet modes.

In Paper Mode (PM), a full animal figure is shown to the participant through a printed sheet. At first, participants receive a hint regarding the first item (red shape) pointing at it. Then, they continue the task autonomously, and the caregiver provides feedback on correct action. Last, after having positioned all shapes correctly, the caregiver congratulates on task completion.

In Tablet Mode (TM), the Tangram task starts when the tablet-app displaying the figure to replicate. Then, the first item (red shape) is highlighted to support participants in initiating the task. When they correctly placed the item, the outline of the shape became bold, and the little Einstein virtual character provides them a positive feedback. As the task has been completed, the virtual character congratulates the participants.

4.3.3 Design comparison between the modes

As in Table 1, in every modality, the activity has a deck-zone (a place where the initial items are located); an instruction-zone (a place where the seven-items-figure to be reproduced is shown) and a play-zone (a place where the participant places the items to be recognized by the application) depicted in Fig. 13. In the paper mode (PM), the instruction-zone is shown on paper, while the deck and play zones are located on the table. The latter in front of the subject and the former on his left (see Fig. 13.a). In Reflex (RM), the instruction-zone is shown on tablet screen, while the deck and play zones are located on the table. The latter in front of the subject and the former on his left (Fig. 13.b). For the tablet modality (TM), the instruction and play zones overlap on the right part of the tablet, while the deck-zone on the left displays the available items needed to be moved and rotated (Fig. 13.c).

4.4 Research questions and variables

Our research questions for the study were the following:

- 1 How do persons with ID perform in Tangram with Phygital Interfaces in comparison with the paper and digital interfaces? How do their performance evolve over time?
- 2 Which version of the game do they like the most? Which the least? How do their preference evolve over time?
- 3 What are the factors promoting, or preventing, the adoption of Phygital Interfaces (like Reflex) at care centers in comparison with paper-based and tablet-based tools?

We defined the research variables taking inspiration from previous works related to the use of tangram-based activities for people with disability [6, 16, 19, 25, 39, 40, 44, 61, 65] as follows:

- 1 *User Performance*: this variable measures the degree of success in accomplishing the Tangram tasks with respect to the time employed within a maximum time; a Tangram task is accomplished when all items are placed in the right position to represent the assigned figure within a predefined threshold time;
- 2 *User Likeability*: this variable considers the user's preference among the three versions of the Tangram - using paper, tablet, and Reflex;
- 3 *Adoptability*: this variable considers the potential for adoptability of Reflex at care centers, as perceived by the therapists.

Table 1 Comparison between PM, RM and TM. Yellow color highlights the differences between one tool and the other two

Design variables	Paper Mode (PM)	Reflex Mode (RM)	Tablet Mode (TM)
Deck zone (Spatial allocation)	On the table-horizontal On the left of play zone	On the table-horizontal On the left of play zone	On the table-horizontal Framed on tablet On the left of play zone Smaller size
Play zone (Spatial allocation)	On the table-horizontal On the right of deck zone Initially empty, progressively populated with the selected items	On the table-horizontal On the right of deck zone Initially empty, progressively populated with the selected items	On the table-horizontal Framed on tablet On the right of deck zone Overlapping with Instruction zone. Contains the target shape
Instruction zone (Spatial allocation)	On the table-horizontal Framed on paper-sheet On top of play zone	On the table- vertical Framed on tablet On top of play zone	On the table-horizontal Framed on tablet Overlapping with Play zone Digital-based image
Shape (in instruction zone)	Paper-based image Same size and visual properties as RM	Digital-based image Same size and visual properties as PM	Same visual properties as RM and PM but smaller size
Items (Deck Components)	Physical items Same as RM	Physical items Same as TM	Digital items Same as RM, PM but smaller size
Action (Move deck item to play zone)	Manipulation: grasp and move physical item	Manipulation: grasp and move physical item	Touch: "Drag-drop digital item"
Hint (to user)	At the beginning of the game the caregiver points to the largest item (the red one) on the deck and invites the user to start with it deck. After this first item is placed, no other hints are provided during the game	At the beginning of the game in the Instruction zone all elements are shadowed but the largest item (the red one), which is bold and blinking. After this first item is placed, no other hints are provided during the game	At the beginning of the game in the shape in the Instruction zone all elements are shadowed but the largest item (the red one), which is bold and blinking. After this first item is placed, no other hints are provided during the game
Feedback (on correct action)	When a component is placed in the right place, the caregiver says "we'll done! Go on!"	When a component is placed in the right place, the color of the component in the Instruction area is bold.	When a component is placed in the right place, the color of the component in the Instruction area is bold.
Feedback (on task completion)	Voice congratulations ("Bravo!") by caregiver	Voice congratulations ("Bravo!") by the animated virtual character	Voice congratulations ("Bravo!") by the animated virtual character

Light-blue color highlights the difference between TM and RM, and PM

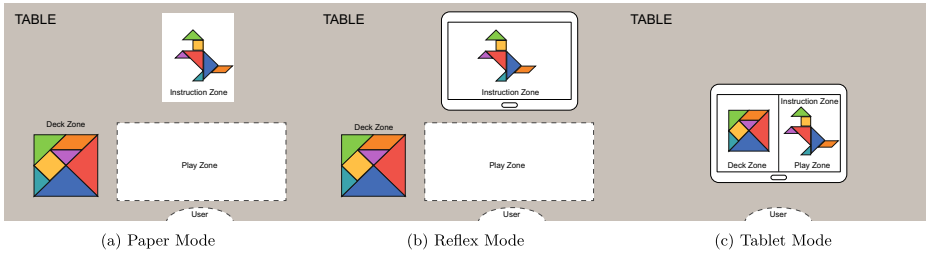


Fig. 13 Interaction modes

4.5 Design

The empirical study was designed as a longitudinal study in which participants performed, in each study session, 2 Tangram tasks in each of the 3 interaction modes (PM: Paper mode, RM: Reflex mode, TM: Tablet mode). The study independent variable was the interaction mode (PM, RM, and TM) while the dependent variables were the *User Performance* and *User Likeability* measured at each session. Each participant attended 6 sessions, once a week to avoid carry-over effects.

For each interaction mode, specialists defined 6 pairs of tasks depicted in Table 2.

In each session, participants performed three task pairs (one for each interaction mode). The task pairs were randomized with respect to the interaction mode as shown in Table 3 to avoid order effects [28].

Specialists decided for two tasks for each interaction mode in each session because a single task was considered too short (at maximum 2 minutes). We opted for a 6 sessions long study (approximately 1.5 months) for two reasons: i) participants should experience equally PM, RM, and TM in a randomized order, therefore the number of sessions should be a multiple of 3; ii) 6 sessions seemed to be a reasonably long period (also compared with other existing studies with persons with ID [32]) to observe an evolution in performance.

5 Empirical study

5.1 Participants

The study was designed and performed in collaboration with Fraternità & Amicizia (F&A), a local social care center. Specialists at F&A recruited 17 participants between 21 and 37 years old (*avg* = 24, *SD* = 3.9) daily attending the care center.

Table 2 Task pairs

Task Pairs	Shape 1	Shape 2
a	Camel	Dog
b	Kangaroo	Horse
c	Swan	Rabbit
d	Cat	Bear
e	Duck	Fish
f	Turtle	Mouse

Given the exploratory nature of the study, specialists followed a labile inclusion approach with the following criteria applied to potential participants:

- being diagnosed with an Intellectual Disability;
- present a typical development of fine motoric movement and coordination (i.e., being able to use a pen to write, or grab and carry an apple in their hands);
- be novel both to Tangram activity and Reflex to avoid any memory-based learning [5] or expertise effects [62];
- have a high attendance to the center activities to be able to easily organize activities at a constant period;
- not having external issues (such as familiar ones) which could affect their psychological sphere.

Table 4 lists participants by unique identifier (ID), age, diagnosis, functioning level and related ICD-11 code [27]. This sample size can be considered as having the same order of magnitude with similar studies in the field [32].

We also involved 8 specialists from 21 to 55 y.o. ($avg = 34.5$, $SD = 11.8$). Of them 2 were voluntary social workers, 2 educators, 3 developmental and behavioral psychologists, and 1 social care center manager. Informed consent for participation and video recording was obtained for all participants and specialists.

5.2 Setting

For the purpose of the study, we established a separated room that was familiar to the participants, but not associated to any specific everyday activity. The room was also selected for its location that assured auditory quietness and all distracting stimuli (e.g., posters, canvases or colorful furniture) were removed. A distance of 80-90 cm was maintained during all sessions from the participant to the instruction-zone and 30-40 cm to deck and play-zone. Components of Tangram kit were always placed in the same position as shown in Fig. 13.

5.3 Procedure

We organized the study according to the following steps:

1. Familiarization phase: homogenization of sample and definition of cut-off time;
2. Study sessions: execution of assigned Tangram tasks by participants, *performance* data gathering during task execution, and collecting *likeability* data at the end of each session through a questionnaire;

Table 3 Pair-rotation system

<i>Sessions</i>			
1	PM(a)	RM(b)	TM(c)
2	TM(b)	PM(c)	RM(a)
3	RM(c)	TM(a)	PM(b)
4	PM(d)	RM(e)	TM(f)
5	TM(e)	PM(f)	RM(d)
6	RM(f)	TM(d)	PM(e)

Table 4 Participants with intellectual disability recruited for empirical study

<i>ID</i>	<i>Age</i>	<i>Diagnosis</i>	<i>Func. Level</i>	<i>ICD-11</i>
P1	22	ID +Kabuki	Moderate	LD2F.1Y
P2	21	ID	Moderate	6A00.1
P3	26	ID	Mild	6A00.0
P4	37	ID	Severe	6A00.1
P5	21	ID + Struge-Weber	Moderate	6A00.1
P6	26	ID	Mild	6A00.0
P7	22	ID	Mild	6A00.0
P8	21	ID + Arnold-Chiari	Mild	LA03
P9	23	ID + Autism	Moderate	6A02
P10	21	ID	Mild	6A00.0
P11	27	ID	Moderate	6A00.1
P12	23	ID	Mild	6A00.0
P13	26	ID + Autism	Mild	6A02
P14	25	ID + Autism	Moderate	6A02
P15	22	ID	Moderate	6A00.1
P16	25	ID	Severe	6A00.1
P17	22	ID	Mild	6A00.0

- Final Group interview: interview of specialists after all study sessions were completed to gather feedback on study results and collect additional information on *adoptability*;

5.3.1 Familiarization phase

A first session was performed to teach participants how to use Tangram. A researcher introduced herself, explained Tangram, and showed its components. The researcher drove the entire session as an External Examiner (EE) - keeping a distant and neutral attitude towards participants - to avoid bias due to interaction with a familiar person and consequent data contamination. During this phase, the EE asked the participants to play the traditional Tangram activity providing them a figure of seven items which did not resemble any of the shapes provided in the study (e.g. as an house, a ship). The maximum time available (t_{max}) to complete the given figure was set to 150seconds as defined in [55] (Fig. 14).



Fig. 14 A participant playing Paper-based (left) and Reflex-based (right) Tangram

5.3.2 Study sessions

At the beginning of each session, participants were invited to enter the room and take a seat. Each session was organized in two distinct moments: Play and Respond.

Play The kit with tablet and Tangram items was placed on the table in front of the them. They were to position items in the space defined by the board and to interact with them within the working space. An average session lasted about 10 minutes. The EE introduced and provided the predefined sequence depicted in Table 2. During the session, EE was instructed not to give any feedback or hint to participants to accomplish the task goal. They reproduced two different figures chosen according to a pair-rotation system described in Tables 3 and 2 for each interaction mode. If participants didn't feel comfortable during the session, they were free to abandon it whenever they needed.

For each task we collected the following interactional data:

- the time needed for a item to be placed in the correct position);
- the total number of items that are placed in the correct position within the predefined threshold set for task completion (150 seconds, as identified in the familiarization phase).

We gathered these performance data automatically for RM and TM, and manually for PM.

Respond At the end of each session, a questionnaire was administered with the following closed questions (multiple-choice):

- Which interaction mode did you like most?
- Which interaction mode did you like least?

Participants were asked to answer by saying or pointing at pictures or elements representing Paper, Tablet or Reflex through visual aids. When possible, they were also asked to provide a why for their choices. If participants could not understand the questions the EE paraphrased and clarified them. Participants answers were collected by the EE through a Google Form.

5.3.3 Final group interview

At the end of the empirical study, three researchers lead a group interview with the participating specialists (N=8). The session was video-recorded and consent was obtained ahead. Each researcher had a given role: facilitator, who lead the group interview; assistant, who assisted the facilitator with detailing questions; recorder, who captured a detailed account of each participant input.

We designed the group interview to last about two hours, and we used a semi-structured protocol. We asked specialists to discuss about their feelings on the study. Our questions investigated key aspects emerged in a previous study [55] and from the current study about adoption of phygital technologies in the center.

Specialists were also prompted to express their opinion regarding the opportunities and challenges raised by the introduction of this kind of technologies in a social care context. The interview started by asking to recall the first moments in which Reflex was introduced.

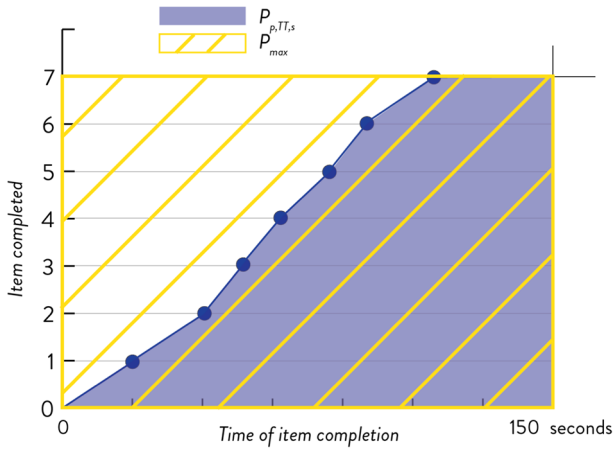


Fig. 15 Example of *User Performance* computation

We then asked specialists to focus on the activities and feelings they had when seeing participants performing activities with the different interaction modes during the study. Finally, we requested to focus on the potential adoption of this tool in regular social care interventions.

For gathering data, we transcribed the interview with automatic speech recognition tools.

6 Data analysis

6.1 Performance

We measured the *User Performance* of a participant P in a Tangram Task TT within a session s ($UP_{p,TT,s}$), as follows:

- 1) we considered the integral of the items placed by P during task TT of session s in the correct position over time within the time interval $[(t_{min} = 0), (t_{max} = 150s)]$. This value ($P_{p,TT,s}$) corresponds to the blue area in Fig. 15.
- 2) we normalized this measure, calculating - for each Participant, Tangram Task, and session, the *User Performance* ($UP_{p,TT,s}$) as as the ratio between ($P_{p,TT,s}$) and *Max Performance* (P_{max}), which corresponds to the whole rectangular area with the yellow texture of Fig. 15:

$$UP_{p,TT,s} = \frac{P_{p,TT,s}}{P_{max}} \tag{1}$$

A Kolmogorov-Smirnov test indicates that the *User Performance* throughout sessions of the three interaction modes do not follow a normal distribution, (e.g., Session 1: $D(252) = 0.12, p < 0.001$). Given the non-normal distribution of the dependent variable, we chose Friedman non-parametric tests to run our statistical analysis. We performed it using IBM SPSS [34]. We compared the differences between interaction mode during Session 1 and Session 2 using two Friedman non-parametric test, and then we computed the post-hoc analysis with a Wilcoxon Signed Ranks test with the Bonferroni correction. We ran three non-parametric tests (one for each condition, RM, TM, and PM) to compare the *UserPerformances* measured across sessions. Then, we also ran post-hoc Wilcoxon Signed Ranks



Fig. 16 Group interview with specialists

test analysis with the Bonferroni correction to evaluate the difference among the sessions for each interaction mode.

6.2 Likeability

For the *User Likeability*, our sample was a discrete (nominal) outcome variable with three response options: RM, PM, and TM. We adopted six one-sample Chi-square tests (one for each session) analysis to evaluate if the user preferences towards PM, RM, and TM changed across the sessions using IBM SPSS [34].

6.3 Adoptability

The audio of the videorecording of the group interview with the 8 therapists was translated into text by means of an automatic speech-to-text tool. Using the video, they also associated - manually - the textual sentences to the caregivers's ID (C1-C8). To analyse these materials, we opted for a thematic coding approach [23]. Thematic coding is a method for analysing qualitative data that involves identifying sections of text that are linked by a common theme or idea, allowing to categorize the text and identify items of analytic interest in the data, tagging these with a coding label. We prepared a “code-book”, in which the main labels were “pro-PM” (in favour of paper-mode), “pro-RM” (in favor of Reflex mode), “pro-TM” (in favour of tablet-mode), and any combination of them (e.g., “pro-PM; pro-RM”) indicates a favourable attitude towards both paper-mode and Reflex. Additional labels were inspired by themes related to adoptability that we had identified in co-design sessions with therapists during past research [55] (Fig. 16).

7 Results

We collected data from 14 participants (P1-P14) out of the 17 recruited ones (see Table 4). P15 and P16 expressed the willingness to interrupt the experimentation since they felt over-stressed during task execution. P17 interrupted the participation for personal and external reasons.

7.1 Performances

7.1.1 Comparison between interaction modes

We performed non-parametric tests to compare the *User Performance* between RM, PM, and TM in all sessions. We made the comparison in each pair of subsequent sessions

(see Tables 5, 7, 6 in Section 7.1.2) and also compared the *User Performance* for each interaction mode in the first and last session (*S1* and *S6*). Below we focus on the latter comparison, since only for *S1* and *S6* the difference in *User Performance* was statistically significant in each interaction mode. The first Friedman's test showed that there was a significant difference between *User Performance* measured with the three modalities in the *S1*, $\chi^2(2) = 19.000$, $p < .05$, see Fig. 17 (top). Post-hoc tests using a Wilcoxon signed-rank test with Bonferroni correction (alpha level of 0.05/3) showed that *User Scores* with RM ($Mdn = .777$) were higher than scores obtained with TM ($Mdn = .617$) during *S1*. This improvement was statistically significant ($T = 105$, $z = -3.296$, $p = .001$). The same results have been obtained between PM ($Mdn = .753$) and TM where the difference was statistically significant ($T = 102$, $z = -3.107$, $p = .002$). However scores with RM did not significantly differ from the scores obtained interacting with PM ($T = 27$, $z = -1.601$, $p = .109$). The second Friedman's test for *S6* showed that there was a significant difference between *User Performance* measured with the three modalities, $\chi^2(2) = 18.429$, $p < .05$, see Fig. 17 (bottom). Post-hoc tests using a Wilcoxon signed-rank test with Bonferroni correction (alpha level of 0.05/3) showed that *User Scores* with RM ($Mdn = .897$) were higher than scores obtained with TM ($Mdn = .696$) during *S6*. This improvement was statistically significant ($T = 104$, $z = -3.233$, $p = .001$). The same results have been obtained between PM ($Mdn = .872$) and TM where the difference was statistically significant ($T = 105$, $z = -3.296$, $p = .001$). However scores with RM did not significantly differ from the scores obtained interacting with PM ($T = 36$, $z = -1.036$, $p = .300$).

In both S1 and S6 there was a statistical difference between PM and TM, and between RM and TM, but no significant difference between PM and RM (see Figure 17).

7.1.2 Finer grained results between session pairs

The first Friedman's test to compare performances between sessions showed that there was a significant difference between *User Performance* measured across the six sessions in the RM, $\chi^2(5) = 21.96$, $p < .05$. Post-hoc tests using a Wilcoxon signed-rank test with Bonferroni correction (alpha level of 0.05/6) showed that user scores with RM during *S6* ($Mdn = .897$) were higher than scores during *S1* ($Mdn = .777$). This improvement was statistically significant $T = 98$, $z = -2.856$, $p = .004$. The *User Performance* with RM was compared between the sessions (as reported in Table 5). No statistical difference was computed among the sessions but between *S1* and *S6*.

The second Friedman's test showed that there was a significant difference between *User Performance* measured across the six sessions in the TM, $\chi^2(5) = 25.633$, $p < .05$. Post-hoc tests using a Wilcoxon signed-rank test with Bonferroni correction (alpha level of 0.05/6) showed that user scores with TM during *S1* ($Mdn = .617$) were lower than scores during *S6* ($Mdn = .696$). This improvement was statistically significant, $T = 99$, $z = -2.919$, $p = .004$. Again, the *User Performance* with TM was compared between the sessions (as reported in Table 6). No statistical difference was computed among the sessions but between *S1* and *S6*.

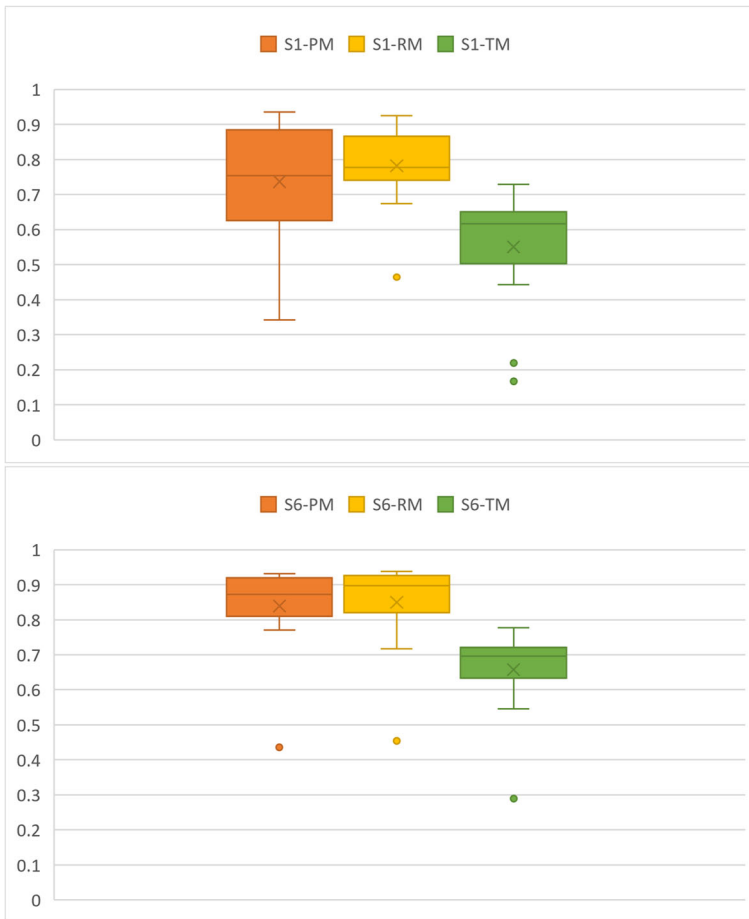


Fig. 17 *User Performance* Session 1 (top) and 6 (bottom)

The third Friedman’s test showed that there was a significant difference between *User Performance* measured across the six sessions in the PM, $\chi^2(5) = 14.490, p < .05$. Post-hoc tests using a Wilcoxon signed-rank test with Bonferroni correction (alpha level of $0.05/6$) showed that user scores with PM during *S1* ($Mdn = .753$) were lower than scores during *S6* ($Mdn = .872$). This improvement was statistically significant $T = 101, z =$

Table 5 Wilcoxon Signed Ranks test results of pair-comparison among Reflex (RM) sessions with a Bonferroni correction ($.05/6=0.008$)

<i>Pair-sessions</i>	<i>T</i>	<i>z</i>	<i>p-value</i>
S1-S2	80	−1.726	.084
S2-S3	64	−0.722	.470
S3-S4	70	−1.099	.272
S4-S5	73	−1.287	.198
S5-S6	42	−0.659	.510
S1-S6	98	−2.856	.004

−3.045, $p = .002$. Also this case, the *User Performance* with PM was compared between the sessions (as reported in Table 7). No statistical difference was computed among the sessions but between S1 and S6.

The three tests showed a statistically significant improvement of the User Performance between the first and the last session in all the interaction modes (see Figure 18)

Figure 18 depicts an overview of the performance, in terms of *User Performance*, of the participants across the sessions. Participants' scores increased over time in every interaction modalities. For the TM, the increasing score curve was steeper with respect to the other two modes, which reached the plateau earlier.

7.2 Likeability

Five participants in S1 and three in S6 preferred not to answer to the questionnaire. Table 8 shows the participants counted preferences over the session, while Table 9 shows which interaction modality the participants liked the least over the session.

Sixteen times participants preferred the PM modality, 48 times participants preferred RM over the other interfaces, while only 10 times participants picked the TM as the most liked modality.

As regards the least preferred modalities, participants chose the PM as the worst for 29 times, the RM one only for 3, and TM for 42 times in total.

Still, the six one sample chi-square tests showed that there was a statistically significant difference in the participants' responses of *User Likeability* in sessions S1, S3, and S4 (S1: $\chi^2(2, 9) = 8.667$, $p < .05$, S2: $\chi^2(2, 14) = 2.000$, $p = .368$, S3: $\chi^2(2, 14) = 7.000$, $p < .05$, S4: $\chi^2(2, 14) = 7.000$, $p < .05$, S5: $\chi^2(2, 14) = 3.571$, $p = .168$, S6: $\chi^2(2, 11) = 3.500$, $p = .174$).

Participants expressed significantly higher preferences towards RM with respect to the paper and digital interfaces.

Table 6 Wilcoxon Signed Ranks test results of pair-comparison among Tablet (TM) sessions, with a Bonferroni correction alpha level of (.05/6=.008)

<i>Pair-sessions</i>	<i>T</i>	<i>z</i>	<i>p-value</i>
S1-S2	94	−2.605	.009
S2-S3	55	−0.157	.875
S3-S4	74	−1.350	.177
S4-S5	38	−0.910	.363
S5-S6	91	−2.417	.016
S1-S6	99	−2.919	.004

Table 7 Wilcoxon Signed Ranks test results of pair-comparison among Paper (PM) sessions, with a Bonferroni correction alpha level of (.05/6=.008)

<i>Pair-sessions</i>	<i>T</i>	<i>z</i>	<i>p-value</i>
S1-S2	100	-2.982	.003
S2-S3	52	-0.031	.975
S3-S4	98	-2.856	.004
S4-S5	18	-2.166	.030
S5-S6	36	-1.036	.300
S1-S6	101	-3.045	.002

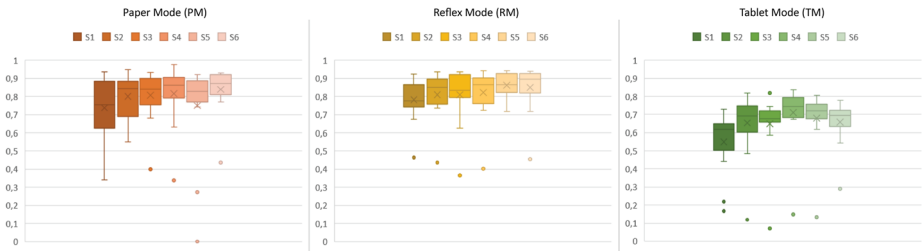


Fig. 18 *User Performances per session*

Table 8 Participants preferences towards the most liked interaction mode for each session

<i>Sessions</i>	<i>PM</i>	<i>RM</i>	<i>TM</i>
S1	2	7	0
S2	3	8	3
S3	4	9	1
S4	4	9	1
S5	3	8	3
S6	3	7	2

Table 9 Participants preferences towards the least liked interaction mode for each session

<i>Sessions</i>	<i>PM</i>	<i>RM</i>	<i>TM</i>
S1	4	1	4
S2	7	0	7
S3	3	0	8
S4	7	0	7
S5	5	0	9
S6	3	2	7

7.3 Group interview results

Three researchers autonomously labelled the interview transcripts using the method described in section and compared their labeling, reaching an high inter-rater agreement of (Cohen's $\kappa = 0.87$). Those contributions which reached a full level of classification agreement are reported below. C6 started the interview session by recalling the very first time she saw Reflex: *"During our very first introduction of Reflex, I thought it was very complex"* [pro-PM, pro-TM] and C7 agreed and added: *"since the beginning I knew paper mode would have been easier because it requires a more natural and common use [...] participants are required just to look at the image and reproduce it"* [pro-PM]. C5 partially disagreed, stating that *"Reflex, as it was presented, required the same manipulative experience and, given the added motivation provided by the digital contents on the tablet, she was sure that it would have been welcomed with great enthusiasm"* [pro-RM]. When referring to the study, C1 agreed to C5 previous statement by confirming that especially for P2, P11 and P12 - who are guests that she is usually taking care - *"Reflex would have been very powerful in engaging them since the beginning"* [pro-RM]. C7 agreed on it by adding as an example that P9 *"was able to focus on the given digital instruction, he performed the assigned task (positioning a piece), and refocused instruction again from the very first time"* [pro-RM]. On this, C3 then added that *"the information was neither overwhelming nor unpredictable at all! Stimuli from the digital app were happening at the right time"* (i.e. only after the user correctly placed the piece) [pro-TM, pro-RM].

The statements of the easiness of use since the beginning provoked an hype during the group interview, mitigated by the assistant asking to recall and explain the exact moments when a particular insight came up. C3 recalled that in the case of P7, P13 and P14 *"were able to adopt their own strategy to come up with the final figure, and this happened more freely with paper and reflex modes"* [pro-PM, pro-RM]. C7 agreed on that by looking at participants shared/joint attention *"in both reflex and paper mode they were noticing us too. In tablet they were only focusing on the device itself"* [pro-PM, pro-RM].

All specialists attending the study sessions recognized that motivation and engagement play a critical role in the study. In particular, C4 referred to the Reflex reward as *"the small Einstein was great! His voice was enough emotional to communicate happiness and enough flat to be understood"* [pro-RM].

Regarding the eventuality of adopting phygital tools for regular social care interventions, the care manager C8, explained that *"current government resources available to social care services do not usually cover the costs of these devices"* [pro-PM] and that with their few funds they always *"favored the choice of paying of an extra caregiver"*. C7, with her experienced role in the center, clarified that the center was applying for funds to be able to afford digital devices but that *"currently we are relying on families to give their children a personal device for interventions"* such as a tablet. C6 acknowledged C7 and C8 thoughts and continued explaining that the center favors paper-based materials because they are *"easy to prepare, easy to maintain and especially very cheap!"* [pro-PM]. No other specialist decided to add on these statements.

8 Discussion

8.1 Potential of using a phygital approach

The quantitative results from the empirical study addressed our first and second research questions and indicated that a phygital approach (RM) was effective for people with ID in a social care center both from a performance and a likability perspective.

Although *User Performance* increased in all interaction modalities, this improvement was higher in Reflex Mode, suggesting that phygital approaches might have a stronger potential than other interaction modes to improve the performance of people with ID. In addition, results indicate lower values on *User Performance* and *User Likeability* for the Tablet mode compared to Reflex and Paper Modes.

Indeed, these quantitative results are in line with the comments and perceptions reported by the therapists. They considered this mode too demanding for people with ID because of the difficulty in moving and rotating virtual shapes on the touch-screen. An additional complexity of the Tablet Mode could be ascribed to the fact that the instruction zone, play zone, and deck zone are merged together in the tablet screen. Compared with RM and PM (in which the three areas have clearly distinguished spatial allocations) this merging does not help the user to identify the specific cognitive tasks involved in the Tangram activity, and might require participants to make a stronger cognitive effort.

Additional findings emerged from qualitative and quantitative analysis were categorized and discussed below.

- *Low Accessibility barriers.* The empirical study confirmed the low accessibility barrier that multimodality provided for persons with ID, as well as supported theoretical research on tangible and digital interaction. Reflex and phygital approaches in general provide a combination of different modalities of representations, thus engaging visual-auditory and motor senses in interaction. Specialists often reported Reflex compliance in giving participants, despite their particular needs, the chance to participate in whichever manner they were able to.
- *Preserved Likeability.* We observed that a hybrid interaction method (RM) can be as effective as a classical method (PM) in terms of performances and more pleasurable in terms of interaction. The majority of participants significantly scored RM as their preferred interaction modes (*User Likeability*) in almost all sessions. As far as we are concerned, the supposed reasons for RM success are linked to a combination of direct-objects manipulation to an interactive feedback and reward.

8.2 Adoption of phygital approaches into a social care context

The group interview and later discussion helped us in answering to the third research question by supposing some potential outcomes for the adoption of phygital approaches into a social care context. Outcomes from the interview were categorized as *pro-PM* (in favour of Paper Mode), *pro-RM* (in favor of Reflex Mode), *pro-TM* (in favor of Tablet Mode), *comparable* (equally favouring RM and PM).

Pro-PM outcomes were:

- *Cost:* paper-based tools are widely adopted in therapeutic centers due to their very low costs. Downloading images over the internet, photocopying books and transforming

them into learning materials (even plasticizing for better durability) is a daily routine in therapeutic centers. The buying cost of these raw materials is around 200\$ per year;

- *Replace-ability*: in case of damage, paper-based materials are easily replaceable with new ones;
- *Contemporaneity*: due to the low costs, many activities can be replicated and contemporary played with different users;
- *Tempestivity (Planning)*: when setting up an activity, specialists spend time in: planning the session, preparing materials, supervising the subject, and analyzing results. Planning a session with paper-based materials takes less time than using the digital tools available in RM and TM; while supervising the participant takes the same amount of time regardless the interaction mode.

Pro-RM insights were:

- *Portability*: whenever specialists are not available in the residential environment to check and correct the users' action, Reflex has the potential to be incorporated in it due to its portability and low cost (200\$ low-end tablet + 50\$ printed materials);
- *Usability*: specialists expected PM to be the method with the highest *User Performance* because it necessitates a “*natural and more common use*” and, as a psychologist noted, “*participants were required just to look at the image and reproduce it*”, highlighting an easier task - compared to TM and RM - both under a cognitive and a motor perspective.
- *Monitoring Capability*: tracking the results achieved during intervention is an important issue for therapists to monitor the evolution of the person with ID and tune future interventions. While in the paper performance mode data are collected manually, in the tablet and Reflex games they are logged automatically in real-time, allowing the analysis of these data ex-post, as noticed by a specialist, C4: “*Collecting results of activities with paper-based materials takes more time than using RM and TM, and their analysis is easier*”
- *Tailorability*: performance data are logged automatically during the training (also at home), allowing customization of parameters as a remote follow-up.
- *Tempestivity (Preparation and Analysis)*: preparation time is significantly less (15 seconds for RM and TM) than the PM (some minutes to find the already prepared materials and organize them). It takes nothing to gather and visualize results, while it takes almost 15 minutes to annotate and process scores in the PM mode;
- *Autonomy*: Reflex gave a level of autonomy that did not establish in any other previous study with neither paper-based nor tablet-based activities. Some users, differently than specialists belief, were able to start, perform and end a Reflex session autonomously without any help;
- *Mappability*: direct manipulation with the system provided by RM facilitated direct mappings between the meaning or semantics of the representing world (the Tangram figures) and the represented world (the resulting output on the tablet screen) as happened in [42];
- *Rewardability*: many participants were stimulated and encouraged by RM (and also TM) feedbacks and, at the same time, got to experience, by direct manipulations of a physical object (which is itself a reward), a concrete and operative performance, more than purely interactive interaction.

9 Limitations

The aim of this study was to investigate the use of a phygital approach into a social care center to support people with Intellectual Disability. Although the research has reached its aims, there were some unavoidable limitations.

First, the data are self-reported introducing several potential sources of bias such as selective memory; telescoping; attribution; exaggeration. The study procedure were mostly organized before performing the study and commented before the results publication. Still we had to face, as common in any study, different unexpected circumstances that may have induced specialists and researchers to hypothesize some facts after the study (known as harking [31]).

Second, this research was conducted on a short time range (6 weeks). Therefore, to generalize the results for larger groups and empirically evaluate the acquired learning skills, the study should be replicated for a longer period. Within a longer time frame, running similar sessions with the same participants would allow analyzing post-novelty effects, analysis and generalization of effects per pathology and consolidation of learning.

Third, although the majority of the empirical studies in HCI field included small samples in their works [7], small samples prevent us from making strong conclusions about the general populations of interest, which is a limitation of both the extent literature and the authors' current study.

Fourth, as technology limitation, we found out during the study that RM and TM introduced some small delays that could have distorted the experience of our participants. On the other hand, the multimodality and multisensoriality of Reflex and tablet-based interaction modes, eventually influenced participants' focus of attention and related performances.

Finally, especially for the most performing participants, given the limited variability of items and shapes, Tangram activity might have had implicitly affected the pleasure to play and so the *User Likeability* at a degree that we are not able to establish.

10 Conclusion and future works

This paper presented an empirical study aimed at investigating the potential of phygital technologies for people with Intellectual Disability (ID) during social care center interventions.

We co-designed [55] and developed a tool that make use of a phygital approach named Reflex, taking inspiration from a commercial system [59]. The proposed approach involved persons with ID through phygital experiences that we have only started to explore but have already highlighted its potential.

The empirical study performed at a social care center provided a number of insights on the benefits of use of phygital activities compared to the use of paper-based only and digital only modalities, and unveils some potential points of strengths and weaknesses of each of the three modalities.

Our comparative approach enabled us to dig more deeply on the characteristics of phygital interfaces that might affect the explored variables - user performance, user likeability, and adoption factors, paving the ground for further research on the design of phygital interfaces for persons with ID.

Our findings indicated better (and similar) results on performance for the tools involving physical manipulation - both paper-only materials and Phygital Interface, compared to digital only tools on tablet. The Phygital Interface also scored highest in likeability. These

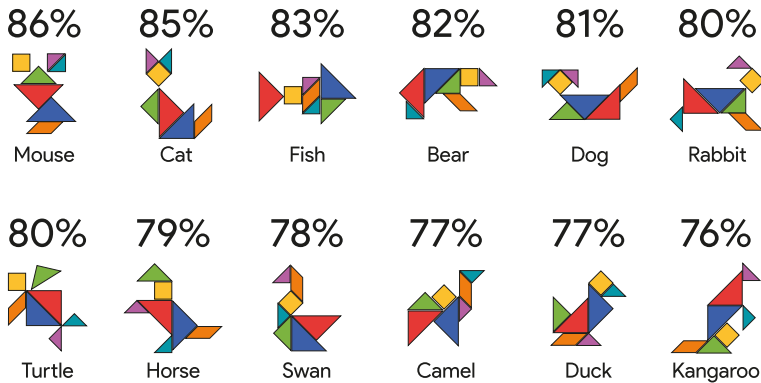


Fig. 19 Shapes by *User Performance* in reflex mode

results are in line with past research that highlights the role of embodiment in learning process and with other comparative studies (see Section 3) in the TUI arena that pinpoint the potential of TUIs to promote engagement and fun.

This work also contributed in understanding the potential of adopting phygital technologies during intervention into a real setting under different logistical and economical lenses.

The outcomes elicited from caregivers confirmed these results, offered illuminating insights on them, and highlighted additional factors (such as low-cost and configurability) for the adoption of phygital interfaces.

The study triggered a beneficial side effect at the social care center: participants of the studies started spreading the voice and telling about their activities to the other care attendants, and even outside the center, so that other people asked to join the study. In addition, other caregivers, besides the one directly collaborating with us, declared their availability to participate in future studies.

Our next steps are in various directions. Reflex was able to derive the best and worst performed shapes as depicted in Fig. 19.

Starting from these results, we are working on the integration of machine learning approaches that exploit the progressively increasing amount of data automatically gathered by the system both for diagnostic purposes and to support self-adaptation of the interactive experience. Secondly, we are adding stronger content management features to the customization tool and expanding the platform with new experiences (e.g., 3D physical objects). Finally, we are planning a wider and longer controlled empirical study to provide more rigorous evidence of the proposed promising benefits.

Acknowledgements Several people helped the authors during the design, development of evaluation of Reflex. Of them, the following are worth to be mentioned.

Thanks to the professionals of l'abilità onlus who participated during the very first requirements elicitation of Reflex.

We are grateful to Fabio Calabretta as the developer of the first Reflex version, successfully evaluated in previous studies [21, 22, 55].

Special thanks to Eleonora Beccaluva (psychologist) who acted as a bridge between Politecnico di Milano and Fratemità & Amicizia, and helped researchers to clarify specialists requirements and participants needs.

Our gratitude goes also to Leonardo Viola for the help received during the the administration of the experimental study.

Finally we warmly thank the specialists, the persons with ID and their families from *Fraternità & Amicizia*, without them this research would never have existed.

Funding Open access funding provided by Politecnico di Milano within the CRUI-CARE Agreement.

Availability of data and material Data and materials are freely available on GitHub at the link <https://github.com/MirkoGelsomini/ReflexPhyigitalApproach>.

With these data, any researcher will be able to run any other type of statistical analysis.

Code Availability The source files of the wizard of oz version are freely available on GitHub at the link <https://github.com/MirkoGelsomini/ReflexPhyigitalApproach>.

With these files, any researcher will be able to run studies with the same approaches we proposed.

Declarations A consent form, dealing with the use of data and images recorded during the study, was signed by both parents or legal tutor of each participant involved. Their information were treated with the highest level of confidentiality, storing only the participant id and related tests results. Personal data were processed only to the extent that they were indispensable in relation to the objective of the study, in compliance with the provisions of current legislation on the protection of personal data and in accordance with the provisions of the general authorizations of the Guarantor Authority for the protection of personal data.

Conflict of Interests The authors have internally developed the proposed phyigital solution called Reflex. Although the studies have been externally administered, potential biases could have played a role in the analysis of data. These limitations have been addressed in Section 9.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. American PA, et al. (2013) Diagnostic and statistical manual of mental disorders (DSM-5®). American Psychiatric Pub
2. Antle AN (2007) The cti framework: informing the design of tangible systems for children. In: Proceedings of the 1st international conference on tangible and embedded interaction. ACM, pp 195–202
3. Antle AN, Wise AF (2013) Getting down to details: Using theories of cognition and learning to inform tangible user interface design. *Interact Comput* 25(1):1–20
4. Beaudouin-Lafon M (2000) Instrumental interaction: an interaction model for designing post-wimp user interfaces. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 446–453
5. Bell BS, Kozłowski SWJ (2008) Active learning: effects of core training design elements on self-regulatory processes, learning, and adaptability. *J Appl Psychol* 93(2):296
6. Bernardo B, Alves-Oliveira P, Santos MG, Melo FS, Paiva A (2016) An interactive tangram game for children with autism. In: International conference on intelligent virtual agents. Springer, pp 500–504
7. Borg J, Lantz A, Gulliksen J (2015) Accessibility to electronic communication for people with cognitive disabilities: a systematic search and review of empirical evidence. *Univ Access Inf Soc* 14(4):547–562
8. Caruso F, Arquilla V, Gelsomini M (2020) Yogo: an hybrid toy-system for children with dcd. In: Fourth European Tangible Interaction Studio 2020 (ETIS)
9. Cerezo E, Marco J, Baldassarri S (2015) Hybrid games: designing tangible interfaces for very young children and children with special needs. In: More playful user interfaces. Springer, pp 17–48

10. Chipman G, Fails JA, Druin A, Guha ML (2011) Paper vs. tablet computers: a comparative study using tangible flags. In: Proceedings of the 10th international conference on interaction design and children, pp 29–36
11. Dandashi A, Karkar AG, Saad S, Barhoumi Z, Al-Jaam J, Saddik AE (2015) Enhancing the cognitive and learning skills of children with intellectual disability through physical activity and edutainment games. *Int J Distrib Sensor Netw* 11(6):165165
12. Dulcan MK, Wiener JM (2006) Essentials of child and adolescent psychiatry. American Psychiatric Pub
13. Eisenberg M, Eisenberg A, Hendrix S, Blauvelt G, Butter D, Garcia J, Lewis R, Nielsen T (2003) As we may print: new directions in output devices and computational crafts for children. In: Proceedings of the 2003 conference on Interaction design and children. ACM, pp 31–39
14. Falcão TP (2017) Action-effect mappings in tangible interaction for children with intellectual disabilities. *Int J Learn Technol* 12(4):294–314
15. Falcão TP, Price S (2012) Tangibles for students with intellectual disabilities. In: Proceedings of the 11th international conference on interaction design and children. ACM, pp 371–374
16. Fianga S (2014) Tangram game activities, helping the students difficulty in understanding the concept of area conservation paper title Proceeding of international conference on research, implementation and education of mathematics and sciences
17. Foglia L, Wilson RA (2013) Embodied cognition. *Wiley Interdiscip Rev Cogn Sci* 4(3):319–325
18. Fröhlich A (2015) La stimolazione basale: per bambini, adolescenti e adulti con pluridisabilità Editrice La Scuola
19. Frutos-Pascual M, García-Zapirain B, Méndez-Zorrilla A (2012) Improvement in cognitive therapies aimed at the elderly using a mixed-reality tool based on tangram game. In: Computer applications for graphics, grid computing, and industrial environment. Springer, pp 68–75
20. Garzotto F, Bordogna M (2010) Paper-based multimedia interaction as learning tool for disabled children. In: Proceedings of the 9th international conference on interaction design and children, pp 79–88
21. Gelsomini M (2018) Reflex: learning beyond the screen in a simple, fun, and affordable way. In: Extended abstracts of the 2018 CHI conference on human factors in computing systems, pp 1–6
22. Gelsomini M, Spitale M, Becaluva E, Viola L, Garzotto F (2019) Reflex: adaptive learning beyond the screen. In: Proceedings of the 24th international conference on intelligent user interfaces: companion, pp 29–30
23. Gibbs GR (2007) Thematic coding and categorizing. *Analyzing Qualitative Data* 703:38–56
24. Hagerman RJ, Polussa J (2015) Treatment of the psychiatric problems associated with fragile x syndrome. *Current Opinion in Psychiatry* 28(2):107
25. Herrero L, Theirs CI, Ruiz-Iniesta A, González A, Sanchez V, Pérez-Nieto MA (2019) Visuospatial processing improvements in students with down syndrome through the autonomous use of technologies. *Br J Educ Technol* 50(4):2055–2066
26. Hornecker E, Buur J (2006) Getting a grip on tangible interaction: a framework on physical space and social interaction. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 437–446
27. ICD-11. <https://icd.who.int/browse11/>
28. Jones B, Kenward MG (2014) Design and analysis of cross-over trials. CRC Press, Boca Raton
29. Joosten AV, Bundy AC (2010) Sensory processing and stereotypical and repetitive behaviour in children with autism and intellectual disability. *Australian Occupational Therapy Journal* 57(6):366–372
30. Jordà S, Kaltenbrunner M, Geiger G, Bencina R (2005) The reactable. In: ICMC
31. Kerr NL (1998) Harking: Hypothesizing after the results are known. *Pers Soc Psychol Rev* 2(3):196–217
32. Kientz JA, Hayes GR, Goodwin MS, Gelsomini M, Abowd GD (2019) Interactive technologies and autism. *Synthesis Lectures on Assistive, Rehabilitative, and Health-Preserving Technologies* 9(1):i–229
33. Kostanjsek N (2011) Use of the international classification of functioning, disability and health (icf) as a conceptual framework and common language for disability statistics and health information systems. In: *BMC public health*, vol 11, p S3. BioMed Central
34. Kremelberg D (2010) Practical statistics: A quick and easy guide to IBM®/SPSS® Statistics, STATA, and other statistical software. SAGE publications
35. Levy S (2017) Graphical user interface (gui). *Encyclopædia Britannica*. Accessed February 16
36. Mahmud AA, Soysa AI (2020) Poma: A tangible user interface to improve social and cognitive skills of sri lankan children with asd. *Int J Human-Computer Stud*, p 102486
37. Marco J, Baldassarri S, Cerezo E (2013) Nikvision: Developing a tangible application for and with children. *J UCS* 19(15):2266–2291
38. Marshall P (2007) Do tangible interfaces enhance learning? In: Proceedings of the 1st international conference on Tangible and embedded interaction, pp 163–170

39. Md M, Abidin R, et al. (2018) Tangram game software in helping understanding of geometric concepts among pre school children. *Adv Sci Lett* 24(7):5338–5341
40. Menendez C, Marques-Villarroya S, Castillo JC, Gamboa-Montero JJ, Salichs MA (2020) A computer vision-based system for a tangram game in a social robot. Springer, pp 61–71
41. Olesen J, Gustavsson A, Svensson M, Wittchen H-U, Jönsson B, CDBE2010 S. G., European B. C. (2012) The economic cost of brain disorders in Europe. *European Journal of Neurology* 19(1):155–162
42. O'Malley C, Fraser DS (2004) Literature review in learning with tangible technologies
43. Petrelli D, O'Brien S (2018) Phone vs. tangible in museums: A comparative study. In: CHI, p 112
44. Renavitasari IRD, Supianto AA (2018) Educational game for training spatial ability using tangram puzzle. In: 2018 international conference on sustainable information engineering and technology (SIET). IEEE, pp 174–179
45. Resnick M (1998) Technologies for lifelong kindergarten. *Educational Technology Research and Development* 46(4):43–55
46. Resnick M, Martin F, Berg R, Borovoy R, Colella V, Kramer K, Silverman B (1998) Digital manipulatives: new toys to think with. In: Proceedings of the SIGCHI conference on Human factors in computing systems, pp 281–287
47. Resnick M, Martin F, Sargent R, Silverman B (1996) Programmable bricks: Toys to think with. *IBM Systems Journal* 35(3.4):443–452
48. Sánchez-Morales A, Durand-Rivera JA, Martínez-González CL (2020) Usability evaluation of a tangible user interface and serious game for identification of cognitive deficiencies in preschool children. *Int J Adv Comput Sci Appl*: 486–493
49. Schalock RL, Borthwick-Duffy SA, Bradley VJ, Buntinx WHE, Coulter DL, Craig EM, Gomez SC, Lachapelle Y, Luckasson R, Reeve A, et al. (2010), Intellectual disability, definition, classification, and systems of supports. ERIC
50. Schneider B, Jermann P, Zufferey G, Dillenbourg P (2011) Benefits of a tangible interface for collaborative learning and interaction. *IEEE Trans Learn Technol* 4(3):222–232
51. Sharma PK, Scholler J, Conrad Verkler. (2017) Activity surface detection, display and enhancement of a virtual scene. US Patent App. 15/081,739
52. Sitdhisanguan K, Chotikakamthorn N, Dechaboon A, Out P (2008) Evaluation the efficacy of computer-based training using tangible user interface for low-function children with autism. In: 2008 Second IEEE International conference on digital game and intelligent toy enhanced learning. IEEE, pp 70–74
53. Sitdhisanguan K, Chotikakamthorn N, Dechaboon A, Out P (2012) Using tangible user interfaces in computer-based training systems for low-functioning autistic children. *Pers Ubiquit Comput* 16(2):143–155
54. Song K, Lee S, Pyun S-B, Kim L (2016) Comparative study of tangible tabletop and computer-based training interfaces for cognitive rehabilitation. In: International conference on universal access in human-computer interaction. Springer, pp 414–424
55. Spitale M, Gelsomini M, Beccaluva E, Viola L, Garzotto F (2019) Meeting the needs of people with neuro-developmental disorder through a phygital approach
56. Starcic AI, Cotic M, Zajc M (2013) Design-based research on the use of a tangible user interface for geometry teaching in an inclusive classroom. *Br J Educ Technol* 44(5):729–744
57. Tam V, Gelsomini M, Garzotto F (2017) Polipo: a tangible toy for children with neurodevelopmental disorders. In: Proceedings of the tenth international conference on tangible, embedded, and embodied interaction. ACM, pp 11–20
58. Tangible P. (2014) Tangram-osmo
59. Tangible P. (2016) Osmo, award-winning educational games system for ipad
60. Ullmer BA, Ishii H (2002) Tangible interfaces for manipulating aggregates of digital information. PhD thesis. Citeseer
61. Urturi ZS, Frutos-Pascual M, Zorrilla AM, Zapirain BG Natural or tangible user interface: Measuring the elderly user experience with a tangram based tool
62. Vicente KJ, Wang JoAnneH (1998) An ecological theory of expertise effects in memory recall. *Psychol Rev* 105(1):33
63. Wilson M (2002) Six views of embodied cognition. *Psychonomic Bulletin & Review* 9(4):625–636
64. Zajc M, Starcic AI (2012) Potentials of the tangible user interface (tui) in enhancing inclusion of people with special needs in the ict-assisted learning and e-accessibility. In: KES international symposium on agent and multi-agent systems: technologies and applications. Springer, pp 261–270

65. Zhang L, Warren Z, Swanson A, Weitlauf A, Sarkar N (2018) Understanding performance and verbal-communication of children with asd in a collaborative virtual environment. *Journal of Autism and Developmental Disorders* 48(8):2779–2789
66. Zuckerman O (2010) Designing digital objects for learning: lessons from froebel and montessori. *Int J Arts Technol* 3(1):124–135
67. Zuckerman O, Arida S, Resnick M (2005) Extending tangible interfaces for education: digital montessori-inspired manipulatives. In: *Proceedings of the SIGCHI conference on Human factors in computing systems*, pp 859–868

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.