ToGather, an interactive website for the stakeholders of school inclusion of children with ASD: an iterative design including user testing

Isabeau Saint-Supery, Cécile Mazon, Eric Meyer, Hélène Sauzéon

ToGather, an interactive website for the stakeholders of school inclusion of children with ASD: an iterative design including user testing

General information

Authors

I. Saint-Supery[1, 2]; H. Sauzéon[1, 2]; E. Meyer[2]; C. Mazon[1, 2]

isabeau.saint-supery@inria.fr; cecile.mazon@inria.fr; eric.meyer@inria.fr; helene.sauzeon@inria.fr

1: Team Ageing, Chronic diseases, Technology, dIsability, and enVironmEnt (ACTIVE), Université de Bordeaux, Inserm, BPH, U1219. Site Carreire, zone Nord, bat. 1B. 146 rue Léo Saignat, 33076, Bordeaux, Gironde, Nouvelle-Aquitaine, France

2: Project team Flowers, Inria Bordeaux Sud-Ouest, 200 Avenue de la Vieille Tour, 33405, Talence Cedex, Gironde, Nouvelle-Aquitaine, France

Corresponding authors:

Correspondence concerning this article should be addressed to Isabeau Saint-Supery or to Cécile Mazon,

Équipe-projet Flowers, Centre Inria de l'Université de Bordeaux, 200 Avenue de la Vieille Tour, 33405 Talence cedex, France

Email: isabeau.saint-supery@inria.fr; cecile.mazon@inria.fr

Abstract

School inclusion of children with Autism Spectrum Disorder (ASD) represents a challenge for all parties involved (family, school and out-of-school actors such as medico-social professionals). To the best of our knowledge, existing tools focus either on the pupil, or the relationships between two parties, always including parents. Considering a systemic approach and knowing that collaboration among stakeholders of school inclusion allows the latter to succeed, we decided to create a tool to enhance and facilitate communication and information sharing among them. For a better usability and to ensure satisfaction of the need expressed by those stakeholders, we used participatory design to create the web-based app *ToGather*. We used (fictitious) realistic cases to create scenarios that contain tasks to do on the web application, adapted to each participant role (parent, teacher and external actors). We used objective (interaction annotations, website navigation inspection) and subjective (questionnaires) measures to assess the design of *ToGather*, with regards to usability, user experience, cognitive load and elicited self-determination. Results indicated that *ToGather* was very usable, with an excellent user experience, and a relatively low cognitive demand. Some design mistakes have been identified and corrected for the enhanced final release.

Keywords

- Special needs education
- Media in education
- Human-computer interface
- Secondary education

1. Introduction

Autism Spectrum Disorder (ASD) is a neurodevelopmental condition, affecting in average 65 persons over 10,000 in the world (Zeidan, Fombonne, Scorah, Ibrahim, Durkin, Saxena, *et al.*, 2022), and a prevalence of 36.5 over 10,000 persons in France (Fombonne, Myers, Chavez, Presmanes & Zuckerman, 2019) with impairments in communication, social skills, and restricted interests and activities (American Psychological Association, 2013). There is a large inter-individual variability in the nature and intensity of ASD appealing to hugely personalize care and support, often running counter to the "one size fits all" functioning of mainstream schools (Fage, Consel, Etchegoyen, Amestoy, Bouvard, Sauzéon, *et al.*, 2016). Today, there is a relative consensus regarding the short and long-term benefits of early school inclusion for students with ASD, i.e., improved socio-adaptive capabilities (Humphrey & Symes, 2011), decreased Autistic symptoms (Osborne & Reed, 2011), higher social self-esteem (Chevallier, Courtinat-Camps & de Léonardis, 2015), better cognitive, language and social development (Ferraioli & Harris, 2011; Osborne & Reed, 2011; Salceanu, 2020), etc. Despite of these tangible benefits, school inclusion remains challenging for children with ASD. For instance, in France, for the first-degree school[1], inclusion is around 66% but collapse considerably for the second degree with 32.85% (Dubois & Jumel, 2019; French Ministry of National Education, Youth and Sports, 2019). Echoing the systemic view of disability (International Classification of Functioning - children and youth version, ICF-CY, WHO, 2001, 2007), this inclusion gap from first-to second degree is explained by both internal and external factors such as the ASD severity as well as the comorbidities (Intellectual deficiency, dys- syndromes) for internal factors, and family and social environments' constraints for external factors (Hodges, Joosten,

[1] French school is divided in three degrees: the first one includes elementary and primary schools (from 3 to 10 years-old); the second includes middle and high schools (from 11 to 17 years-old), and the third one corresponds to higher-education level.

Bourke-Taylor & Cordier, 2020; Stephenson, Browne, Carter, Clark, Costley, Martin, *et al.*, 2021). As a result, intervention research seeks to find levers for bypassing these internal and external barriers (De, Chakraborty, Bhattacharya, Bhunia & Basu, 2018). In this vein, computer-based education for special learning needs has a growing development for addressing challenges related to internal factors, i.e., traditional person-technology match approach (*Mazon, et al., 2019*; Odom, Thompson, Hedges, Boyd, Dykstra, Duda, *et al.*, 2015; Roberts-Yates & Silvera-Tawil, 2019; Suhaila & Nordin, 2022). In contrast, researches addressing the external factors are less numerous (i.e., peers collaborative technology supports; technologies for ASD awareness and anti-stigma, etc.) or non-existent regarding the collaborative issues of proximal social environments of children including family, school and care (Vlcek, Somerton & Rvner, 2020). This reveals that it is yet difficult to move from a model where disability is the matter of the individual to a model where disability is the matter of all, but also the complexity of having multiple end-users for a given technology. From a previous analysis of co-education needs of stakeholders of schooling children with ASD leading to the co-designing of a new web app, i.e., *ToGather;* (*Mazon, et al., 2021*), the present study aims to first, draw up an original method based on realistic students cases for a three-parties-user assessment (parents, teachers and caregivers) with both objective (interaction data) and subjective metrics (standardized questionnaires), and secondly, to provide results from the three parties assessment for extracting relevant and to-be-improved features of *ToGather* app.

1.1. Related work

Due to the heterogeneity and severity of ASD symptoms, various educative adaptations (i.e., pedagogical adaptations, compensatory tools, specific teaching) are required for meeting the pupils' specific needs for their school inclusion. In France, if the personalization need for education is too great students are enrolled in specialized device but can be included in regular classes for a defined period of time (from 1 hour per week to the whole). In many countries,

these adaptations are gathered in the Individual Education Plan (IEP) paper document, which is also a monitoring tool throughout the schooling process (i.e., French Ministry of National Education, Youth and Sports, 2015, 2020). The IEP is critical by gathering all the educational levers for ensuring as best as possible a successful achievement of the child's educational project. It should therefore be as accessible and understandable as possible to teachers and all those involved in school inclusion, and should be updated as and when the student's progress, life events, or preferences and choices in terms of life plans are made. IEP updating over time is a byproduct of process of co-education and mobilizes collaboration and close links between the family and the school, as well as the care services the student receives. Unfortunately, there are several barriers to co-education process. Without listing them exhaustively, most of them are related to the structural design of schools (tailored services and logistical issues) and or to the school-specific social environment (i.e., Coster, Law, Bedell, Liljenquist, Kao, Khetani, *et al.*, 2013) which together negatively influence the child's feeling of school connectedness (Saggers, Klug, Harper-Hill, Ashburner, Costley, Clark, *et al.*, 2016), i.e., "the extent to which students feel personally accepted, respected, included, and supported by others in the school social environment" (p.80, Goodenow, 1993).

First, *a cultural information gap on ASD* is widely documented in studies of the school's social environment (Kinnear, Link, Ballan & Fischbach, 2016; Mitter, Ali & Scior, 2019; Roberts & Simpson, 2016). Whether it is the teacher or the peer students, a lack of knowledge about ASD is noticeable (i.e., Fage, Moullet, Consel & Sauzéon, 2017; Gómez-Marí, Sanz-Cervera & Tárraga-Mínguez, 2021). They have no idea on how to interact with a child with ASD or how to react in case of repetitive behaviors for example. Stereotypes on ASD and disabilities still remains a barrier to an inclusive school where teachers who have no experience with ASD believe it is challenging to welcome a child with ASD by considering them unable to remain autonomous, or worse, being sure that the pupil will create some sort of chaos in the

class (Alexandri, Papailiou & Nikolaou, 2017; Paraskevi, 2021). This is even more true for pupils with behavioral issues or multiple difficulties: regular teachers are less incline to include them in comparison with special education teachers (Desombre, Lamotte & Jury, 2019; Noreen, Intizar & Gulzar, 2019). For dealing with ASD-related stigma, several studies leveraged from technologies as training tool for informed representations on ASD and for enhanced empathy and positive attitudes toward ASD (i.e., *Auti-Sim* Intervention, Sarge, Kim & Velez, 2020). However, such anti-stigma interventions are not widely widespread today among teaching professionals and students and therefore the social impact is still limited. In addition, they are very generic about ASD and may not always reflect the heterogeneity of the disorders and the specific needs of a given child with ASD. Hence, a direct communication between school's actors and child' parents or ASD-trained mediator is reported as impacting more positively. Indeed, a trained teacher, an autism association or parent-school interaction can turn out to be facilitators (Stephenson, *et al.*, 2021). For example, encouraging the pupil to work in small groups provides a facilitator to their school inclusion whereas the lack represents a barrier (Stephenson, *et al.*, 2021). As reported in (*Mazon, et al., 2021*), to empower their mission teachers would like the child with ASD to be delivered with an instruction manual containing their interests, relaxing tips, description of their stereotypical behaviors, learning specificities, etc.

Second, *the time availability of parents or teachers* is also an issue in co-education. Parents often take on the role of coordinating their child's schooling, which is costly in terms of time and energy when combined with daily parental activities (work, household chores, mental workload, etc.) (Behnia, Rassafiani, Nakhai, Mohammadpour & Ahmadi Kahjoogh, 2017; Budi Santoso, 2021). As the level of education is a major factor to accomplish a self-determined professional project, it generates stress for the parents until affecting their health and well-being as well as their quality of life (Abolkheirian, Sadeghi & Shojaeizadeh, 2022;

ten Hoopen, de Nijs, Duvekot, Greaves-Lord, Hillegers, Brouwer, *et al.*, 2022). Similarly, teachers face hierarchical pressure by having to do the best they can within a given time and often with a lack of time for training and gathering pedagogical resources (Azad, Marcus, Sheridan & Mandell, 2018; Minke, Sheridan, Kim, Ryoo & Koziol, 2014). Finally, each stakeholder in the educational pathway faces time managing difficulties of their own while tensions are then created between the two parties due to a lack of time or place for communication, often having a direct impact on the child's progress and even success at school (Crowell, Keluskar & Gorecki, 2019; Li, Xu, Wu, Tang, Zhang, Liu, *et al.*, 2022). Indeed, a relationship of trust between parents and other stakeholders facilitates the achievement of the child's educational goals (Bateman, *et al.*, 2022; Garcia-Melgar, *et al.*, 2022; Rispoli, *et al.*, 2019). Therefore, the paucity of continuous school-parent communication opportunities is an impediment of co-education process of children with ASD. Surprisingly, very few studies have focused on the design of digital tools to enrich the parent-teacher relationship in order to foster co-education process (*Mazon, et al., 2021*; Azad, *et al.*, 2018). The mainstream tools of communication between the family and the school are fairly traditional, such as the home liaison diary, social networks or software packages often tailored for a passive curriculum follow-up of students without IEP. Yet, Olmstead (2013) showed that parents and teachers each have positive attitudes towards digital tools (i.e., e-mails, websites, social networks) in the school context, even if they agree on their drawbacks (i.e., time allocated to update sites, to write messages; "invasion" of technologies in everyday life). As a result, there is a lack of a real communication tool for the family-school relationship fitting with the findings of the needs reported by parents and teachers (*Mazon, et al., 2021*) namely, a crucial lack of information sharing on the pupil in terms of consistency and developments with regard to the IEP.

Third, the *multi-sectorial information gap* is also a hindrance to co-education. In numerous countries of OECD, health care professionals involved in the rehabilitation of ASD

(social skills, language, writing, etc.) do not have direct access to the school and parents remain their representatives to explain the disorders and educational needs of the child with ASD (Baker, Galemore & Lowrey, 2020; Shahidullah, McClain, Azad, Mezher & McIntyre, 2020). Once again, several studies revealed the technology interests for promoting the communication between caregivers and parents of children with ASD such as tools to help announce ASD diagnosis and understand its consequences (Gibbons, 2015; Sarge, *et al.*, 2020) and their impact on schooling (*Mazon, et al., 2019*; Pontikas, Tsoukalas & Serdari, 2022) or even to favor care follow-up (*Mazon, et al., 2021*; Shahidullah, *et al.*, 2020; Vlcek, *et al.*, 2020), etc. Care-school relationships would gain in efficiency from sharing a genuine communication place to serve the child schooling while respecting of privacy rules of medical information. Hence, an efficient co-education requires not only parent-teacher dialogue but a three-way (parent, school, care) channel of communication to foster at best the support to the child. Additionally, coordination and cooperation from all persons involved in a process (i.e., care or education) is beneficial regarding the time it takes (Gomes & McVilly, 2019) but also regarding stakeholder's well-being (Lefebvre, Pelchat, Swaine, Gélinas & Levert, 2005), thus makes it a key to successful inclusive schooling of a child with ASD. Involving multiple disciplines to cooperate in following-up a child (Anthony & Campbell, 2020; Choi & Pak, 2006) by creating a technology to enhance the communication between these systems, and considering their members as a whole team, may constitute a new lead to improve the efficiency of inclusive schooling for children with ASD.

Four, *the evolutivity of child's educational needs* is hugely challenging for co-education process. The specificity of children with ASD and the needs generated are constantly evolving (Mintz, Seleznyov, Peacey, Brown & White, 2021; Stepney, 2008), highlighting the need for a three-way, continuous and rapid communication tool. In their recent educational needs analysis amongst a large sample of students with ASD, of parents, of teachers and of specialists, Saggers,

et al., 2016 stressed the crucial role of collaboration and communication about educational needs changes and curriculum adjustments and support. Importantly, they reported that the stakeholders themselves perceived this role as critical to support the school connectedness of child. Additionally, they invoked a better collaboration between them thanks to a whole school approach as success keys. At the best of our knowledge, no technology-based study has addressed the collaborative follow-up of educational needs changes of children with ASD.

Overall, technologies, as they stand, do not support the long-term collaborative monitoring of the pupil and the joint commitment of parents and teachers to support the pupil's educational projects (Olmstead, 2013). Industrial schooling software packages have been designed to meet the needs of neurotypical pupils but are difficult to adapt to the special needs of children with disabilities. They cover some of the requirements regarding the child's educational pathway, but do not, for example, make it possible to observe changes in social skills, since they are usually acquired at primary level (Casartelli, Federici, Fumagalli, Cesareo, Nicoli, Ronconi, *et al.*, 2020).

1.2. A Technology based on the Whole school approach: the *ToGather* app

Moving the educational technology field for children with ASD forward into a whole school approach, *Mazon, et al., 2021* have codesigned with participatory methods (Brajić, Bećirović, Marić, Avdibegovic, Delic & Bećirović, 2012; van der Velden & Mörtberg, 2021) a web-based application promoting stakeholders collaboration for co-education, i.e. *ToGather* apps (*Anonymous label*). Inspired by the ecosystemic model for pupils with ASD (Cappe & Boujut, 2016), this app places the child with ASD at the center of all proximal social environments systems (family, school, care) in respect of child's system is in constant communication with these environments. *ToGather* offers both the child follow-up and the three-way collaboration of the stakeholders through interactive services and which content

always may be updated, regarding (Figure 1): 1 - the pupil's profile, with their strengths and weaknesses; 2 – assessment and follow-up on targeted skills ; 3 – a meeting manager allowing updates (mostly the child's progress) to prepare the official meeting for the schooling follow-up teams; 4 – a shared notebook on efficient psychoeducational solutions for the pupil and 5 – a shared messaging for the pupil's or family's "news".

ToGather application is a breakthrough for co-education because it is supported by an upstream analysis of the multiple end-users' needs and its functionality and features are driven by a participatory design method (Limayem, Hier & Cheung, 2007; Rahimi & Ibarra, 2014). Nevertheless, *ToGather* app has not been documented by an extensive user assessment studying deeply its usability, its user experience as well as the cognitive load elicited by its use. The main reason of that is clearly related to the limitations of available user assessment methods. First, it is ethically questionable to deploy a technology in a real situation with a fragile public and social environment already under strong tension in order to obtain a user field evaluation. Second, by definition, *ToGather* apps involves multiple end-users. It involves multiple end-users, who are certainly motivated by the same co-educational purpose but who take on different functions (parents, teachers, caregivers) and therefore make qualitatively and quantitatively different contributions to co-education depending on the student's educational needs. Thus, standardized evaluation methods with simple web browsing scenarios (Allison, Hayes, McNulty & Young, 2019; Pope, 2009) would not be very informative on the ergonomics value of *ToGather* in terms of usability and user experience. Additionally, an important property that was unanimously praised by all three parties is a low cognitive load to the use of *ToGather* (*Mazon, et al., 2021*) to achieve the co-education process embedded into the app. For all these reasons, a new assessment method based on usage scenarios has been drew up from the creation of three realistic cases of children with ASD varying by their educational needs

(Appendix A). From it, we collected various objective measures of *ToGather* use (interaction data) and traditional standardized subjective measures in terms of usability, user experience and cognitive load.

From the overall data, through a new method based on realistic cases of children with ASD, this study aimed to conduct user testing on the *ToGather* app and examine whether the app does match with the future end-users needs for a collaborative follow-up of school inclusion of children with ASD.

2. Material and methods

We recruited end-users representing the different roles of future users of the *ToGather* application, and asked them to use an interactive mock-up. We designed usage scenarios of the web application, allowing the exploration of the different functionalities it offers. The study was conducted in remote experimentation to maintain ecological conditions and future contexts of app use (home, work places; own computers and web browsers).

2.1. Participants

Recruitment of the participants was ensured in collaboration with institutions and organizations, via an email campaign. According to Nielsen & Landauer (1993), 5 to 10 subjects are the minimum required to avoid making ergonomics design errors. Then, the participants panel was composed of 21 adults (3M; 18F), with 7 parents of pupils with ASD enrolled in middle or high school, 7 specialized teachers, and 7 external school actors involved in the educational inclusion of students with ASD (i.e., specialized educator, psychologist). Participants were aged between 27 and 60 years-old with an average of 44.81 years old (SD = 6.95).

Their consent was collected after they were informed of the aim and conditions of the study. This consent as well as the whole of research purpose and design, have been approved by the

ethical committee of the research center and the CNIL (French national committee securing the individual rights in terms of issues related to person privacy and technology).

2.2. Materials

2.2.1. The *ToGather* mock-up

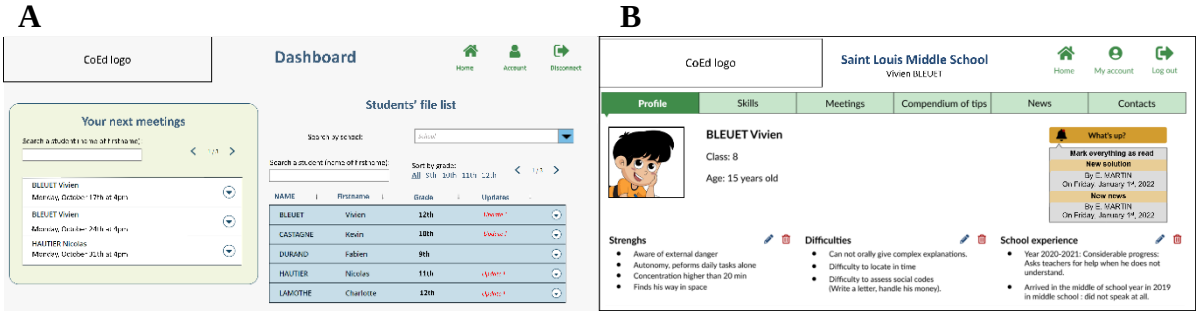

Figure 1: *Dashboard and Profile interfaces of the mock-up used for the user tests*

A: Dashboard interface of the mock-up with a list of student files and a list of the meetings to come for the authenticated user

B: Profile interface, the first tab of a student's file

To ensure that we were compliant with GDPR (European general data protection regulation) during the user testing study, we opted for an interactive mock-up creation of the *ToGather* app. We used Axure RP 9 software, because it provides very realistic mock-up with interactive features that can be shared online with the participants, without cloud processing of interaction data (thus respecting GDPR) while enabling their local collecting and storing via a video-based recording.

The mock-up of *ToGather* app contained the sections listed below as proposed by *Mazon, et al., 2021*. Once authenticated, the user faces a dashboard divided in two blocks; the right one contains a list of student files accessible by the user; and on the left is a list of next meetings regarding those pupils. Each student's file contains the following sections:

1) ***Profile:*** information on the pupil and their particularities on various themes (for example their health, previous schools or center of interest);

2) ***Skills****:* classified in two categories that are school-related and socio-adaptive, skills evolution is displayed in a graph;
3) ***Meetings****:* skills assessment, evaluation matrix, generating meeting report, history of the previous meetings;
4) ***Compendium of tips****:* effective aides, strategies for the child;
5) ***News****:* communication channel for any member of the support team who wants to share a past, present or future event concerning the pupil.
6) ***Contacts***: list of stakeholders and their contact details

Access to these sections can only be guaranteed for authenticated users, which was simulated with a log-in page before accessing any other content. Each user is identified according to their role (i.e., teacher, parent, external) and a referent is designated to regularly monitor the pupil's record and coordinate the stakeholders. Though, the whole support team can add or update any information regarding the child, to place them in equal terms.

For more details on the content of the *ToGather* mock-up please refer to Appendix B.

2.2.2. Usage scenarios based on realistic cases of children with ASD

Three fictitious student profiles, derived from real cases of children who participated in previous studies related to assistive technologies for school (*Anonymized references 1 & 2*). This method is inspired from the persona method classically used in design research to simulate real persons as an important support mechanism in interfaces design (Miaskiewicz & Kozar, 2011). Classically, personas represent, hypothetically, behavior synthesis, main characteristics, motivations and expectations of real users (Antle, 2006; Melo, Rivero, Santos & Barreto, 2020). Here in our study, personas are used not to create empathy between designers and future app end-users but to support a realistic use projection for future end-users. A main reason of that is the diverse audience of children with ASD. The three personas of child with ASD have been selected in our own data base to refer to three kinds of educational needs depending on child's

functioning and ASD severity (for a detailed description of realistic children's cases, see Appendix A).

From these realistic cases, four usage scenarios have been designed for each persona, in order to cover every functionality of the app, with each of the different roles (parents, teachers, health professionals and referent). Participants were given tasks to do on each persona. Personas' order was randomized between participants of the same role. Tasks were attributed depending on the participant's role(s) – teachers and health professionals also had the referent role – and each task was adapted using perspective-taking in order to match what they would have to do or share on a pupil's file based on their role (Wetzels, Kester & van Merriënboer, 2011). Construction of usage scenarios is represented in Figure 2.

Most of the time, one of the tasks necessitated actions from them such as clicking and eventually typing, and the other was simply visually finding and reading information on the displayed page. The different functionalities of the tool were distributed across the three personas (i.e.., finding information, adding information, editing the profile) and the user role (i.e., parent, teacher, external actor). As the role of referent can be ascribed to either a specialized teacher or an external actor, their scenarios contained referent-related functionalities (i.e.., deleting, planning meeting). To prevent any order effect, scenarios' order has been randomized between participants.

See Appendix C for a more detailed description of the actions of each scenario depending on the user role.

2.2.3. Measures

Each testing session was video-recorded to provide different objective use measures. Objective measures of usability were assessed through a video analysis performed by a researcher according to a specific grid to collect the measures of interest, based on the cursor path and

participants oral responses. Kirsh & Joy (2020) showed that using mouse-tracking instead of eye-tracking could be even more relevant for examining webpage exploration. Videos were annotated using ELAN software version 6.1 to extract the time spent per task from the scenario. We recorded participants' response time as from the moment they started to move their mouse until the end of their answer, whether it was correct or not. When the cursor remained steady, we started recording response time when they began to speak or make sound (i.e.., onomatopoeia) (Sears & Jacko, 2009). In parallel, we filled out a spreadsheet to collect data regarding each participants' response (i.e., correctness, number of active interactions [clicking and typing], cursor path and potential comments). From this data, we extracted objective performance measures, in addition to qualitative remarks on the mock-up appearance or functionalities.

Objective measurement – Performance measures were collected through the extraction of video analysis data:

- **Usability measures:** two dimensions of usability were assessed:
    - *Effectiveness*: Assessed using the proportion of correct responses (based on oral responses and active interactions) and the completion time across the different scenario.
    - *Efficiency*: Assessed using the number of active interactions per minute; and an optimality score based on the ratio between participants' speed and a theoretical optimal speed (computed in a within-subject fashion by dividing their maximal number of active interactions by their minimal time to complete a scenario).
- **Website navigation inspection**: the cursor path was annotated as where the participant was looking, in order to get an insight of the website's usability: if the information to look for and the mouse were in the same area (i.e.., heading menu) from the beginning, we could assume that content was properly organized on the page (Kirsh & Joy, 2020). When

participants wandered on the website or gave wrong answers, we noted hovered elements and items clicked in the wrong place, while they were interacting with the website. As participants were doing the asked tasks on the mock-up, the experimenter was writing down participant's comments and/or reactions to gather qualitative data. All those elements are qualitatively analyzed in order to revise the displayed interfaces.

Subjective measurement – Specific and standardized questionnaires were added to complete use measurements and assess user satisfaction

- **Scenario-based Subjective usability**: for each task, the participant was invited to self-assess their interaction difficulty to perform the task with a 10-points Likert scale (from 0 "very easy" to 10 "very difficult").
- **System Usability Scale** (Brooke, 1996; French version, Gronier & Baudet, 2021): The SUS is currently one of the most used scales to assess an interactive system's usability. It consists in 10 items rated using a 5-points Likert scale. The max total score is 100: the higher the score is, the better the usability is.
- **User Experience Questionnaire** (French version, Hassenzahl & Tractinsky, 2006; Laugwitz, Held, Schrepp & Holzinger, 2008). The UEQ aims at evaluating easily and rapidly users' experience on several points. There are 26 items, rated on 7-points Likert scales and gathered in the six following categories: 1) *Attractiveness* (i.e.., Overall impression of the product. Do users like the product or not?); 2) *Perspicuity* (i.e.., How easy is it to learn about the product? Is it easy to learn how to use the product?); 3) *Efficiency* (i.e.., Can users solve their tasks without unnecessary effort?) 4) *Dependability* (i.e.., Does the user feel in control of the interaction); 5) *Stimulation* (i.e.., s using the product exciting and motivating?) 6) *Novelty* (i.e.., Is the product innovative and creative? Does the product engage users?).

Efficiency, Perspicuity and Dependability are part of the task-related or *Pragmatic Quality*, Stimulation and Novelty are part of the *Hedonic Quality*. It is important to evaluate hedonic dimensions of the tool, even if the latter is mainly aimed at the pragmatic side, because user pleasure and satisfaction are equally important in the user experience.

- **NASA-TLX – Task Load Index** (6-items version, French version, Cegarra & Morgado, 2009; Hart & Staveland, 1988): This widely used test assesses the cognitive load required to use a system with 6 workload factors (i.e., mental demand, physical demand, temporal demand, performance, effort, frustration) to rate on a linear scale. Each dimension of the NASA-TLX is rated between 0 and 100. A raw total score (/100) can be computed by averaging the 6 scores. The higher the score is, the higher the cognitive load is.
- **Elicited self-determination for a collaborative work:** it was measured using a custom-built questionnaire inspired from (Vallerand, Pelletier, Blais, Briere, Senecal & Vallieres, 1992). This assessment comprises 9 items to be rated on a 5-points scale ranging from *strongly disagree (0)* to *strongly agree (4)*. The 9 items were distributed into the three main concepts of Self-Determination Theory (SDT) (Ryan & Deci, 2020): *Autonomy (*i.e.*, "I would be more effective in following-up the student/my child by using this tool"*) , *Competence* (i.e.*, "I would have the ability to organize the support of the student/my child with this tool"*) and *Relatedness (*i.e.*, "With this tool, I could better participate in the follow-up of the student/my child")*, and we computed a raw total score by averaging these sub-scores.

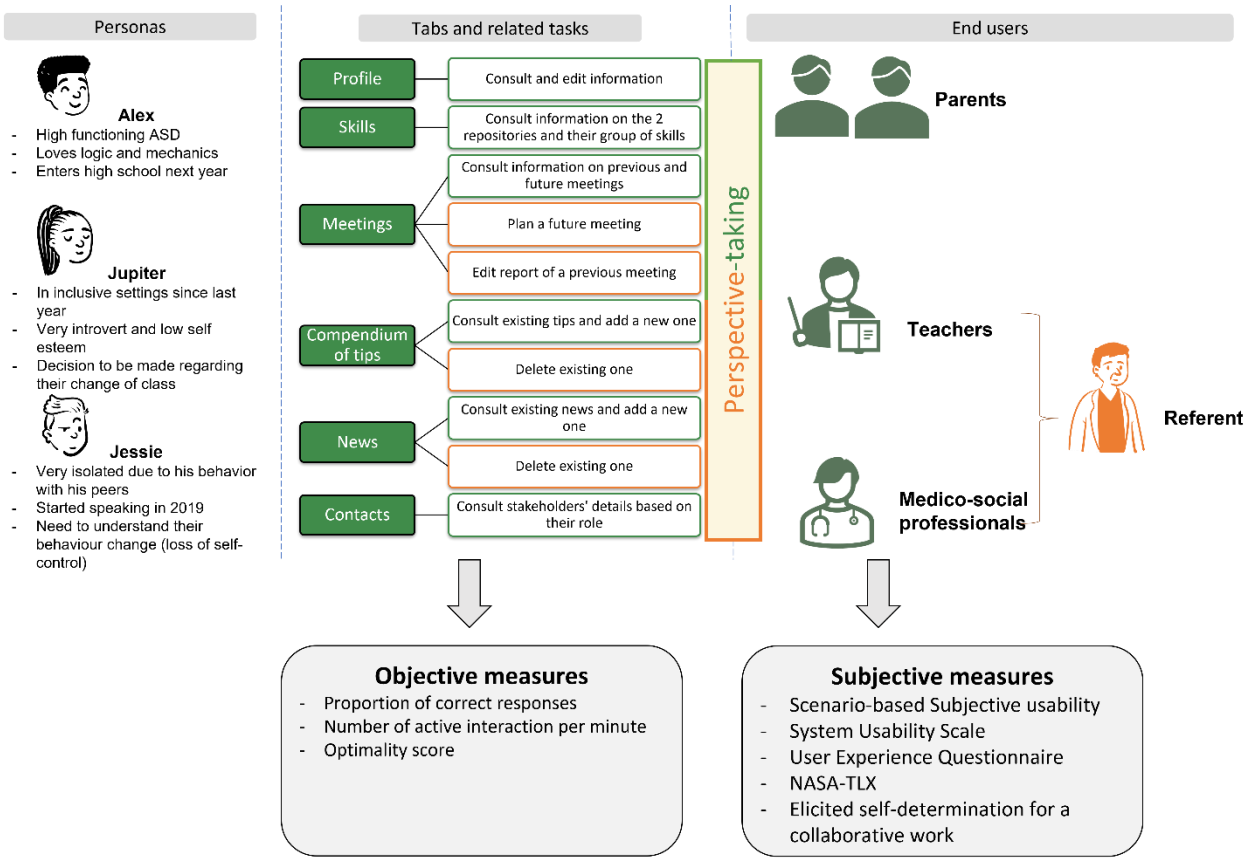

Figure 2: *Scenarios construction based on persona and participant's role, with the associated measures*

2.3. Procedure

The study was conducted between April and June 2020, in remote conditions. Participants and researchers met using videoconference, and the participants shared their screen in order to monitor/record their use of the mock-up. After the researcher reminded the participant of the purpose and process of the study, the participant was invited to connect to the mock-up through the provided URL and begin the first usage scenario. Each scenario began by a description of the fictious pupil, followed by the list of tasks that the participants had to conduct.

Once the participants read the case description, they can begin to use the mock-up following the usage scenario. To avoid back and forth between the scenario file and the mock-up, participants were given oral instructions while they were on the mock-up, with 2 to 3 tasks to do per tab. After the participant finished the three scenarios, they were invited to answer to the questionnaires through an online form. Once completed, researchers asked questions to gather

participants' feedback on the mock-up in a non-directive interview. In average, the full experimental session lasted about 1.5 hours.

The experimenter supervised the participants' performance, and intervened only if the participant was stuck, or when a glitch occurs in the mock-up. Otherwise, the experimenter let the participant perform the tasks by themself, even if the participant performs wrong actions. When participants asked questions, the experimenter answered by eventually giving clues, but was careful not to give the solution. The screen and audio were recorded on the experimenter's computer during the whole experiment, using the software Monosnap (participants were invited to turn off their camera to anonymize the recording).

2.4. Data analysis

Statistical analyses were conducted using R software version 4.2.1.

The video recordings analysis was carried out by one researcher, but two other researchers performed the video analysis on a randomly picked video to ensure inter-rater reliability. Inter-rater agreement was calculated using intra-class correlation, based on the ratings of the three researchers. The intra-class correlation is commonly used to assess the interrater agreement on continuous measures, and is easy to interpret as this coefficient is a correlation (Gisev, Bell & Chen, 2013). The inter-rater agreement on the annotation grid was excellent (r = .968), which means that the annotation grid was clear and valid.

For objective and subjective measures, average scores for each measure were screened in the whole user group to conclude on global usability of the web application. We also explored the differences between each users' role, that is between parents, teachers and medico-social practitioners. As objective measures and subjective difficulty questions were distributed on the three scenarios, we carried out two-way mixed ANOVA, setting the Role as a between-group variable (i.e., parent, teacher, external) and the Time as a within-group variable (i.e., T1, T2, T3) for determining whether usability differs across time (training effects) and depending on

the specific needs of each category of user targeted by the *ToGather* app. In addition, a separate ANOVA was carried out to examine the effect of the Referent's role, as a between-group variable, and the Time as a within-group variable (T1 vs. T3 only). For the result of the questionnaires, we carried out ANOVA with the Role as a between-group variable. When appropriate, the Dimension was added as a within group variable (only for questionnaire composed with several dimensions). For each effect, we computed the η^2 to examine the effect size, according to the following classes (Cohen, 1988): small (η^2<.01), medium (.01 < η^2 < .06), and large (η^2 > .14).

3. Results

Descriptive and two-way ANOVA statistics are depicted in Appendix D for objective measures extracted from annotations' results, for subjective measures extracted from questionnaires' results, and for subjective use measures in the whole group and each category of user.

3.1. Objective Use measures

3.1.1. Effectiveness

Completion time. Overall, the participants took in average 10 min to complete a scenario (range: 5.13min-19.33min). There was an improvement across the 3 scenarios [Time: $F(2,36) = 20.842$; $p < .001$; $\eta^2 = .340$], but there was no difference for the global completion time between users' category [Role: $F(2,18) = 1.396$; $p > .200$; $\eta^2 = .079$]. The interaction effect was not significant. Pairwise t-test revealed significant differences between T1 – T2 [$t(20) = 5.896$; $p_{adj} < .01$] and T1 – T3 [$t(20) = 4.942$; $p_{adj} < .001$]. The results depicted in Figure 3 showed that the completion time decreased over time, in compliance with a large training effect, irrespective of the participant's role. Like the 3 main roles (parent, teacher, external actor), there is no effect for the Referent role [$F(1,18) = 0.052$; $p > .800$; $\eta^2 = .002$].

Training effects between T1 and T3 are found when the teacher's or external actor's role was combined with being Referent [$F(1,18) = 33.404$; $p < .001$; $\eta^2 = .346$].

Proportion of correct responses. Overall, the participants reached 92% of correct responses (range: 77%-100%). There was an improvement across the 3 scenarios [$F(2,36) = 19.467$; $p < .001$; $\eta^2 = .324$], but there was no difference for the global proportion of correct responses between users' category [$F(2,18) = 0.286$; $p > .700$; $\eta^2 = .017$]. The interaction effect was not significant. Pairwise t-test revealed significant differences between T1 – T2 [$t(20) = -3.389$; $p_{adj} < .001$] and T1 – T3 [$t(20) = -5.786$; $p_{adj} < .001$]. The results depicted in Figure 4 showed that the proportion of correct responses increased over time, in compliance with a training effect, regardless the participant's role. Like the 3 main roles (parent, teacher, external actor), there is no effect for the Referent role [$F(1,18) = 0.155$; $p > .600$; $\eta^2 = .006$]. Training effects between T1 and T3 are found when the teacher's or external actor's role was combined with being Referent [$F(1,18) = 25.248$; $p < .001$; $\eta^2 = .327$].

In summary, effectiveness measures showed that participants successfully managed - in terms of correct responses and time per scenario - the scenarios, and improved their performance after completing the first one.

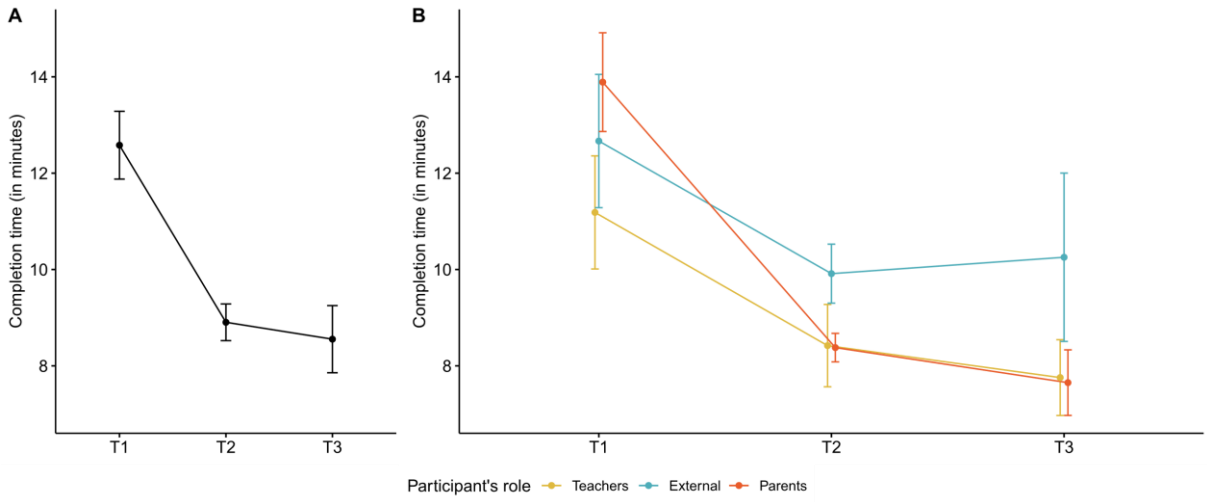

Figure 3: *Completion time (ss.ms) across time in average (A) and differentiated by role (B)*

A: Line chart for the whole group with a significant drop in completion time between T1 and T2, and a slow decrease between T2 and T3

B: Line chart differentiated by role, completion time significantly decreases for each role between T1 and T2. Between T2 and T3, it slowly decreases for Parents and Teachers but increases a little bit for External.

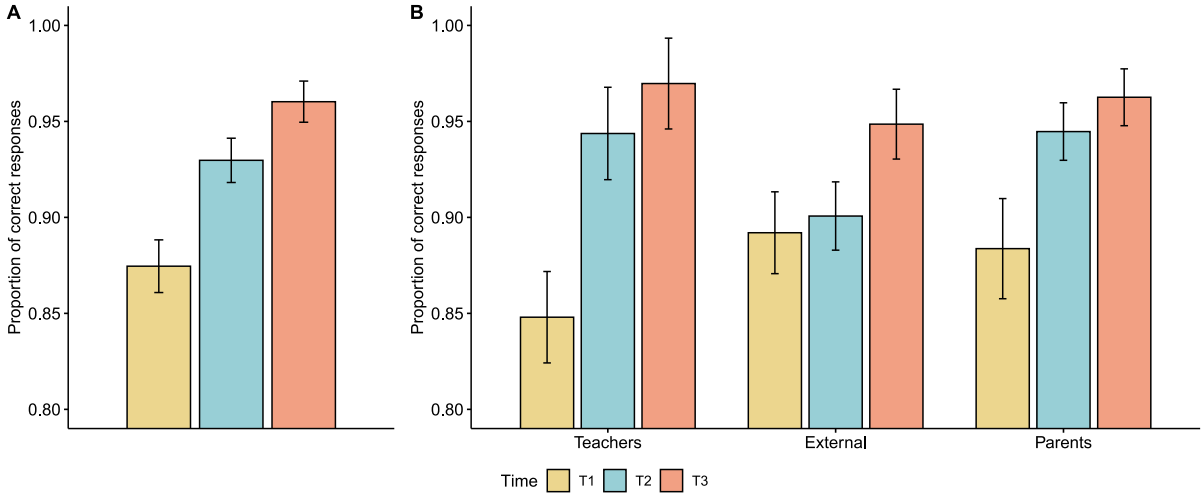

Figure 4: *Proportion of correct responses across time in average (A) and differentiated by role (B)*

A: Histogram for the whole group. In average, the proportion of correct responses increases at each scenario

B: Histograms differentiated by role. Teachers' and parents' proportion of correct responses had a huge increase between T1 and T2, improvements slowed down for T3. This noticeable increase happened between T2 and T3 for external actors, as their progression was low between T1 and T2.

3.1.2. Efficiency

Number of active interactions per minute. Overall, the participants reached in average a speed of 4.754 active interactions per minute (range: 2.010-9.547). There was an improvement across the 3 scenarios [Time: $F(2,36) = 15.545$; $p < .001$; $\eta^2 = .252$], but there was no difference for the number of actions per minute between users' category [Role: $F(2,18) = 2.399$; $p > .100$; $\eta^2 = .140$]. The interaction effect was not significant. Pairwise t-test revealed significant differences between T1 – T3 [$t(20) = -4.536$; $p_{adj} < .001$] and T2 – T3 [$t(20) = -3.858$; $p_{adj} < .005$]. The results depicted in Figure 5 showed that the number of active interactions per minute increased over time, in compliance with a large training effect, despite the participant's

role. Like the 3 main roles (parent, teacher, external actor), there is no effect for the Referent role [$F(1,18) = 2.085$; $p > .100$; $\eta^2 = .078$]. Training effects between T1 and T3 are found when the teacher's or external actor's role was combined with being Referent [$F(1,18) = 23.800$; $p < .001$; $\eta^2 = .266$].

Optimality score. Overall, the participants reached in average 63.9% of their optimal speed (range: 43.8%-100%). There was an improvement across the 3 scenarios [Time: $F(2,36) = 13.880$; $p < .001$; $\eta^2 = .315$], but there was no difference for the optimality score between users' category [Role: $F(2,18) = 0.740$; $p > .400$; $\eta^2 = .032$]. The interaction effect was not significant. Pairwise t-test revealed significant differences between T1 – T3 [$t(20) = -4.276$; $p_{adj} < .002$] and T2 – T3 [$t(20) = -3.662$; $p_{adj} < .006$]. The results depicted in Figure 6 showed that the optimality increased over time, in compliance with a large training effect, despite the participant's role. Like the 3 main roles (parent, teacher, external actor), there is no effect for the Referent role [$F(1,18) = 0.839$; $p > .300$; $\eta^2 = .024$]. Training effects between T1 and T3 are found when the teacher's or external actor's role was combined with being Referent [$F(1,18) = 23.800$; $p < .001$; $\eta^2 = .388$].

In summary, as the participants were progressing in the experiments, they were more efficient in the use of the interface. Their speed has increased from the 2nd scenario, and there were approaching their maximum efficiency.

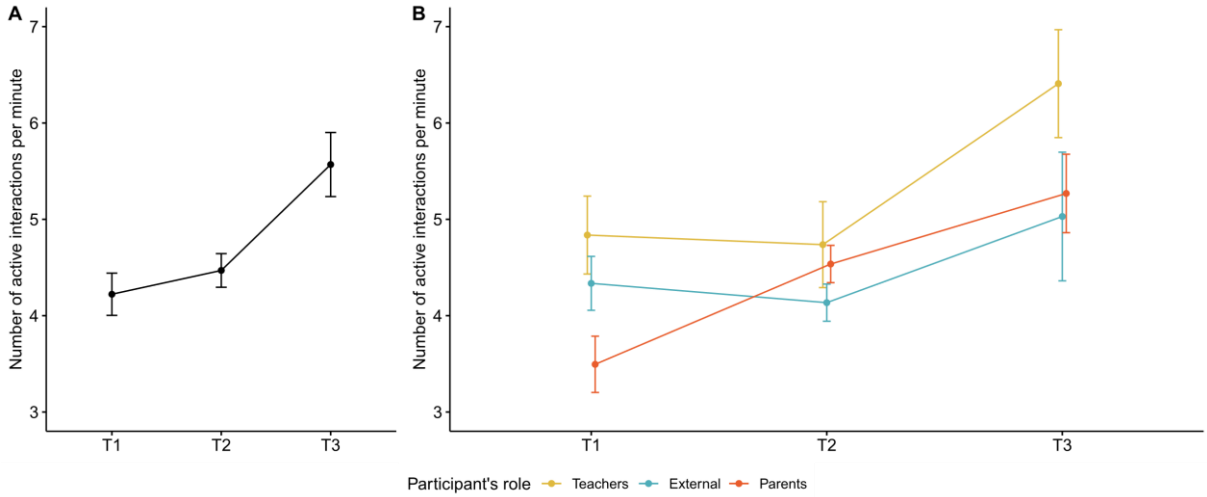

Figure 5: *Number of active interactions per minute across scenarios in average (A) and differentiated by role (B)*

A: Line chart for the whole group. In average, participants' number of active interactions per minute improved over time with a slow increase between T1 and T2, and a noticeable increase between T2 and T3.

B: Line chart differentiated by role. Participants' number of active interactions per minute increased over time for parents. Teachers' and external actors' number of active interactions per minutes slowly decreased between T1 and T2 and increased for T3.

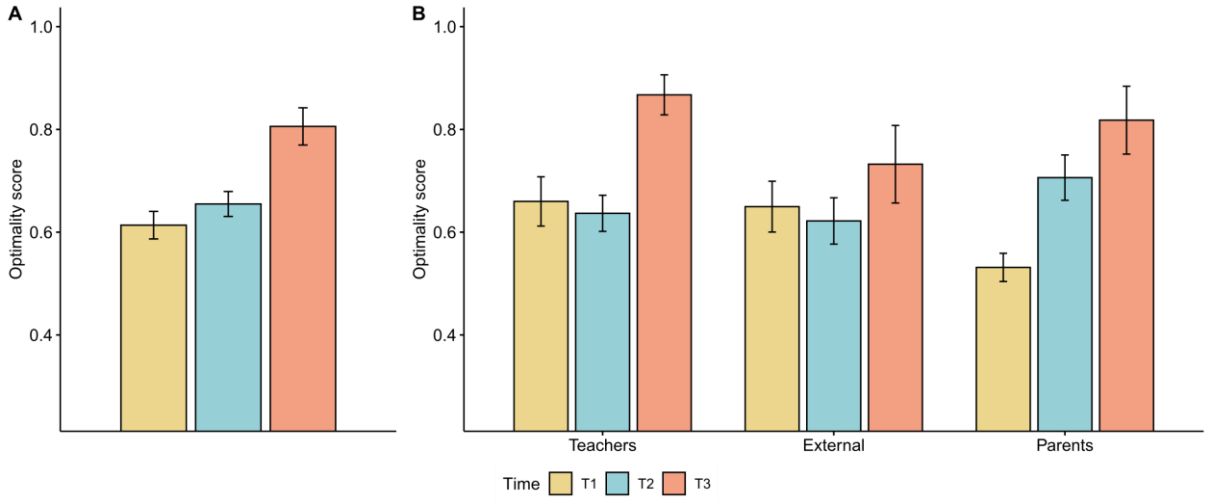

Figure 6: *Optimal score (participants'speed / theoretical optimal speed) across scenarios in average (A) and differentiated by role (B)*

A: Histogram for the whole group. In average, participants' optimal score improved over time with a slow increase between T1 and T2, and a noticeable increase between T2 and T3.

B: Histogram differentiated by role. Teachers' and externals' optimal score decreased between T1 and T2 and significantly improved at T3. Parents' optimal score improved over time.

3.2. Subjective Use measures

3.2.1. Scenario-based subjective usability

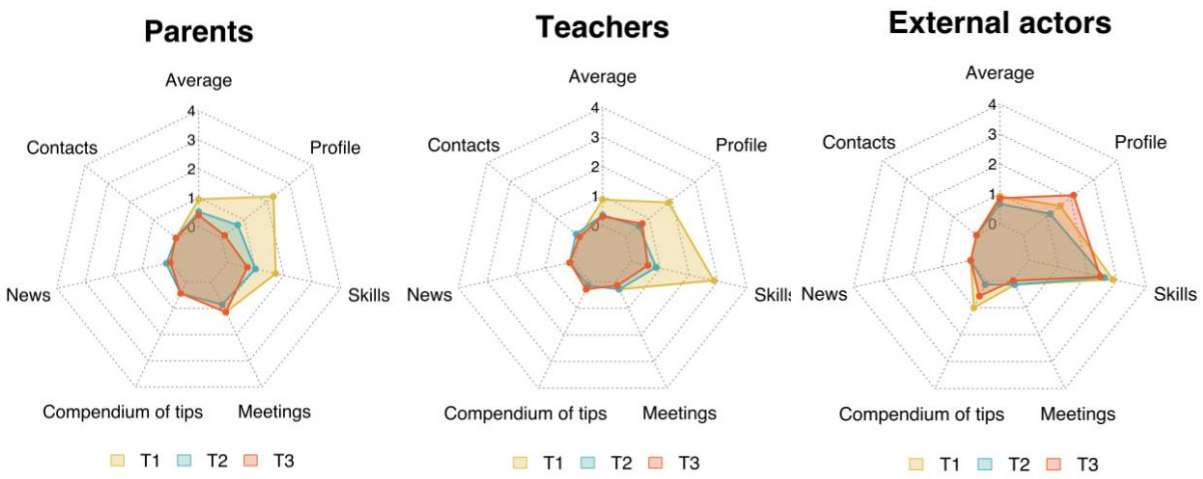

Figure 7: *Perceived difficulty answers' distribution depending on the role, the assessed tab and the test instance on a scale from 0 (very easy) to 10 (very difficult)*

Parents' and teachers' perceived difficulty was relatively higher for T1 than for T2 and T3. Level of perceived difficulty for external actors remained quite similar across time.

The average scenario-based subjective usability score for the whole group was 0.73/10 (0.83), meaning that user perceived interactions as easy. For the subsequent analysis, three independent variables have been considered for this score using a mixed three-way ANOVA: the Role as a between-group factor, and the Time and the Tab as within-group factors (See Appendix E for detailed results).

The global Role effect was not significant [$F(2,18) = 0.479$; $p > .600$; $\eta^2 = .011$], meaning that globally, the perceived interaction difficulty did not differ depending on the user role. The ANOVA yields a significant Tab effect [$F(5,90) = 11.537$; $p < .001$; $\eta^2 = .208$]: interactions in the "skills" tab and the "profile" tab were perceived slightly more difficult than the four other tabs ($p_{adj} < .01$). The Time effect [$F(2,36) = 3.697$; $p < .01$; $\eta^2 = .024$] was significant too, with a diminished score between the first and the second instances [$t(125) = 3.304$, $p_{adj.} > .01$], and

a tendency between the first and the third instance [$t(125) = 2.407$, $p_{adj} = .053$]. Finally, we found a significant Tab x Time interaction effect [$F(10,180) = 2.353$; $p < .01$; $\eta^2 = .035$]: the higher difficulty perceived in "profile" and "skills" tabs compared to others was only present during the first scenario ($p_{adj} < .05$). None of the other interaction effects were significant.

3.2.2. System Usability Scale

The average SUS score for the whole group was 92.02 (8.04), revealing an excellent usability of the web application. Between-group comparisons did not reveal significant differences [$F(2,18) = 0.447$; $p > .600$; $\eta^2 = .047$], meaning that the usability is equally assessed depending on the user role.

3.2.3. User Experience Questionnaire (UEQ)

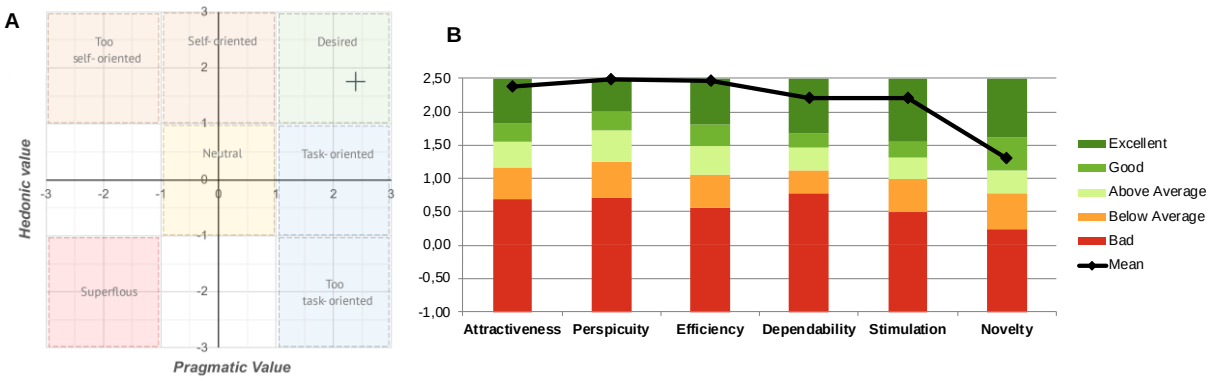

Figure 8: *Visual representation of UEQ results using Benchmark tool, in average (A) and for each scale (B)*

A: Mean UEQ results along hedonic and pragmatic dimension, it placed the web application *ToGather* in the "Desired" area

B: Each scale of the UEQ had "Excellent" scores except for the "Novelty" scale that was "Good"

The UEQ scores for the whole group was 2.385/3 (SD = 0.642) for pragmatic qualities, 1.750/3 (SD = 0.808) for hedonic qualities, and 2.373/3 (SD = 0.643) for the attractiveness. The ANOVA revealed that these differences are significant [$F(2,18) = 14.538$; $p < .001$; $\eta^2 = .167$], but there are no differences depending on the user's role [$F(2,18) = 0.761$; $p > .400$; $\eta^2 = .060$]. Indeed, the hedonic score is significantly lower than both the pragmatic [$t(20) = -4.122$;

$p_{adj} < .005$] and the attractiveness scores [$t(20) = 4.514$; $p_{adj} < .001$]. Nevertheless, hedonic and pragmatic scores placed the web application is in the desired area (Figure 8 A), even if the web application have more pragmatic qualities than hedonic ones.

The analysis of the 6 dimensions with the Benchmark tool is displayed on the Figure 8 B. We observed that Novelty is good and all the other dimensions are excellent. The ANOVA revealed only a significant effect of Dimension [$F(5,90) = 11.496$; $p < .001$; $\eta^2 = .219$]. Pairwise comparisons revealed that the Novelty dimension is significantly different from the other 5 dimensions [$.05 > p_{adj} > .001$]. From these overall data, we can conclude that the web application elicited excellent user experience.

3.2.4. NASA - Task Load Index

The average NASA-TLX raw score for the whole group was 12.98/100 (SD = 9.71), meaning that the web application elicited few use-related cognitive load (Figure 9). The ANOVA showed that the effect of user's role was not significant [$F(2,18) = 1.421$; $p > .200$; $\eta^2 = .066$].

However, the effect of NASA-TLX Dimensions was significant [$F(6,108) = 7.588$; $p < .001$; $\eta^2 = .189$]. The physical demand score is lower than the temporal demand score [$t(20) = -4.334$; $p_{adj} < .01$], the mental demand score [$t(20) = 3.728$; $p_{adj} < .05$], and the total raw score [$t(20) = 4.306$; $p_{adj} < .01$]. Also, the temporal demand score is significantly higher than the frustration score [$t(20) = 4.002$; $p_{adj} < .02$]. Indeed, the physical demand (M = 4.048) and the frustration (M = 6.905) were the lowest scores observed, while the mental demand (M = 21.667) and the temporal demand (M = 21.905) were the highest scores (albeit fairly low compared to the max score).

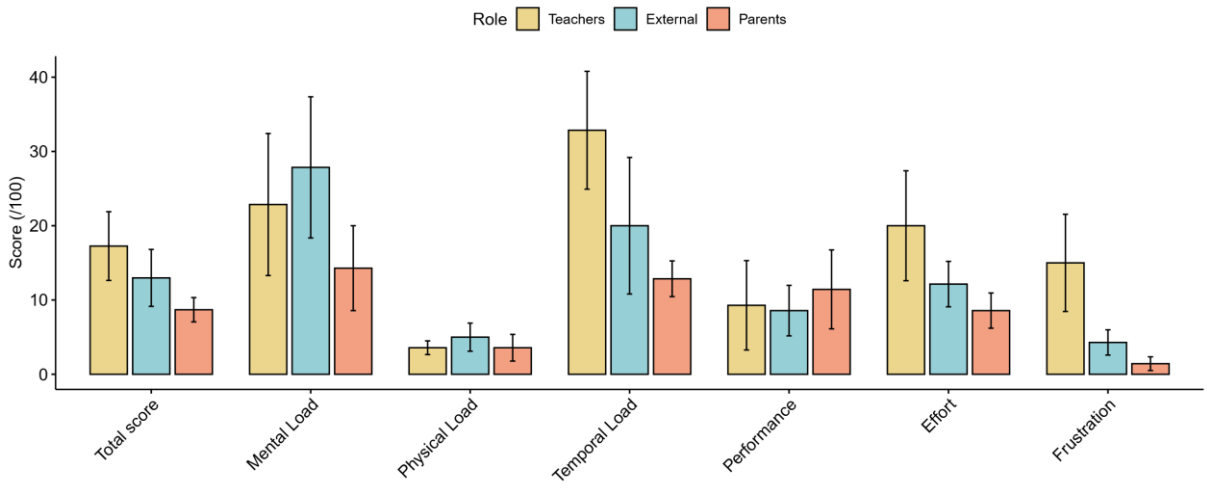

Figure 9: *NASA-TLX raw score and subscales*

Total raw score indicates a higher cognitive load for teachers, a lower one for external actors and a lower one for parents. Temporal and Mental load show the highest (but still low as it is below 40/100) scores for each role. Physical load shows the lowest score for each role.

3.2.5. Elicited self-determination

The average self-determination score for the whole group was 3.84 (SD = 0.26), meaning that the web application elicited a high self-determination score. Such high score means that the interface elicited a great level of self-determination feeling for collaborative activities. The ANOVA did not reveal significant differences depending on the user's role [$F(2,18) = 1.147$; $p > .300$; $\eta^2 = .094$], nor between autonomy, competence and relatedness [$F(3,54) = 1.859$; $p > .100$; $\eta^2 = .019$].

4. Discussion

The objective of the present work aimed a user assessment study to ultimately achieve the participatory design of the *ToGather* web application., designed to promote collaboration and improve the follow-up of the school inclusion of children with ASD, by all IEP stakeholders (i.e., family, school, and external actors).

Overall, the results show high effectiveness and efficiency of use, according to our objective measures, across scenarios and user roles. Additionally, the user testing reveals that the final release elicited an excellent user experience, as revealed by the scores on the SUS and the UEQ. The latter scale provides more detailed information on the user experience, with the distinction between several dimensions. As we expected, the web application elicited more task-oriented, pragmatic values, than hedonic ones, meaning that the web application has been appreciated for its usefulness qualities. Moreover, 20 out of 21 participants (95%) said that they would recommend it to their friends and/or colleagues, and use it every day, or at least once a week. The NASA-TLX score was relatively low, meaning that the use of the website requires low cognitive load, but this result is contrasted by the fact that mental and temporal exigence are the items with the higher scores. This result is confirmed by informal discussions at the end of the experiments, which revealed concerns from the participant regarding data entry time. However, when a technology meets the needs of its users by proposing functionalities allowing them to perform a task – collaborate in this case -, it makes the effort acceptable (Goodhue & Thompson, 1995; Spies, Grobbelaar, Botha, Hattingh, Matthee, Smuts, *et al.*, 2020).

The great level of elicited self-determination by the app underlines the need for a support tool to cooperate by gathering information regarding the pupil and any ongoing intervention in order to coordinate actions lead towards them. Such a tool also facilitates inclusion of every stakeholder in the follow-up team by empowering them as they are all invited to contribute to the pupil's file and thus giving them a voice to participate equally in every collective decision for the child.

From the above results, user-testing for the new release of *ToGather* appears conclusive.

Despite of this positive results, we identified a particular difficulty with the "Profile" and the "Skills" tabs, for which the perceived difficulty was slightly higher than for other tabs. Also, 6 out of 21 participants expressed this difficulty in the questionnaire, but this feeling was

often expressed at the first sight of the web application, and attenuated over the course of scenario-based testing, as highlighted by significant Time effect across objective measures revealing an efficient learning of *ToGather* handling. Hence, we can assume that the *ToGather* is suitable and ergonomic, as difficulties disappeared entirely or at least considerably reduced for all participants during the three scenarios. This was supported by participants' ease of use whereas most of them do not feel very comfortable while using their computer (we noticed they were lost to use Zoom, fill a PDF document for the consent form, they do not have the reflex to use sort or filter functionalities while encountering multiple information on a single page, etc.). According to the Rogers' theory of diffusion (1995) which offers a comprehensive view regarding the processes involved in accepting or discontinuing use of technology, the prior trialability is a key dimension for continued use of technology, particularly for individuals with cognitive impairments (i.e., Parette & VanBiervliet, 1992; Riemer-Reiss & Wacker, 2000). Trialability is the degree to which the user has prior use experiences with the technology for its long-term adopting. From this key concept, it could be assumed that observed first positive interaction experience with the features of *ToGather* website are decisive elements for expecting its long-term adoption amongst the different targeted end-users.

Additionally, we observed minor different needs expressed by the three parties (family, school and medico-social domains). Some people expressed the desire to make certain content visible only to certain user profiles, but we have decided not to exclude anyone by making the information available to all stakeholders, as this tool is intended to foster collaboration between all those involved in the student's educational inclusion. Taken together, these observations supported PD methods as fruitful for offering functionalities consensually desired by the end-users. The *ToGather* app elicited a real enthusiasm of the participants. As highlighted in a previous work (*Mazon, et al., 2021*), this enthusiasm could be partly explained by the empowerment linked to the feeling of having their needs heard and the opportunity to contribute

to change things. This means that the PD within UCD methods has fulfilled their role, as they allowed to design an innovative tool, derived from the users' expressed needs.

Moreover, the PD process appears insightful for the field of technology-based interventions for children with ASD. By its highly field user need-driven nature, the PD method has enabled a step change in this field of research from an individual approach to ASD to a whole social approach where social environments are stressed as essential ingredients of an optimal school inclusion (i.e., Frauenberger, Good & Keay-Bright, 2011; Robb, Boyle, Politis, Newbutt, Kuo, Sung, *et al.*, 2021). As already mentioned, the current frameworks on child's functioning such as ICF-CY or ecosystemic models of school inclusion of children with ASD, fostering school inclusion of pupils with ASD promotes a vision in which socioenvironmental factors are identified as key levers for a better social participation of children with ASD (Bateman, *et al.*, 2022; Bowman, Suarez & Weiss, 2021). By extension, this amounts to embedding in the technologies not an individual agentivity as already done, but a collective agentivity at the service of co-education.

Conducting user tests before developing the final version of a website or an app is essential to ensure a design that really users' needs, as it allows them to be confronted with a physical tool – here interface – and not the idea of it. As a mock-up can appear as interactive as a website, participants were confronted to it as if they were on an actual web application. Doing so helps revealing design mistakes or functionality oversights. For example, we reviewed the dashboard by using a need approach depending on the type of authenticated user (i.e.., teacher or parent) based on the qualitative data collected during video annotation. Taken together, the present results provided useful and extensive insights from the three kinds of end-users. Thanks to our new user-testing method mixing personas and usage scenarios methods on an interactive mock-up, it was possible to collect objective and subjective measures while respecting GDPR. A such method could be fruitful for future user assessment studies.

4.1. Limitations and future works

Several limitations need to be addressed and they provide avenues for future work.

- *Limit related to non-homogeneous interactive features across tabs:* depending on the tab of the website, more active interactions are solicited from the user (i.e., "skills" tab vs. "compendium of tips" tab) whose characteristic is to make the task more difficult and then error-prone. One way to address this limitation would be to avoid any free entry by the user, and propose a drop-down menu of possible entries, in accordance with Nielsen and Landauer/Bastien and Scapin's criteria for error prevention in web navigation (1993). A specific study comparing a free-entry versus an entry selection system would allow to evaluate if it would enable to overcome this difficulty without frustrating the user who is stuck in a limited field of possible entries.
- *Limit related to the collection of interaction data:* As a millisecond accuracy measurement does not provide an increased insight for the study purpose, we opted to track interactions with some interaction logs and video analysis of navigation. Despite of this, our method could be improved by extending the tracking of interactions in the HTML code to gain more accuracy (Atterer, Wnuk & Schmidt, 2006). It might be interesting to have a comparative study between our simple method and a method of exhaustive and accurate tracking of all interactions in order to evaluate whether our method is a good proxy for what would be observed with a more exhaustive and accurate method of collecting the interaction logs.
- *Limit of use scenario method:* Like any ergonomic study based on usage scenarios, an essential limit is that this method allows us to get closer to real life but for all that the conditions of real life are not met. It is therefore a simulation of use with its strengths and weaknesses. It is interesting to note that the 3 scenarios elaborated were judged by our participants as very heterogeneous (from easy to difficult), which was desired to be representative of the real conditions. In any case, this last limitation highlights the importance of a further field study to

evaluate the long-term adoption and the usefulness of *ToGather* in order to have a full real-life assessment.

5. Conclusions

The design of *"ToGather"* app was based on a user-centered methodology involving participatory process, beginning with an analysis of the information needs of family and professional caregivers of student with ASD included in mainstream school and finishing with a conclusive user-testing in terms of both objective measures (usability effectiveness and efficiency) and subjective measures regarding usability, cognitive load, self-determination as well as user experience. Future work will pursue the assessment of the *ToGather* app, with a field study including a large panel of users in order to cover the evaluation of its efficacy in supporting parents, teachers, and professional caregivers, as well as the interactions between the three settings (i.e., family, school, medico-social). This future study may highlight whether the *ToGather* app may or not improve the quality of support provided to students with ASD included in mainstream schools.

Acknowledgements

First of all, I would like to thank Caroline Simonpietri, Miliana Rahouadj and Kamélia Belassel, temporary study and research associate, design research intern and cognitive science research intern, for their help during experiments and video annotations.

Our gratitude is also expressed to Bénédicte Mendiboure, Austim & Inclusive Projects Director at the ARI association; the Academy of Bordeaux and especially Maïc Sagel from the ASH mission as well as Sébatien Gouleau from the digital mission, for their help in designing the web-app, recruiting participants and disseminating information on the project.

Thanks to parents, teachers and health and well-being professionals for their time and feedbacks on the web app.

We thank also the support provided by our financial partners, the FIRAH, the Klésia, the CCAH and the Nouvelle-Aquitaine region.

Fundings

This work was supported by the FIRAH[2] and their partners Klésia and CCAH[3], Paris, France; the Nouvelle-Aquitaine Region, France [convention n° AAPR2021-2020-11922210]; the French national PhD program into the EDSP2 - doctoral school of University of Bordeaux, Talence, France.

References

Abolkheirian, S., Sadeghi, R., & Shojaeizadeh, D. (2022). What do parents of children with autism spectrum disorder think about their quality of life? A qualitative study. *Journal of Education and Health Promotion*, *11*, 169. https://doi.org/10.4103/jehp.jehp_887_21

Alexandri, M., Papailiou, C., & Nikolaou, E. (2017). Teachers' Social Representations of Inclusion of Children with Autism Spectrum Disorder (ASD) in Regular Class. *Open Journal for Educational Research*, *1*(2), 45–56.

Allison, R., Hayes, C., McNulty, C. A. M., & Young, V. (2019). A Comprehensive Framework to Evaluate Websites: Literature Review and Development of GoodWeb. *JMIR Formative Research*, *3*(4), e14372. https://doi.org/10.2196/14372

American Psychological Association. (2013). *Autism spectrum disorder*. Https://Www.Apa.Org. https://www.apa.org/topics/autism-spectrum-disorder

[2] International Foundation for Applied Research on Disability
[3] national Coordinating Committee on Disability Action

Anthony, N., & Campbell, E. (2020). Promoting Collaboration Among Special Educators, Social Workers, and Families Impacted by Autism Spectrum Disorders. *Advances in Neurodevelopmental Disorders*, *4*(3), 319–324. https://doi.org/10.1007/s41252-020-00171-w

Antle, A. N. (2006). Child-personas: Fact or fiction? *Proceedings of the 6th ACM Conference on Designing Interactive Systems - DIS '06*, 22. https://doi.org/10.1145/1142405.1142411

Atterer, R., Wnuk, M., & Schmidt, A. (2006). Knowing the user's every move: User activity tracking for website usability evaluation and implicit interaction. *Proceedings of the 15th International Conference on World Wide Web*, 203–212. https://doi.org/10.1145/1135777.1135811

Azad, G. F., Marcus, S. C., Sheridan, S. M., & Mandell, D. S. (2018). Partners in School: An Innovative Parent-Teacher Consultation Model for Children with Autism Spectrum Disorder. *Journal of Educational and Psychological Consultation: The Official Journal of the Association for Educational and Psychological Consultants*, *28*(4), 460–486. https://doi.org/10.1080/10474412.2018.1431550

Baker, C., Galemore, C. A., & Lowrey, K. M. (2020). Information Sharing in the School Setting During a Public Health Emergency. *Nasn School Nurse (Print)*, *35*(4), 198–202. https://doi.org/10.1177/1942602X20925031

Bastien, C., & Scapin, D. (1993). *Ergonomic criteria for the evaluation of human-computer interfaces* (p. 79) [Report]. INRIA. https://hal.inria.fr/inria-00070012

Bateman, K. J., Schwartz, I. S., & Gauvreau, A. N. (2022). It Takes a Team: Working Together to Meet the Needs of Young Children With Autism Spectrum Disorder in an Inclusive Setting. *Inclusive Practices*, 27324745221097350. https://doi.org/10.1177/27324745221097354

Behnia, F., Rassafiani, M., Nakhai, S., Mohammadpour, M., & Ahmadi Kahjoogh, M. (2017). Time Use of Mothers of Children With an Autism Spectrum Disorder: A Comparative Study. *Iranian Rehabilitation Journal*, *15*(1), 49–56. https://doi.org/10.18869/nrip.irj.15.1.49

Bowman, K. S., Suarez, V. D., & Weiss, M. J. (2021). Standards for Interprofessional Collaboration in the Treatment of Individuals With Autism. *Behavior Analysis in Practice*, *14*(4), 1191–1208. https://doi.org/10.1007/s40617-021-00560-0

Brajić, A., Bećirović, S., Marić, B., Avdibegovic, M., Delic, S., & Bećirović, D. (2012). *ENSURING SOCIAL COHESION-CITIZENS AS PLANNERS OF URBAN GREEN SPACES*. *27–30*. https://www.researchgate.net/publication/331872088_ENSURING_SOCIAL_COHESION-CITIZENS_AS_PLANNERS_OF_URBAN_GREEN_SPACES

Brooke, john. (1996). SUS: A "Quick and Dirty" Usability Scale. In *Usability Evaluation In Industry*. CRC Press.

Budi Santoso, T. (2021). Daily Occupational Balance of Mothers of a Child with Autism Spectrum Disorders. *Annals of Physiotherapy & Occupational Therapy*, *4*(2). https://doi.org/10.23880/APhOT-16000194

Cappe, É., & Boujut, E. (2016). L'approche écosystémique pour une meilleure compréhension des défis de l'inclusion scolaire des élèves ayant un Trouble du Spectre de l'Autisme. *ANAE - Approche Neuropsychologique Des Apprentissages Chez l'Enfant*, *143*, 391–401.

Casartelli, L., Federici, A., Fumagalli, L., Cesareo, A., Nicoli, M., Ronconi, L., Vitale, A., Molteni, M., Rizzolatti, G., & Sinigaglia, C. (2020). Neurotypical individuals fail to understand action vitality form in children with autism spectrum disorder. *Proceedings*

of the National Academy of Sciences, *117*(44), 27712–27718. https://doi.org/10.1073/pnas.2011311117

Cegarra, J., & Morgado, N. (2009). Étude des propriétés de la version francophone du NASA-TLX. *EPIQUE 2009: 5ème Colloque de Psychologie Ergonomique*, 233–239.

Chevallier, E., Courtinat-Camps, A., & de Léonardis, M. (2015). Estime de soi chez des élèves scolarisés en classe d'inclusion scolaire (CLIS). *Neuropsychiatrie de l'Enfance et de l'Adolescence*, *63*(2), 76–83. https://doi.org/10.1016/j.neurenf.2014.12.002

Choi, B. C. K., & Pak, A. W. P. (2006). Multidisciplinarity, interdisciplinarity and transdisciplinarity in health research, services, education and policy: 1. Definitions, objectives, and evidence of effectiveness. *Clinical and Investigative Medicine. Medecine Clinique Et Experimentale*, *29*(6), 351–364.

Coster, W., Law, M., Bedell, G., Liljenquist, K., Kao, Y.-C., Khetani, M., & Teplicky, R. (2013). School participation, supports and barriers of students with and without disabilities. *Child: Care, Health and Development*, *39*(4), 535–543. https://doi.org/10.1111/cch.12046

Crowell, J. A., Keluskar, J., & Gorecki, A. (2019). Parenting behavior and the development of children with autism spectrum disorder. *Comprehensive Psychiatry*, *90*, 21–29. https://doi.org/10.1016/j.comppsych.2018.11.007

De, M., Chakraborty, R., Bhattacharya, M., Bhunia, P., & Basu, I. (2018). *Challenges for teachers in inclusive education vis-à-vis the barriers faced by students with Autism Spectrum Disorder (ASD)*. *3*, 8.

Desombre, C., Lamotte, M., & Jury, M. (2019). French teachers' general attitude toward inclusion: The indirect effect of teacher efficacy. *Educational Psychology*, *39*(1), 38–50. https://doi.org/10.1080/01443410.2018.1472219

Dubois, J., & Jumel, S. (2019). *Rapport fait au nom de la commission d'enquête sur l'inclusion des élèves handicapés dans l'école et l'université de la République, quatorze ans après la loi du 11 février 2005* (No. 2178). French National Assembly. https://www.assemblee-nationale.fr/dyn/15/rapports/ceincleh/l15b2178_rapport-enquete

Fage, C., Consel, C., Etchegoyen, K., Amestoy, A., Bouvard, M., & Sauzéon, H. (2016). Applications Numériques pour la Cognition Sociale pour Favoriser l'Inclusion Scolaire des Élèves avec Troubles du Spectre Autistique (TSA). In V. Brun, J.-M. Mazaux, P.-A. Joseph, & A. Prouteau (Eds.), *La Cognition Sociale* (pp. 85–109). SAURAMPS MEDICAL. https://hal.inria.fr/hal-01611303

Fage, C., Moullet, P., Consel, C. Y., & Sauzéon, H. (2017). School inclusion of children with special needs in France. In Wehmeyer & Patton (Eds.), *The Praeger International Handbook of Special Education* (Vol. 2, pp. 263–280). ABC-CLIO.

Ferraioli, S. J., & Harris, S. L. (2011). Effective Educational Inclusion of Students on the Autism Spectrum. *Journal of Contemporary Psychotherapy*, *41*(1), 19–28. https://doi.org/10.1007/s10879-010-9156-y

Fombonne, É., Myers, J., Chavez, A., Presmanes Hill, A., & Zuckerman, K. (2019). Épidémiologie de l'autisme: Où en sommes-nous ? (R. Pry, Trans.). *Enfance*, *1*(1), 13–47. https://doi.org/10.3917/enf2.191.0013

Frauenberger, C., Good, J., & Keay-Bright, W. (2011). Designing technology for children with special needs: Bridging perspectives through participatory design. *CoDesign*, *7*(1), 1–28. https://doi.org/10.1080/15710882.2011.587013

French Ministry of National Education, Youth and Sports, M. de l'Éducation, de la Jeunesse et des Sports. (2015). *Scolarisation des élèves en situation de handicap*. Ministère de

l'Education Nationale et de la Jeunesse. https://www.education.gouv.fr/bo/15/Hebdo31/MENE1504950C.htm

French Ministry of National Education, Youth and Sports, M. de l'Éducation, de la Jeunesse et des Sports. (2019). *Repères et références statistiques sur les enseignements, la formation et la recherche 2019*. Ministère de l'Education Nationale et de la Jeunesse. https://www.education.gouv.fr/reperes-et-references-statistiques-sur-les-enseignements-la-formation-et-la-recherche-2019-3806

French Ministry of National Education, Youth and Sports, M. de l'Éducation, de la Jeunesse et des Sports. (2020). *Mettre en œuvre un plan d'accompagnement personnalisé*. Ministère de l'Éducation nationale et de la Jeunesse - Direction générale de l'enseignement scolaire. https://eduscol.education.fr/1214/mettre-en-oeuvre-un-plan-d-accompagnement-personnalise

Garcia-Melgar, A., Hyett, N., Bagley, K., McKinstry, C., Spong, J., & Iacono, T. (2022). Collaborative team approaches to supporting inclusion of children with disability in mainstream schools: A co-design study. *Research in Developmental Disabilities, 126*, 104233. https://doi.org/10.1016/j.ridd.2022.104233

Gibbons, S. (2015). Disability, Neurological Diversity, and Inclusive Play: An Examination of the Social and Political Aspects of the Relationship between Disability and Games. *Loading...*, *9*(14), Article 14. https://journals.sfu.ca/loading/index.php/loading/article/view/150

Gisev, N., Bell, J. S., & Chen, T. F. (2013). Interrater agreement and interrater reliability: Key concepts, approaches, and applications. *Research in Social and Administrative Pharmacy, 9*(3), 330–338. https://doi.org/10.1016/j.sapharm.2012.04.004

Gomes, M. F., & McVilly, K. R. (2019). The Characteristics of Effective Staff Teams in Disability Services. *Journal of Policy and Practice in Intellectual Disabilities*, *16*(3), 191–200. https://doi.org/10.1111/jppi.12280

Gómez-Marí, I., Sanz-Cervera, P., & Tárraga-Mínguez, R. (2021). Teachers' Knowledge Regarding Autism Spectrum Disorder (ASD): A Systematic Review. *Sustainability*, *13*(9), Article 9. https://doi.org/10.3390/su13095097

Goodenow, C. (1993). The psychological sense of school membership among adolescents: Scale development and educational correlates. *Psychology in the Schools*, *30*(1), 79–90. https://doi.org/10.1002/1520-6807(199301)30:1<79::AID-PITS2310300113>3.0.CO;2-X

Goodhue, D. L., & Thompson, R. L. (1995). Task-Technology Fit and Individual Performance. *MIS Quarterly*, *19*(2), 213–236. https://doi.org/10.2307/249689

Gronier, G., & Baudet, A. (2021). Psychometric Evaluation of the F-SUS: Creation and Validation of the French Version of the System Usability Scale. *International Journal of Human–Computer Interaction*, *37*(16), 1571–1582. https://doi.org/10.1080/10447318.2021.1898828

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Eds.), *Advances in Psychology* (Vol. 52, pp. 139–183). North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9

Hassenzahl, M., & Tractinsky, N. (2006). User experience—A research agenda. *Behaviour & Information Technology*, *25*(2), 91–97. https://doi.org/10.1080/01449290500330331

Hodges, A., Joosten, A., Bourke-Taylor, H., & Cordier, R. (2020). School participation: The shared perspectives of parents and educators of primary school students on the autism

spectrum. *Research in Developmental Disabilities*, *97*, 103550. https://doi.org/10.1016/j.ridd.2019.103550

Humphrey, N., & Symes, W. (2011). Peer interaction patterns among adolescents with autistic spectrum disorders (ASDs) in mainstream school settings. *Autism*, *15*(4), 397–419. https://doi.org/10.1177/1362361310387804

Kinnear, S. H., Link, B. G., Ballan, M. S., & Fischbach, R. L. (2016). Understanding the Experience of Stigma for Parents of Children with Autism Spectrum Disorder and the Role Stigma Plays in Families' Lives. *Journal of Autism and Developmental Disorders*, *46*(3), 942–953. https://doi.org/10.1007/s10803-015-2637-9

Kirsh, I., & Joy, M. (2020). Exploring Pointer Assisted Reading (PAR): Using Mouse Movements to Analyze Web Users' Reading Behaviors and Patterns. In C. Stephanidis, M. Kurosu, H. Degen, & L. Reinerman-Jones (Eds.), *HCI International 2020—Late Breaking Papers: Multimodality and Intelligence* (pp. 156–173). Springer International Publishing. https://doi.org/10.1007/978-3-030-60117-1_12

Laugwitz, B., Held, T., & Schrepp, M. (2008). Construction and Evaluation of a User Experience Questionnaire. In A. Holzinger (Ed.), *HCI and Usability for Education and Work* (pp. 63–76). Springer. https://doi.org/10.1007/978-3-540-89350-9_6

Lefebvre, H., Pelchat, D., Swaine, B., Gélinas, I., & Levert, M. J. (2005). The experiences of individuals with a traumatic brain injury, families, physicians and health professionals regarding care provided throughout the continuum. *Brain Injury*, *19*(8), 585–597. https://doi.org/10.1080/02699050400025026

Li, F., Xu, M., Wu, D., Tang, Y., Zhang, L., Liu, X., Zhou, L., Li, F., & Jiang, L. (2022). From child social impairment to parenting stress in mothers of children with ASD: The role of parental self-efficacy and social support. *Frontiers in Psychiatry*, *13*. https://www.frontiersin.org/articles/10.3389/fpsyt.2022.1005748

Limayem, M., Hirt, S. G., & Cheung, C. M. K. (2007). How Habit Limits the Predictive Power of Intention: The Case of Information Systems Continuance. *MIS Quarterly*, *31*(4), 705–737. https://doi.org/10.2307/25148817

Mazon, C., Etchegoyhen, K., Saint-Supery, I., Amestoy, A., Bouvard, M., Consel, C., & Sauzéon, H. (2021). Fostering parents-professional collaboration for facilitating the school inclusion of students with ASD: Design of the "ToGather" web-based prototype. *Educational Technology Research and Development*. https://doi.org/10.1007/s11423-021-10073-w

Mazon, C., Fage, C., & Sauzéon, H. (2019). Effectiveness and usability of technology-based interventions for children and adolescents with ASD: A systematic review of reliability, consistency, generalization and durability related to the effects of intervention. *Computers in Human Behavior*, *93*, 235–251. https://doi.org/10.1016/j.chb.2018.12.001

Melo, Á. H. da S., Rivero, L., Santos, J. S. dos, & Barreto, R. da S. (2020). PersonAut: A personas model for people with autism spectrum disorder. *Proceedings of the 19th Brazilian Symposium on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3424953.3426651

Miaskiewicz, T., & Kozar, K. A. (2011). Personas and user-centered design: How can personas benefit product design processes? *Design Studies*, *32*(5), 417–430. https://doi.org/10.1016/j.destud.2011.03.003

Minke, K. M., Sheridan, S. M., Kim, E. M., Ryoo, J. H., & Koziol, N. A. (2014). Congruence in parent-teacher relationships: The role of shared perceptions. *The Elementary School Journal*, *114*, 527–546. https://doi.org/10.1086/675637

Mintz, J., Seleznyov, S., Peacey, N., Brown, C., & White, S. (2021). Evidence informed practice for autism, special educational needs and disability in schools: Expanding the

scope of the research learning community model of professional development. *Support for Learning*, *36*(2), 159–182. https://doi.org/10.1111/1467-9604.12349

Mitter, N., Ali, A., & Scior, K. (2019). Stigma experienced by families of individuals with intellectual disabilities and autism: A systematic review. *Research in Developmental Disabilities*, *89*, 10–21. https://doi.org/10.1016/j.ridd.2019.03.001

Nielsen, J., & Landauer, T. K. (1993). A mathematical model of the finding of usability problems. *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*, 206–213. https://doi.org/10.1145/169059.169166

Noreen, H., Intizar, F., & Gulzar, S. (2019). Teachers' Multidimensional Attitude Towards Inclusive Education. *UMT Education Review*, *02*, 72–89. https://doi.org/10.32350/uer.22.04

Odom, S. L., Thompson, J. L., Hedges, S., Boyd, B. A., Dykstra, J. R., Duda, M. A., Szidon, K. L., Smith, L. E., & Bord, A. (2015). Technology-Aided Interventions and Instruction for Adolescents with Autism Spectrum Disorder. *Journal of Autism and Developmental Disorders*, *45*(12), 3805–3819. https://doi.org/10.1007/s10803-014-2320-6

Olmstead, C. (2013). Using Technology to Increase Parent Involvement in Schools. *TechTrends*, *57*(6), 28–37. https://doi.org/10.1007/s11528-013-0699-0

Osborne, L. A., & Reed, P. (2011). School factors associated with mainstream progress in secondary education for included pupils with Autism Spectrum Disorders. *Research in Autism Spectrum Disorders*, *5*(3), 1253–1263. https://doi.org/10.1016/j.rasd.2011.01.016

Paraskevi, G. (2021). The Inclusion of Children with Autism in the Mainstream School Classroom. Knowledge and Perceptions of Teachers and Special Education Teachers. *Open Access Library Journal*, *8*(9), Article 9. https://doi.org/10.4236/oalib.1107855

Parette, H. P., & VanBiervliet, A. (1992). Tentative findings of a study of the technology needs and use patterns of persons with mental retardation. *Journal of Intellectual Disability Research: JIDR*, *36 (Pt 1)*, 7–27. https://doi.org/10.1111/j.1365-2788.1992.tb00468.x

Pontikas, C.-M., Tsoukalas, E., & Serdari, A. (2022). A map of assistive technology educative instruments in neurodevelopmental disorders. *Disability and Rehabilitation-Assistive Technology*, *17*(7), 738–746. https://doi.org/10.1080/17483107.2020.1839580

Pope, J. (2009). *The significance of navigation and interactivity design for readers' responses to interactive narrative: Some conclusions from an empirical study of readers' responses*. https://doi.org/10.25969/MEDIAREP/17729

Rahimi, N., & Ibarra, M. (2014). A review of multiple user center design methods for new product development in smart and connected health applications. *Proceedings of PICMET'14 Conference: Portland International Center for Management of Engineering and Technology; Infrastructure and Service Integration*, 3498–3510.

Riemer-Reiss, M. L., & Wacker, R. R. (2000). *Factors associated with assistive technology discontinuance among individuals with disabilities*. *66*(3), 44–50.

Rispoli, K. M., Lee, G. K., Nathanson, E. W., & Malcolm, A. L. (2019). The parent role in school-based teams for adolescents with autism spectrum disorder. *School Psychology, 34*, 458–467. https://doi.org/10.1037/spq0000289

Robb, N., Boyle, B., Politis, Y., Newbutt, N., Kuo, H. J., & Sung, C. (2021). Participatory Technology Design for Autism and Cognitive Disabilities: A Narrative Overview of Issues and Techniques. In A. L. Brooks, S. Brahman, B. Kapralos, A. Nakajima, J. Tyerman, & L. C. Jain (Eds.), *Recent Advances in Technologies for Inclusive Well-Being: Virtual Patients, Gamification and Simulation* (pp. 469–485). Springer International Publishing. https://doi.org/10.1007/978-3-030-59608-8_25

Roberts, J., & Simpson, K. (2016). A review of research into stakeholder perspectives on inclusion of students with autism in mainstream schools. *International Journal of Inclusive Education*, *20*(10), 1084–1096. https://doi.org/10.1080/13603116.2016.1145267

Roberts-Yates, C., & Silvera-Tawil, D. (2019). Better Education Opportunities for Students with Autism and Intellectual Disabilities through Digital Technology. *International Journal of Special Education*, *34*(1), 197–210.

Rogers, E. M. (1995). *Diffusion of innovations* (4th ed). Free Press.

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective: Definitions, theory, practices, and future directions. *Contemporary Educational Psychology*, *61*, 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Saggers, B., Klug, D., Harper-Hill, K., Ashburner, J., Costley, D., Clark, T., Bruck, S., Trembath, D., A. Webster, A., & Carrington, S. (2016). *Australian Autism Educational Needs Analysis – What are the needs of schools, parents and students on the autism spectrum?* (Executive Summary No. 978-0-9941641-2–4; p. 16). Cooperative Research Centre for Living with Autism (Autism CRC). https://www.autismcrc.com.au/sites/default/files/inline-files/Educational%20needs%20analysis%20-%20Executive%20summary.pdf

Salceanu, C. (2020). Development and Inclusion of Autistic Children in Public Schools. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, *11*(1), Article 1. https://doi.org/10.18662/brain/11.1/12

Sarge, M. A., Kim, H.-S., & Velez, J. A. (2020). An Auti-Sim Intervention: The Role of Perspective Taking in Combating Public Stigma with Virtual Simulations.

Cyberpsychology, Behavior, and Social Networking, *23*(1), 41–51. https://doi.org/10.1089/cyber.2019.0678

Sears, A., & Jacko, J. A. (Eds.). (2009). *Human-Computer Interaction: Development Process*. CRC Press. https://doi.org/10.1201/9781420088892

Shahidullah, J. D., McClain, M. B., Azad, G., Mezher, K. R., & McIntyre, L. L. (2020). Coordinating Autism Care Across Schools and Medical Settings: Considerations for School Psychologists. *Intervention in School and Clinic*, *56*(2), 107–114. https://doi.org/10.1177/1053451220914891

Spies, R., Grobbelaar, S., & Botha, A. (2020). A Scoping Review of the Application of the Task-Technology Fit Theory. In M. Hattingh, M. Matthee, H. Smuts, I. Pappas, Y. K. Dwivedi, & M. Mäntymäki (Eds.), *Responsible Design, Implementation and Use of Information and Communication Technology* (pp. 397–408). Springer International Publishing. https://doi.org/10.1007/978-3-030-44999-5_33

Stephenson, J., Browne, L., Carter, M., Clark, T., Costley, D., Martin, J., Williams, K., Bruck, S., Davies, L., & Sweller, N. (2021). Facilitators and Barriers to Inclusion of Students With Autism Spectrum Disorder: Parent, Teacher, and Principal Perspectives. *Australasian Journal of Special and Inclusive Education*, *45*(1), 1–17. https://doi.org/10.1017/jsi.2020.12

Stepney, M. (2008). The Evolution of Autism in Education. In *Making Waves in Education* (pp. 169–174). ESCalate. http://escalate.ac.uk/downloads/5811.pdf#page=169

Suhaila, N. A., & Nordin, N. M. (2022). Assistive Technology for Autism Spectrum Disorder: Systematic Literature Review. *International Journal of Advanced Research in Education and Society*, *4*(2), Article 2.

ten Hoopen, L. W., de Nijs, P. F. A., Duvekot, J., Greaves-Lord, K., Hillegers, M. H. J., Brouwer, W. B. F., & Hakkaart-van Roijen, L. (2022). Caring for Children with an

Autism Spectrum Disorder: Factors Associating with Health- and Care-Related Quality of Life of the Caregivers. *Journal of Autism and Developmental Disorders*, *52*(11), 4665–4678. https://doi.org/10.1007/s10803-021-05336-7

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1992). The Academic Motivation Scale: A Measure of Intrinsic, Extrinsic, and Amotivation in Education. *Educational and Psychological Measurement*, *52*(4), 1003–1017. https://doi.org/10.1177/0013164492052004025

van der Velden, M., & Mörtberg, C. (2021). Participatory Design and Design for Values. In J. van den Hoven, P. E. Vermaas, & I. van de Poel (Eds.), *Handbook of Ethics, Values, and Technological Design: Sources, Theory, Values and Application Domains* (pp. 1–22). Springer Netherlands. https://doi.org/10.1007/978-94-007-6994-6_33-1

Vlcek, S., Somerton, M., & Rayner, C. (2020). Collaborative Teams: Teachers, Parents, and Allied Health Professionals Supporting Students With Autism Spectrum Disorder in Mainstream Australian Schools. *Australasian Journal of Special and Inclusive Education*, *44*(2), 102–115. https://doi.org/10.1017/jsi.2020.11

Wetzels, S. A. J., Kester, L., & van Merriënboer, J. J. G. (2011). Adapting prior knowledge activation: Mobilisation, perspective taking, and learners' prior knowledge. *Computers in Human Behavior*, *27*(1), 16–21. https://doi.org/10.1016/j.chb.2010.05.004

Zeidan, J., Fombonne, E., Scorah, J., Ibrahim, A., Durkin, M. S., Saxena, S., Yusuf, A., Shih, A., & Elsabbagh, M. (2022). Global prevalence of autism: A systematic review update. *Autism Research*, *15*(5), 778–790. https://doi.org/10.1002/aur.2696

Appendices

Appendix A. Fictitious personas inspired by real cases

A.1. Alex

Alex, 15 years old

Background:

- Diagnosed with high level ASD (formerly Asperger), Alex was adopted when they were one.
- Their host family takes care of them daily and has tried to reinforce social links with Alex's teachers and companions thanks to *ToGather* tool.
- Collective work done paid off as Alex starts to chat with other pupils. Parents remain worried as their child will enter high school next year, in a new institution. They are trying to figure out how the tool will help with the transition.

School career:

- Arrived during 2019 school year in middle school: did not speak at all.
- 2020-2021 year: Significant progress as they ask help from teachers when they do not understand and communicate with the other pupils.

Weaknesses:

- No notion of money, costs
- Cannot write a simple letter
- Cannot provide complex explanations
- Does not situate in time, has difficulty to organize his agenda

Main points:

- Security: aware of danger (heat sources, sidewalks, …)
- Autonomy: tries to perform tasks on his own
- Concentration > 20 minutes
- Find their way in space

Progress/Acquisition:

- Emotion expression and recognition
- Able to ask for help
- Communication with other pupils

Objectives/Roadmap:

- Prepare their high school start: inform (train) future teachers to the modalities of the web app
- Continue to work on their emotions and socialization
- Regularly give them logic or mechanics exercises because Alex likes it

A.2. Jessie

Jessie, 15 years old

Background:

- Jessie is a young teenager who has difficulty integrating. They constantly look away and remain isolated in the courtyard. In class, if they do not understand, they answer, but miss the point, which amuses the other students who make fun of them. In the canteen, they tend to come and disturb the others who lose their patience, which does not encourage inclusion, especially since Jessie does not like to share their belongings, nor do they like teamwork. However, Jessie sometimes hugs other when they are sad and is very sensitive to arguments. Jessie starts to cry and blame themselves, even when they are not guilty or responsible. They say “it’s Jessie’s fault. They then need an adult to explain the situation to them to reassure them.
- After making great progress, they started to show unusual emotional reactions, such as crying suddenly for no apparent reason or putting things in their mouth, which they

had not done before. Parents and teachers both wonder what is going on at school and at home. They search the platform for explanatory information

School career:

- Arrived during the 2019 year (in December) in middle school: did not speak at all. Implementation of rituals correctly learned and then performed by Jessie.
- Back to school 2020: Less agitated than last year because more rituals were engaged. + Planning assistance workshops (cooking recipes, etc.) done.

Weaknesses:

- Cannot write or calculate + language disorders (which lead to communication problems)
- Concentration in 10 minutes step
- Say everything they think without realizing the impact it can cause

Main points:

- Enjoy learning although difficult to achieve: ok if ritual and repetition
- Knows the politeness codes. Polite student, raises hand in class to ask for help

Progress/Acquisition:

- Progress in math (Jessie has learned to use a double entry table!)
- Some progress in communication/language according to the aid

Objectives/Roadmap:

- Understand the reasons for Jessie's change in behavior. Needs to continue the accompaniment via the aid who seems very "leading" according to the IEP

NB: often has a col (plan to have more tissues in the classroom)

A.3 Jupiter

Jupiter, 12 years old

Background:

- While Jupiter mostly stays out of the groups, they love coming to the inclusion class. Often say that they are happy to be here in the third person: “Jupiter is happy!”
- “*No need to speak, their beautiful blue eyes speak for them in case of difficulty*” explains the history-geography teacher. According to the teacher, Jupiter never really shows any anxiety because they do not show any stereotypical movements.
- In reality, Jupiter is very introverted. They do not dare to express themselves or ask for help. So, the parents alerted the teacher after reading their comments on the platform. After the misunderstanding was resolved, it was decided to teach Jupiter to raise their finger, which was quickly acquired. However, the inclusion class is very undisciplined. How do you say raise your hand when other students do not respect politeness codes and speak without permission? After a few unsuccessful attempts with the other students, the teacher concludes that it would be better to place Jupiter in another class that is less dissipated because the class is not adapted to Jupiter’s problems.

School career:

- Started to come in inclusion class in September 2020 in 6th grade
- 2020-2021 year: Awareness by the teacher after alerting the parents

Weaknesses:

- Rather messy student with little vocabulary. Teachers regularly have to explain and repeat words. They do not give Jupiter homework
- Low self-esteem

Main points:

- Enjoys learning. No difficulty in learning school routines
- Very sharing: puts their tablet between them and their classmate
- Pleasant company for teachers and other students

Progress/Acquisition:

- Learned to ask for help by raising a finger

Objectives/Roadmap:

- Decide whether or not Jupiter should change classes next year

Appendix B. Images of the final release mock-up

B.1. Global features

Figure 1: *Authentication*

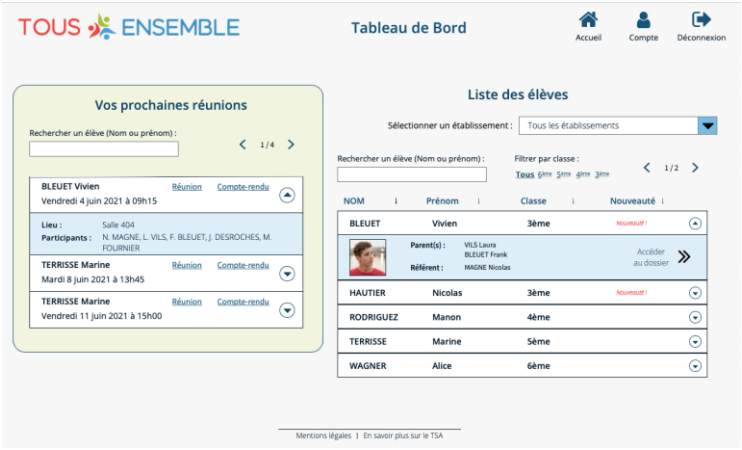

Figure 2: *Dashboard*

Figure 3: *My account*

B.2. Tabs of a pupil's file

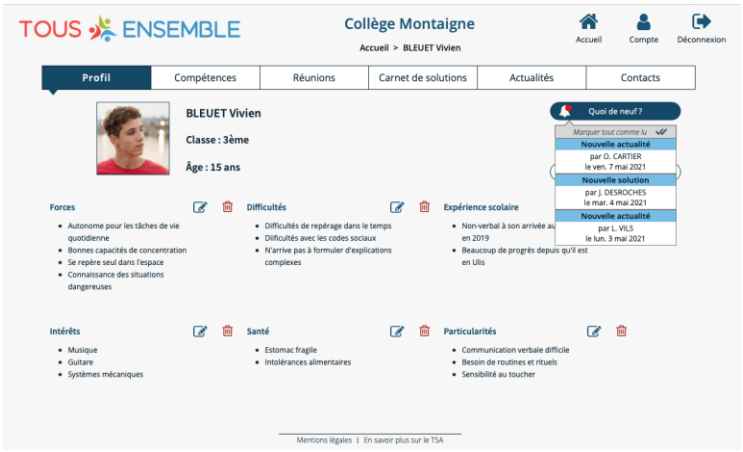

Figure 1: *Profile*

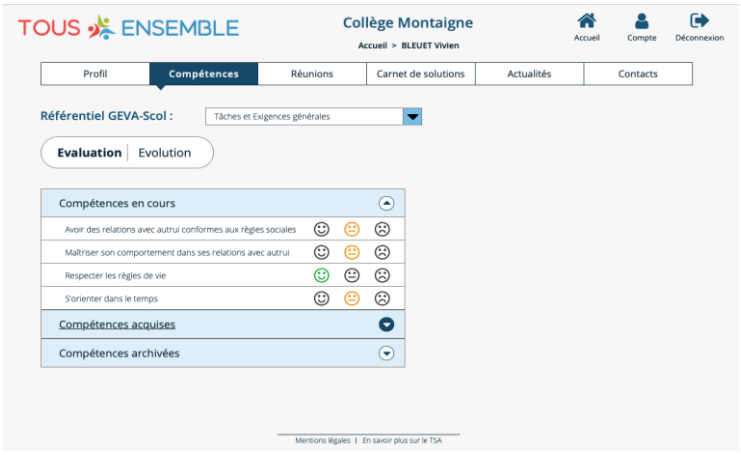

Figure 2: *GEVA-Sco Skills – Evaluation*

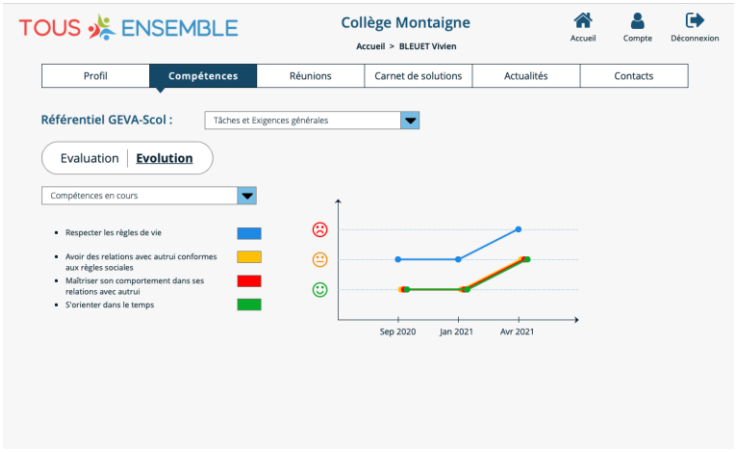

Figure 3: *Skills – Evolution*

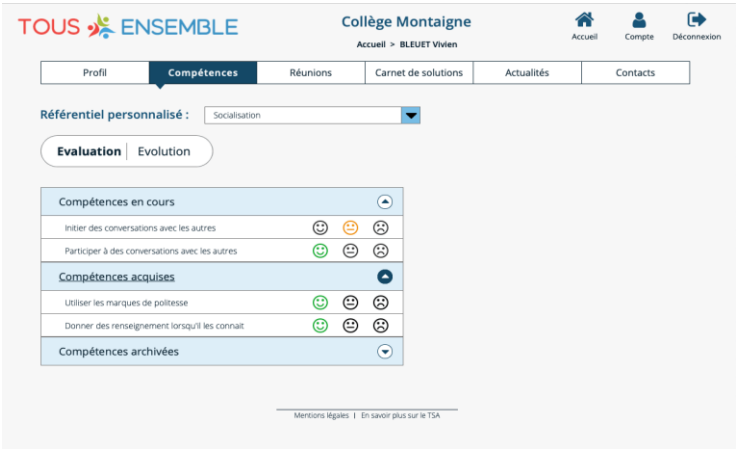

Figure 4: *Personalized Skills - Evaluation*

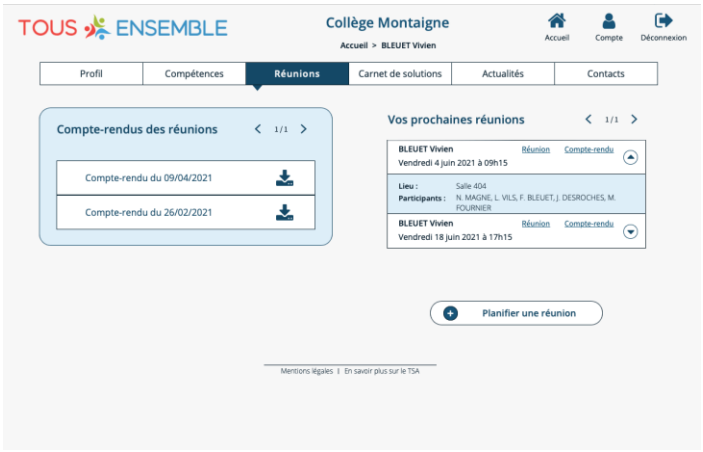

Figure 5: *Meetings*

Figure 6: *A meeting report*

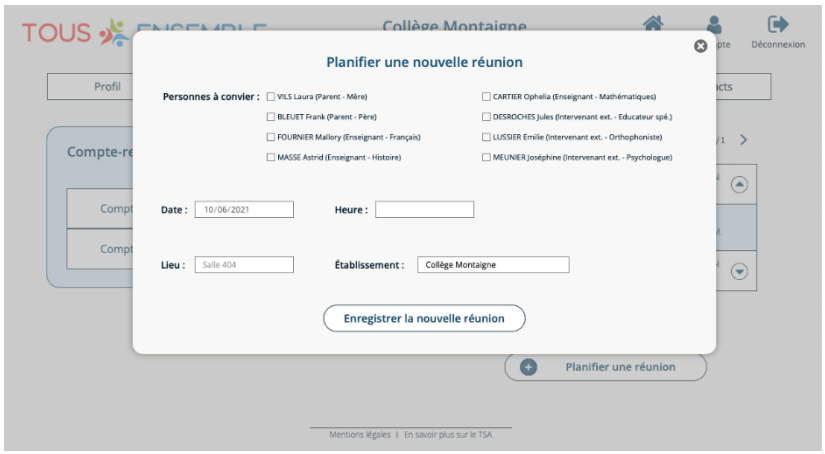

Figure 7: *Meetings – Plan one*

Figure 8: *Compendium of tips*

Figure 9: *News*

Figure 10: *Contacts*

B.3. Editing information

Figure 11: *Profile – Edit a category*

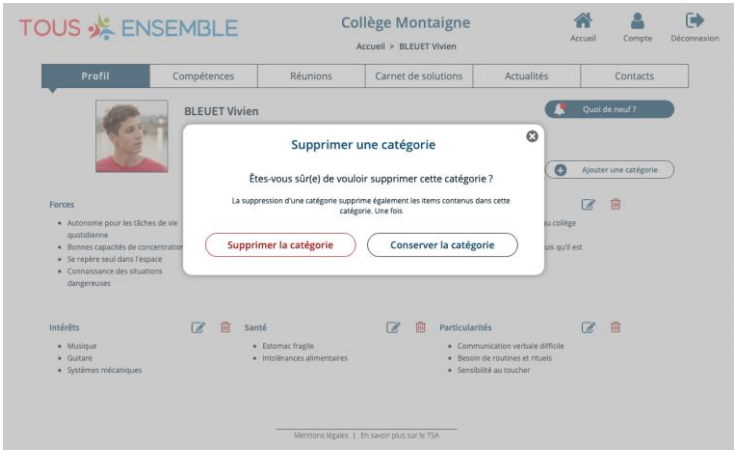

Figure 12: *Profile – Delete a category*

Figure 13: *Solution – Add a new one*

Appendix C. Detailed description of scenarios with optimal active interactions needed

C.1 Parents' scenario

Scenario	Page	Action	Number of interactions
1 (Alex)	Log-in	Log-in	5
	Dashboard	Read a meeting date	0
		Access to the folder	2
	Profile	Read notification (What's up)	1
		Add an item to the profile	4
		Read an item in the profile	0
	Skills	Access to Geva-Sco skills	2
		Read name and evaluation of a skill	3
		Access to custom skills	2
		Read name and evaluation of a skill	3
	Meetings	Access to the tab	1
		Access to and read a meeting report	1
		Read a meeting place	1
	Compendium of tips	Access to the tab	1
		Read a tip item	0

Scenario	Page	Action	Number of interactions
2 (Jupiter)	Log-in	Log-in	5
	Dashboard	Read a meeting date	0
		Access to the folder	2
	Profile	Read notification (What's up)	1
		Add an item to the profile	4
		Read an item in the profile	0
	Skills	Access to Geva-Sco skills	2
		Read name and evaluation of a skill	3
		Access to custom skills	2
		Read name and evaluation of a skill	3
	Meetings	Access to the tab	1
		Read a meeting date	0
		Read meeting participants	1
	Compendium of tips	Access to the tab	1
		Read a tip item	0

Scenario	Page	Action	Number of interactions
3 (Jessie)	Log-in	Log-in	5
	Dashboard	Read a meeting date	0
		Access to the folder	2
	Profile	Read notification (What's up)	1
		Add an item to the profile	4
		Read an item in the profile	0
	Skills	Access to Geva-Sco skills	2
		Read name and evaluation of a skill	3
		Access to custom skills	2
		Read name and evaluation of a skill	3
	Meetings	Access to the tab	1
		Read a meeting date	0
		Read meeting participants	1
	Compendium of tips	Access to the tab	1
		Read a tip item	0

		Give Feedback to a tip item	1
	News	Access to the tab	1
		Read a news item	0
		Add a news item	4
	Contacts	Access to the tab	1
		Read a person's name	0
		Read a person's phone number	1
	Log-out	Log-out	1
TOTAL			**35**

		Give Feedback to a tip item	1
	News	Access to the tab	1
		Read a news item	0
		Add a news item	4
	Contacts	Access to the tab	1
		Read a person's name	0
		Read a person's phone number	1
	Log-out	Log-out	1
TOTAL			**34**

		Add a tip item	6
	News	Access to the tab	1
		Read a news item	0
		Add a news item	4
	Contacts	Access to the tab	1
		Read a person's name	0
		Read a person's phone number	1
	Log-out	Log-out	1
TOTAL			**39**

C.2 Teachers' scenario

Scenario	Page	Action	Number of interactions
1 (Alex)	Log-In	Log-In	5
	Dashboard	Access to the folder	2
		Read notification (What's up)	1
	Profile	Add an item to the profile	4
		Read an item in the profile	0
		Access to Geva-Sco skills	2
	Skills	Read name and evaluation of a skill	3
		Access to custom skills	2
		Read name and evaluation of a skill	3
		Access to the tab	1
	Meetings	Read a meeting date	0
		Add a new meeting	10
		Access to the tab	1
	Compendium of tips	Read a tip item	0
		Delete a tip item	2

Scenario	Page	Action	Number of interactions
2 (Jupiter)	Log-In	Log-In	5
	Dashboard	Read a meeting date	0
		Access to the folder	2
	Profile	Add an item to the profile	4
		Read an item in the profile	0
		Access to Geva-Sco skills	2
	Skills	Read name and evaluation of a skill	3
		Access to custom skills	2
		Read name and evaluation of a skill	3
		Access to the tab	1
	Meetings	Read a report date	0
		Read meeting participants	1
		Access to the tab	1
	Compendium of tips	Read a tip item	0
		Give feedback to a tip item	1

Scenario	Page	Action	Number of interactions
3 (Jessie)	Log-In	Log-In	5
	Dashboard	Access to the folder	2
		Read notification (What's up)	1
	Profile	Delete an item to the profile	4
		Read an item in the profile	0
		Access to Geva-Sco skills	2
	Skills	Read name and evaluation of a skill	3
		Access to custom skills	2
		Read name and evaluation of a skill	3
		Access to the tab	1
	Meetings	Read a report date	0
		Add a new meeting	10
		Access to the tab	1
	Compendium of tips	Read a tip item	0
		Give feedback to a tip item	1

		Access to the tab	1
	News	Read a news item	0
		Delete a news item	2
		Access to the tab	1
	Contacts	Read a person's name	0
		Read a person's phone number	1
	Log-out	Log-out	1
TOTAL			**42**

		Access to the tab	1
	News	Read a news item	0
		Add a news item	4
		Access to the tab	1
	Contacts	Read a person's name	0
		Read a person's phone number	1
	Log-out	Log-out	1
TOTAL			**33**

		Access to the tab	1
	News	Read a news item	0
		Delete a news item	2
		Access to the tab	1
	Contacts	Read a person's name	0
		Read a person's phone number	1
	Log-out	Log-out	1
TOTAL			**41**

C.3 External actors' scenario

Scenario	Page	Action	Number of interactions
1 (Alex)	Log-In	Log-In	5
	Dashboard	Access to the folder	2
		Read notification (What's up)	1
	Profile	Add an item to the profile	4
		Read an item in the profile	0
		Access to Geva-Sco skills	2
	Skills	Read name and evaluation of a skill	3
		Access to custom skills	2
		Read name and evaluation of a skill	3
		Access to the tab	1
	Meetings	Read a meeting date	0
		Add a new meeting	10
		Access to the tab	1
	Compendium of tips	Read a tip item	0
		Delete a tip item	2

Scenario	Page	Action	Number of interactions
2 (Jupiter)	Log-In	Log-In	5
	Dashboard	Read a meeting date	0
		Access to the folder	2
	Profile	Add an item to the profile	4
		Read an item in the profile	0
		Access to Geva-Sco skills	2
	Skills	Read name and evaluation of a skill	3
		Access to custom skills	2
		Read name and evaluation of a skill	3
		Access to the tab	1
	Meetings	Read a report date	0
		Read meeting participants	1
		Access to the tab	1
	Compendium of tips	Read a tip item	0
		Give feedback to a tip item	1

Scenario	Page	Action	Number of interactions
3 (Jessie)	Log-In	Log-In	5
	Dashboard	Access to the folder	2
		Read notification (What's up)	1
	Profile	Delete an item to the profile	4
		Read an item in the profile	0
		Access to Geva-Sco skills	2
	Skills	Read name and evaluation of a skill	3
		Access to custom skills	2
		Read name and evaluation of a skill	3
		Access to the tab	1
	Meetings	Read a report date	0
		Add a new meeting	10
		Access to the tab	1
	Compendium of tips	Read a tip item	0
		Give feedback to a tip item	1

		Access to the tab	1
	News	Read a news item	0
		Delete a news item	2
		Access to the tab	1
	Contacts	Read a person's name	0
		Read a person's phone number	1
	Log-out	Log-out	1
TOTAL			**42**

		Access to the tab	1
	News	Read a news item	0
		Add a news item	4
		Access to the tab	1
	Contacts	Read a person's name	0
		Read a person's phone number	1
	Log-out	Log-out	1
TOTAL			**33**

		Access to the tab	1
	News	Read a news item	0
		Delete a news item	2
		Access to the tab	1
	Contacts	Read a person's name	0
		Read a person's phone number	1
	Log-out	Log-out	1
TOTAL			**41**

Appendix D. Descriptive and two-way ANOVA statistics

D.1 Mean, Standard deviation and Two-way ANOVA Statistics for objective measures

Variable	Whole		Parents		Teachers		External		ANOVA			
	M	SD	M	SD	M	SD	M	SD	Effect	F	df	$\eta 2$
Efficacy												
Proportion of correct responses												
Time 1	0.875	0.063	0.884	0.069	0.848	0.063	0.892	0.056	Role	0.286	2,18	0.017
Time 2	0.930	0.053	0.945	0.040	0.944	0.064	0.901	0.047	Time	19.467***	2,36	0.324
Time 3	0.960	0.049	0.963	0.039	0.970	0.063	0.949	0.048	Role x Time	1.865	4,36	0.084
Completion time (ss.ms)												
Time 1	754.844	193.519	833.404	162.704	671.11	186.389	760.017	219.561	Role	1.396	2,18	0.079
Time 2	534.245	104.640	502.708	46.989	505.186	135.686	594.842	97.211	Time	20.842***	2,36	0.340
Time 3	513.223	191.683	459.066	108.062	465.352	124.924	615.253	277.399	Role x Time	1.585	4,36	0.073
Efficiency												
Active interactions per minute												

Time 1	4.222	1.004	3.495	0.773	4.837	1.070	4.336	0.740	Role	2.399	2,18	0.140
Time 2	4.470	0.800	4.536	0.512	4.737	1.179	4.135	0.510	Time	15.545***	2,36	0.252
Time 3	5.569	1.524	5.269	1.078	6.408	1.482	5.029	1.766	Role x Time	1.777	4,36	0.071
Performance score (speed/optimal)												
Time 1	0.614	0.123	0.531	0.073	0.660	0.127	0.650	0.131	Role	0.740	2,18	0.032
Time 2	0.655	0.111	0.706	0.117	0.637	0.093	0.622	0.119	Time	13.880***	2,36	0.315
Time 3	0.806	0.166	0.818	0.175	0.867	0.103	0.732	0.200	Role x Time	2.162	4,36	0.125

Note: *p < .05, **p < .01, ***p<.001

D.2 Mean, Standard deviation and Two-way ANOVA Statistics for subjective measures

Variable	Whole		Parents		Teachers		External		ANOVA			
	M	SD	M	SD	M	SD	M	SD	Effect	F	df	$\eta 2$
System Usability Scale (/100)	92.024	8.047	89.643	11.033	92.857	7.962	93.571	4.532	Role	0.447	2, 18	0.047
User Experience Questionnaire												
Pragmatic Value (/3)	2.385	0.642	2.119	0.878	2.571	0.496	2.464	0.478	Role	0.761	2, 18	0.060
Hedonic Value (/3)	1.750	0.808	1.607	1.081	1.839	0.756	1.804	0.633	Dim	14.538***	1.49, 26.89	0.167
Attractiveness (/3)	2.373	0.643	2.095	0.952	2.595	0.302	2.429	0.480	Role x Dim	0.193	2.99, 26.89	0.005
Self-Determination												
Average Total score (/4)	3.847	0.257	3.952	0.059	3.841	0.256	3.746	0.356	Role	1.147	2, 18	0.094
Autonomy (/4)	3.905	0.261	3.952	0.126	3.905	0.252	3.857	0.378	Dim	1.859	1.92, 34.49	0.019
Competence (/4)	3.825	0.291	3.905	0.162	3.809	0.325	3.762	0.371	Role x Dim	1.409	3.83, 34.49	0.028
Relatedness (/4)	3.810	0.326	4.000	0.000	3.810	0.262	3.619	0.448				
NASA-TLX												
Raw Total Score (/100)	12.976	9.710	8.690	4.326	17.262	12.245	12.976	10.136	Role	1.421	2, 18	0.066

Mental Demand (/100)	21.667	21.985	14.286	15.119	22.857	25.308	27.857	25.142	Dim	7.588*	2.99, 53.88	0.189
Physical Demand (/100)	4.048	4.068	3.571	4.756	3.571	2.440	5.000	5.000	Role x Dim	1.145	5.99, 53.88	0.066
Temporal Demand (/100)	21.905	19.842	12.857	6.362	32.857	20.988	20.000	24.324				
Performance (/100)	9.762	12.696	11.429	14.058	9.286	15.924	8.571	8.997				
Effort (/100)	13.571	13.052	8.571	6.268	20.000	19.579	12.143	8.092				
Frustration (/100)	6.905	11.562	1.429	2.440	15.000	17.321	4.286	4.499				

Note: *p < .05, **p < .01, ***p<.001

Appendix E. Detailed results on scenario-base subjective usability (perceived difficulty during interactions).

		Average	Profile	Skills	Meetings	Compendium of tips	News	Contact
User role	*Parents*	0,63	1,05	1,14	1,05	0,43	0,10	0,00
	Ext. actors	0,82	1,62	2,62	0,10	0,57	0,00	0,00
	Teachers	0,52	1,05	1,43	0,24	0,19	0,14	0,05
Test instance	*T1*	0,92	1,90	2,48	0,52	0,52	0,10	0,00
	T2	0,52	0,81	1,48	0,43	0,24	0,10	0,05
	T3	0,52	1,00	1,24	0,43	0,43	0,05	0,00
TOTAL		0,65	1,24	1,73	0,46	0,40	0,08	0,02

Table 1: Scenario-based subjective difficulty scores – marginal means

User role	Test instance	Average	Profile	Skills	Meetings	Compendium of tips	News	Contact
Parents	*T1*	0,95	2,29	1,71	1,14	0,43	0,14	0,00
	T2	0,52	0,71	1,00	0,86	0,43	0,14	0,00
	T3	0,40	0,14	0,71	1,14	0,43	0,00	0,00
External actors	*T1*	0,93	1,57	2,86	0,14	1,00	0,00	0,00
	T2	0,67	1,14	2,57	0,14	0,14	0,00	0,00
	T3	0,86	2,14	2,43	0,00	0,57	0,00	0,00
Teachers	*T1*	0,88	1,86	2,86	0,29	0,14	0,14	0,00
	T2	0,36	0,57	0,86	0,29	0,14	0,14	0,14
	T3	0,31	0,71	0,57	0,14	0,29	0,14	0,00

Table 2: Scores for each user role depending on the time point (test instance)