Evidence-based development and first usability testing of a social serious game based multi-modal system for early screening for atypical socio-cognitive development

A private copy of the author-final manuscript.
Reference of the final, published version:
Gyori, M, Borsos, Zs, Stefanik, K (2015): Evidence-based development and first usability testing of a social serious game based multi-modal system for early screening for atypical socio-cognitive development. In: Sik-Lányi, C., Hoogerwerf, E-J, and Miesenberger, K. (eds.) *Assistive Technology: Building Bridges.* Studies in Health Technology and Informatics, 217. pp. 48-54. IOS Press, Amsterdam.
The final publication is available at IOS Press via http://dx.doi.org/10.3233/978-1-61499-566-1-48

Miklos GYORI[a,1], Zsófia BORSOS[a] and Krisztina STEFANIK[b]
[a] *Institute for the Psychology of Special Needs, ELTE University, Budapest, Hungary*
[b] *Institute of Special Education for Atypical Cognition and Behavior, ELTE University, Budapest, Hungary*

Abstract. At current, screening for, and diagnosis of, autism spectrum disorders (ASD) are based on purely behavioral data; established screening tools rely on human observation and ratings of relevant behaviors. The research and development project in the focus of this paper is aimed at designing, creating and evaluating a social serious game based multi-modal, interactive software system for screening for high functioning cases of ASD at kindergarten age. The aims of this paper are (1) to summarize the evidence-based design process and (2) to present results from the first usability test of the system. Game topic, candidate responses, and candidate game contents were identified via an iterative literature review. On this basis, the 1st partial prototype of the fully playable game has been created, with complete data recording functionality but without the decision making component. A first usability test was carried out on this prototype (n=13). Overall results were unambiguously promising. Although sporadic difficulties in, and slightly negative attitudes towards, using the game occasionally arose, these were confined to non-target-group children only. The next steps of development include (1) completing the game design; (2) carrying out first large-*n* field test; (3) creating the first prototype of the decision making component.

Keywords. Autism spectrum disorders, early screening, evidence-based design, eye-tracking, serious game, usability test.

---

[1] Corresponding author, Institute for the Psychology of Special Needs, ELTE University, Ecseri road 3, H-1097 Budapest, Hungary; E-mail: gyorimiklos@elte.hu.

1. Introduction

Autism spectrum disorders (ASD) are defined and diagnosed on the basis of limitations in reciprocal social interactions including social communication, and in adaptive, flexible organization of behaviors and interests, including atypical sensory sensitivities in a part of cases. Symptoms as well as support and educational needs are present from early childhood on, and persist throughout life [1]. Now-standard interventions are evidence-based, complex psycho-educational approaches, used in highly individualized ways to help affected people experience and increase their autonomy, competence and development [2].

Early identification is a key goal, since the earlier the adequate intervention begins, the more positive the long-term outcome is [3]. To date, several psychometric tools have been developed to increase recognition rate of high probability cases in early childhood [4]. Recent surveys, however, still show that there are two major waves of diagnosis in terms of the age when diagnosis is assigned: more severe cases are identified largely between 30 and 60 months of age, but so-called 'high functioning' (henceforth: HF) cases, including cases with Asperger's syndrome, typically receive diagnosis later; most of them at school-age [5][6][7]. Bringing this later wave to an earlier age therefore remains an important goal.

At current, screening for, and diagnosis of, ASD are based on purely behavioral data. Widely used screening tools are still 'low-tech' in the sense that they rely on human observation and ratings of relevant behaviors. They either collect information from parents or other informants from the child's close social environments, and so-gained data are then rated by a professional, or it is the professional who both makes the observations and rates observation data [4]. Several research and development projects have been aimed at developing screening and diagnostic tools that utilize up-to-date technology. In terms of the technological basis, such attempts have included neuro-imaging, eye-tracking, decision support systems, machine learning, and others, in various combinations with each other and with established screening tools [8][9][10][11]. To our knowledge, none of such technology-based screening tools has been introduced into daily practice.

2. Aims

The overall aim of the R+D project which is in the focus of this paper is designing, creating and evaluating a social serious game based, multi-modal, interactive screening tool (a software system) for screening for HF cases of ASD at kindergarten age. The specific aims of the present paper are (1) to summarize the evidence-based process of designing this system and (2) to present results from the first usability test of its first partial but fully playable prototype.

3. Methods of the design process

3.1. The target usage scenario

The design process began by identifying a target usage scenario. According to this, an adult user with an assistant-level competence would manage the low-cost, technologically robust, portable, autonomously functioning screening system which runs on a laptop computer equipped with eye-tracking hardware and software. The system would offer the child user the possibility to play with a motivating game for about 15-30 minutes, alone and independently. During the game the system would record (a) the overt behavioral responses of the child player (as provided via the mouse or the touch-screen); (b) her/his visual scanning behaviors, via the eye-tracking application; and (c) the emotional facial expressions of the child, via a face camera. After completing the game, the system would then make an estimate about the probable risk of ASD being present at the child user, solely on the basis of an autonomous, intelligent processing of the recorded data. Importantly, the role of the adult assistant user would be confined to setting up the system and supervising the otherwise autonomous game playing.

3.2. Procedures of designing

After formulating the target user scenario, an iterative literature review was implemented in order to identify

- specific candidate behavioral responses that could serve with relevant input data for the decision process, in the given technological context;
- specific candidate game contents ('presses') that could potentially evoke such responses;
- and a suitable overall game theme that could serve with a framework for these presses and with adequate motivation for the children in the target group to play with the game.

Two broad fields of evidence-based literature were systematically reviewed: (1) that of the early diagnosis and screening for ASD; and (2) that of clinical and experimental studies on behavioral responses that are significantly different in children with ASD and in neurotypical children at the target age. Results of this review were iteratively reported to a broader expert team which discussed them and identified promising behaviors, candidate contents as potential presses, and candidate overall game topics. Finally, a narrower expert team of the present authors made the final decisions about the components to be implemented in the first prototype.

After making the above decisions, a similarly iterative next phase of the design process took place. In this, the detailed script of the game has been elaborated. It was done on two levels. First, the overall structure of the game script was fixed, breaking it down into scenes. Then a script for each scene has been elaborated, in the form of 'event trees' (story graphs), where each event and each transition condition between events (e.g., elapsed time and/ or player action and/or player gaze focus) was precisely defined.

After completing the scripting phase of the design process, graphic elements of the game were designed in a similarly iterative fashion: the sketches created by the graphic designer were periodically reviewed by the expert team. Candidate designs chosen by the expert team were further developed into the final graphic elements of the game.

4. Results of the design process

The game narratives from an experimental developmental psychological study by Sodian and Frith on strategic deception and sabotage in typically developing and autistic children [12] were chosen as the main theme of the game.

On the basis of identified candidate responses and press contents the detailed, algorithmic game script was elaborated. The game script consisted of an introductory-instruction scene, a closing scene, and 9 scenes serving as 'micro-experiments'. The introductory-instruction scene was designed to teach the child the goals and rules of the game, familiarize the child with the characters, and teach her/him to control the events in the game via mouse actions. A narrator figure introduces these to the player. In the short closing scene the narrator praises the child and says good-bye to her/him.

In 8 of the 9 micro-experiment scenes (social micro-experiments) the player has to influence the behaviors of one of two actors (a competitor or a co-operator) strategically, in order to maximize own reward. Mostly (but not only) these scenes contain those presses that are expected to evoke behavioral responses potentially relevant to making the screening estimates. Another micro-experiment does not expect strategic social behavior from the part of the player, but contain a few purely perceptual presses that are expected to evoke (probabilistically) autism-specific visual scanning responses (perceptual micro-experiment).

5. The first prototype

Based on the complete game script, the first, partial prototype of the fully playable game has been developed, with full data recording functionality. Five of the nine micro-experiments were implemented at this phase: the perceptual and four social micro-experiments. The introductory-instruction and the closing scenes were also operationalized. The game software was developed for the Unity game engine (Unity Technologies). A special-purpose authoring software tool was developed to generate the JSON script for the Unity engine, and another software to serve as an interface between the eye-tracker software (EyeGaze by LC Technologies) and the Unity engine, and to implement data recording; see also the *Acknowledgements*. The decision making component has not yet been designed and implemented.

6. First usability test: methods

The first low-*n* usability test has been carried out in order to test the technical stability and the data recording functionality of the game; to gain data on the usability of the prototype and on the motivating potential of the game; to identify design features that may hamper children's immersion; and to determine expectable game time. Testing screening functionality was not aimed in the first usability test.

6.1. Subjects

Five children with ASD, 4 children with developmental language impairment and 4 typically developing children participated (mean age was 6.3 years, intentionally slightly above the final target group's age). From among the three subject groups, children with developmental language impairment are not within the pre-defined target group; they were involved in order to gain insights about the role of receptive linguistic competence of the player in her/his understanding of, and success in, the game.

6.2. Equipment

A desktop-mounted binocular EyeFollower II eye-tracker PC from LC Technologies was used, with 120 c/s data acquisition frequency (provided by measuring the gaze focus of the two eyes alternately at 60 c/s speed). This eye-tracker equipment allowed the subjects relatively free head movements (in a space of 76 x 51 x 40 cm) and, accordingly, the tests did not require fixing their head by any means. Children could control the game by a small-sized mouse. A face camera recorded players' facial expressions, another camera made a full-body recording about the players' behaviors in the play situation.

6.3. Setting, procedure and materials

Tests took place in a small room in the presence of an experimenter, another researcher sitting in the background and making observations, and the child's caretaker, if she/he wished to be present during the test – otherwise she/he stayed in a neighboring room.

An informed, written consent was collected from the child's caretaker and a verbal consent from the child before starting the test. Following a short session to practice mouse control in a simple 'find-and-click' game, developed for this purpose, children played independently with the first prototype of the screening game. The experimenter intervened only if and when the child asked for her assistance, of when (occasional and minor) technical difficulties made it necessary.

Background developmental and digital proficiency data were collected via a short parental questionnaire, administered in the recruitment phase. Mouse responses and gaze behaviors were recorded by the game interface software. Emotional facial expressions and full-body behaviors were video-recorded for later processing. At the end of the game, a short user experience questionnaire, devised specifically for this purpose, was administered by the experimenter.

7. Results

All children successfully completed the game, and all but one (having developmental language impairment) performed at ceiling or close-to-ceiling level. Mean total playing time was 706 seconds, well within the intended range. In the post-test user experience questionnaire all but one participant children indicated that they found the game entertaining and would play it again. The only child who gave below-ceiling ratings on some aspects of the game was highly above the final target user age. Eye-gaze, mouse response, and emotional facial expression raw data were successfully recorded.

Observations revealed a few user behaviors that require further considerations in the next phases of development. Namely, a few children initiated interactions with the experimenter (such as joint attention for sharing experience) or turned to the experimenter for information or help. These behaviors are seemingly in conflict with our intention to make the game fully playable alone and independently.

On the other hand, no significantly maladaptive design feature was revealed by any user response or other behavior.

8. Discussion

Although the first usability test was a low-*n* study, this seems adequate for exploring fundamental user experiences and resulting user attitudes. This is partly so as a fully playable version of the game was tested in which all the user-relevant features were implemented; also, this sample size is not unusual in testing for similar purposes in similar phases of development [13][14].

Results seem unambiguously promising. The child who had difficulties in understanding the game fully, and the child who expressed devaluating attitudes towards the game were not members of the final target group. Although certain user behaviors were seemingly in conflict with the fully independent game play situation targeted in the final user scenario, there are several potential ways to attempt to manage these (e.g., via emphasizing autonomous playing in the instructions, and/or via implementing Wizard-of-Oz elements to handle occasional user difficulties, etc.).

9. Conclusions and perspectives

A successful first usability test confirmed that the first partial, but fully playable prototype of the game is engaging enough for children in the target group, its successful completion is well within their abilities, and that it is able to record data as planned. No need for any significant change in game design emerged. Next future steps of development include (1) implementing five further micro-experiments in order to complete the game design; (2) carrying out first large-*n* field test for exploring usability in the target usage scenario; and using data collected in this forthcoming field test as input to (3) creating the first prototype of the decision making component of the screening system.

Acknowledgements

Certain elements of the project reported here has been funded by a grant within the EIT ICT Labs Hungarian Node (PI: András Lőrincz at Faculty of Informatics, ELTE University), and via a TÁMOP grant, co-financed by the European Union and the government of Hungary (TÁMOP 4.2.1./B-09/KMR-2010-0003). Authors wish to thank András Lőrincz for his various contributions and support; Tibor Gregorics (Faculty of Informatics, ELTE University) for coordinating software development; and members of the Autism Group (TAM) at 'Bárczi Gusztáv' Faculty of Special Education, ELTE University. Authors are highly grateful to Fanni Varga for her immense work in creating the graphic elements of the game.

References

[1] APA [American Psychiatric Association] *Diagnostic and Statistical Manual of Mental Disorders, Fifth edition (DSM-5).* American Psychiatric Association, Washington DC (2013).

[2] P. Howlin, The effectiveness of interventions for children with autism, *J Neural Transm. Suppl.* **69** (2005), 101-119.

[3] S. Eikeseth, Intensive early intervention, In: J. L. Matson & P. Strumey (eds.) *International Handbook of Autism and Pervasive Developmental Disorders*, Springer New York, 2011, pp. 321-338.

[4] P. Garcı´a-Primo, A. Hellendoorn, T. Charman, H. Roeyers, M. Dereu, B. Roge, S. Baduel, F. Muratori, A. Narzisi, E. Van Daalen, I. Moilanen, M. Posada de la Paz, R. Canal-Bedia, Screening for autism spectrum disorders: State of the art in Europe, *European Child & Adolescent Psychiatry* **23** (2014), 1005-1021.

[5] E. Fombonne, L. Heavey, L. Smeeth, L. C. Rodrigues, C., Cook, P. G. Smith, L. Meng, A. J. Hall, Validation of the diagnosis of autism in general practitioner records, *BMC Public Health* **4** (2004), doi:10.1186/1471–2458–4–5.

[6] D. Mandell, M. Mayrali, N. Zubritsky, & C. Zubritskx, Factors associated with Age of Diagnosis among Children with Autism Spectrum Disorders, *Pediatrics* **116** (2005), 1480–1486.

[7] M. Noterdaeme, & A. Hutzelmeyer-Nickels, Early symptoms and recognition of pervasive developmental disorders in Germany, *Autism* **14** (2010), 575–588.

[8] J.A. Kientz, M.S. Goodwin, G.R. Hayes, & G.D. Abowd, Interactive Technologies for Autism. Synthesis Lectures on Assistive, Rehabilitative, and Health-Preserving Technologies. **2** (2013) 1-177.

[9] N.S. Bauer, L.A. Sturm, A.E. Carroll, & S.M. Downs. Computer Decision Support to Improve Autism Screening and Care in Community Pediatric Clinics, *Infants & Young Children*, **26** (2013), 306-317.

[10] A. Klin, Identifying risk for ASD in the first year of life. Conference keynote talk. IX. Autism-Europe Conference, Catania, Italy, 8-10 October, 2010.

[11] C. Ecker, A. Marquand, J. Mourão-Miranda, P. Johnston, E. M. Daly, M.J. Brammer, . . . D.G.M. Murphy, Describing the Brain in Autism in Five Dimensions—Magnetic Resonance Imaging-Assisted Diagnosis of Autism Spectrum Disorder Using a Multiparameter Classification Approach. *The Journal of Neuroscience* **30** (2010), 10612-10623.

[12] B. Sodian, U. Frith, Deception and Sabotage in Autistic, Retarded and Normal Children, *Journal of Child Psychology and Psychiatry* **33** (1992), 591-605.

[13] R. Khaled, A. Vasalou, Bridging serious game and participatory design, *International Journal of Child-Computer Interaction* **2** (2014), 93-100.

[14] O. De Troyer, E. Janssens, Supporting the requirement analysis phase for development of serious games for children, *International Journal of Child-Computer Interaction* **2** (2014), 76-84.