Eye-tracking test battery for newborns: A pilot feasibility study

Marta Malavolta[1,*], Andrea De Gobbis[2], Emiliano Trimarco[3], Isabel Benavente-Fernández[3,4,5], Simón Pedro Lubián-López[3,4], Vida Groznik[1,2,6] and Aleksander Sadikov[1,2]

[1]*University of Ljubljana, Faculty of Computer and Information Science, Ljubljana, Slovenia*
[2]*NEUS Diagnostics, d.o.o., Ljubljana, Slovenia*
[3]*Biomedical Research and Innovation Institute of Cádiz (INiBICA) Research Unit, Puerta del Mar University, Cádiz, Spain*
[4]*Division of Neonatology, Department of Paediatrics, Puerta del Mar University Hospital, Cádiz, Spain*
[5]*Area of Paediatrics, Department of Child and Mother Health and Radiology, Medical School, University of Cádiz, Cádiz, Spain*
[6]*Faculty of Mathematics, Natural Sciences and Information Technologies, University of Primorska, Koper, Slovenia*

Abstract
Preterm birth is one of the leading causes of neurodevelopmental disabilities. Many efforts have been made to improve the well-being and quality of life of preterm infants and their families, especially during the first months of life. Several authors investigated cognitive impairments in children, such as social disorders or attention deficit, using various remote eye-tracking techniques. However, this tool remains poorly used in newborn infants, particularly in the first three months old children. Therefore, we aim to create a neuropsychological test battery using screen-based eye-tracking that can also be used on the above-mentioned population.

The aim of this study is to analyse the feasibility of the created pilot eye-tracking test battery and the suitability of the different stimuli used. We also investigate how the current paradigm evolved based on observations made during data acquisition, and how it was modified to achieve an appropriate test in terms of composition and length to keep children's attention.

Keywords
Eye-tracking test battery, Cognitive Impairments, Feasibility Study, Premature Children

1. Introduction

Every year 15 million babies worldwide are born preterm, 7.1% of them with some degree of impairment. The World Health Organization (WHO) considers an infant as preterm if

DETERMINED 2022: Neurodevelopmental Impairments in Preterm Children — Computational Advancements, August 26, 2022, Ljubljana, Slovenia
*Corresponding author.
marta.malavolta@fri.uni-lj.si (M. Malavolta); andrea.dg@neus-diagnostics.com (A. De Gobbis); emiliano.trimarco@inibica.es (E. Trimarco); isabel.benavente@uca.es (I. Benavente-Fernández); simonp.lubian.sspa@juntadeandalucia.es (S. P. Lubián-López); vida@neus-diagnostics.com (V. Groznik); aleksander.sadikov@fri.uni-lj.si (A. Sadikov)
0000-0002-7462-0822 (M. Malavolta); 0000-0002-3993-5116 (A. De Gobbis); 0000-0001-9276-1912 (I. Benavente-Fernández); 0000-0001-8925-926X (V. Groznik); 0000-0001-8697-3556 (A. Sadikov)

birth occurs before the 37^{th} week of gestational age (GA). Furthermore, the WHO provided a classification of preterm infants according to the gestational time window, thereby further dividing children into late and moderate preterm, very preterm and extremely preterm [1, 2]. As the time span of gestation decreases (lower GA), the risk of complications caused by preterm birth significantly increases. The primary causes of long-term disabilities, such as behavioural alterations or neurological disorders (e.g. cerebral palsy), are related to brain injury occurring in the neonatal period in the case of premature births [3]. Furthermore, the burden of preterm birth extends beyond the well-being and overall health of the infants themselves as it also includes economic impact on the healthcare system, e.g. in the form of a longer stay in neonatal intensive care units (NICU) as well as an increased overall burden on the family [1, 2].

To date, a troubling dichotomy arises: despite the great improvement of the quality of prenatal care in recent years, which dramatically improved the survival of preterm babies, the diagnosis of motor and cognitive impairments has not progressed at the same pace, mostly relying on monitoring the clinical parameters. Indeed, clinicians need a more robust and timely set of biomarkers to assess the risk of disability of preterm infants to provide early intervention to reduce the developmental delays associated with this condition.

Many authors in the past decades relied on eye-tracking to investigate the possible connection between cognitive deficits and gaze behaviour in response to selected stimuli, both in adults and children. Indeed, eye-tracking has been employed in children to study social orientation [4], attention and memory [5, 6] or pursuit of moving objects [7]. In the specific context of the term- vs. preterm-born infants, Stand-Brodd et al. showed that premature children would appear to have delayed eye movements in following objects in comparison with term-born children of the same age [8]. Other works highlighted that the ability to visually follow a moving object at 4 months of age not only has a robust predictive power for neurodevelopment at 3 years in children born very preterm [9], but can also predict future memory and attention problems at 6 years of age [10]. Nevertheless, very few of these works are performed in the very first months after birth.

Whithin this framework of investigation, we created a neuropsychological test battery to study different types of cognitive processes in children born prematurely, using screen-based eye-tracking. The underlying goal is to give a quantitative and early prediction of possible future impairments. For the development of this prototype, we studied 23 children with an average age of 11 months (std 6 months), including eight babies born prematurely (8 months old, 6 months old CA). The purpose of this study is to analyse the evolution of the test battery and to pinpoint and implement any needed changes, to obtain a tool that can capture and hold the child's attention and at the same time ensure the correct gathering of sufficient data for analysis. Lastly, despite the young age of the studied population, we demonstrate the data acquisition feasibility of the proposed test battery by measuring missing values and attention for each child of the piloting cohort.

The article is organised as follows. First, we analyse the population studied during the implementation of the test battery. In the subsequent section, we introduce the neuropsychological test battery with a brief explanation and representation of the tasks. We explain in detail how we optimised the testing protocol, explaining the technical aspects in the fourth section. Finally, the last paragraph concludes the article.

2. Subjects

During the implementation of the neuropsychological test battery, we tested 23 subjects born at the Hospital Universitario Puerta del Mar, Cadiz (ethical committee code PIEBA 0672-N-22, register number 44.22), of whom 8 were premature. Table 1 summarises the subject data during the implementation of the neuropsychological test battery.

Table 1
Summary of the performed tests with the different setups, including information about the involved subjects. Total test times and inattention of the children during the test are also reported.

Prototype (test length)	*Setup*	*Sex*	*Preterm*	*Age [CA] (months)*	*Calibration*	*Test time (min)*	*Inattention time (min)*	*Missing values (%)*
#1 (10 min)	All-in-one room, 24-inch monitor	F	No	12	Completed	04:42	02:29	57%
	Partitioned room, 24-inch monitor, Artificial light	M	No	12	Completed	02:53	00:30	17%
		F	No	6	Completed	02:32	00:29	20%
		M	Yes	5 [3]	Completed	03:10	00:57	27%
#2 (3 min)	Partitioned Room, 27-inch Monitor, Artificial Light, Covered Table	M	No	9	Failed			
		M	No	7	Completed	03:00	01:33	52%
		M	No	24	Completed	03:00	00:31	17%
		M	No	12	Completed	03:00	00:52	29%
		F	No	18	Failed			
		F	No	6	Completed	03:00	00:16	9%
	Partitioned Room, 24-inch Monitor, Artificial Light, Covered Lateral Table	M	No	9	Completed	03:00	01:21	45%
		F	No	18	Completed	03:00	00:23	13%
		M	Yes	15 [12]	Completed	02:17	00:45	35%
		F	Yes	8 [6]	Completed	03:00	00:16	9%
#3 (4 min)	Partitioned Room, 24-inch Monitor, Artificial Light, Covered Lateral Table	M	Yes	23 [20]	Completed	04:00	00:48	20%
		F	No	4	Completed	04:00	00:58	24%
		M	Yes	12 [10]	Completed	04:00	00:34	14%
		M	Yes	6 [5]	Completed	04:00	00:02	1%
		M	No	18	Completed	02:26	00:15	19%
		F	Yes	6 [3]	Completed	04:00	00:26	11%
	Partitioned Room, 18-inch Monitor, Artificial Light, Covered Lateral Table	M	No	15	Completed	04:00	01:00	25%
		F	Yes	6 [3]	Completed	**Test aborted due to child crying.*		
		F	No	3	Completed	04:00	02:02	51%

There were 10 females and 13 males, and their ages ranged from 3 to 24 months. The premature babies ranged from 3 months CA to 20 months CA. Only three of the children did not want to take the test (two of these completed the calibration). We also noticed that it is useful to help the children perform the instrument calibration. During the actual test, however, the children were mostly well-engaged. In fact, only in few cases the parents had to call the child's attention to the screen. The initial prototype was updated with minor changes as described later, mostly regarding test design and not task design. The first child differed the most from the other children in terms of the test setup. As a matter of fact, we noticed that there was a need for separation between the examination area, where the child was tested, and the examiners' space. In fact, the child was very distracted as witnessed by the percentage of missing values in Table 1. As can be seen from the table, some differences in the length of the test emerge between the prototypes. Specifically, with prototypes #1 and #2, where the tasks were consecutive, we could get a maximum of three minutes of testing. Instead, with the last prototype, where task repetitions were alternated, we were able to get up to four minutes of testing. The maximum distraction was approximately 25 seconds which corresponds to the length of one task. In conclusion, we also noticed that although it is not evident from the data, from 15-18 months of age it was more challenging to keep the children's attention and therefore, there is likely a need to develop a different test for that age group (and older).

3. The neuropsychological test battery

The test is composed of four different types of tasks (smooth pursuit task, attention task, memory task and social orienting task), taken from the literature and adapted to test the children's cognitive development. Each task is consecutively repeated four times, obtaining the total test battery duration of four minutes. As described in the subsequent paragraph, the procedure proved able to engage the subjects for meaningful periods of time. The calibration process was also largely successful.

3.1. Smooth pursuit task

In the smooth pursuit task, the child has to follow with the gaze a smiling face moving horizontally with sinusoidal movements at a frequency of 0.4 Hz. This task is widely used for the detection of possible cognitive deficits in adults [11]. An illustration of the task is shown in Figure 1. The specific parameters which we plan to sample and analyse are related to how the child follows the moving smiling face: examples include e.g. the quantitative assessment of the anticipation of the movement of the dot, the delay in following the object and the latency in starting the smooth pursuit movement.

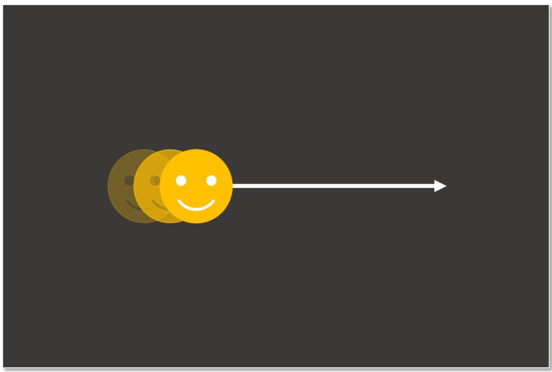

Figure 1: A graphical representation of the smooth pursuit task.

3.2. Attention task

The attention task was created to analyse the differences in the children's response to specific variations of the stimulus, inspired by the IOWA test [12, 13]. In detail, the test first presents a white cross for a fraction of a second (a visual cue) along with a warning sound to keep the child's attention. Subsequently, a black screen is shown briefly, followed by a target image located either at the same location or at the opposite position of the initial white cross. In one version of the task, the white cross is shown on both sides of the screen. Furthermore, an additional baseline test instead consists of only the auditory stimulus prior to the appearance of the target, without the appearance of the cross. The task is visually summarised in Figure 2.

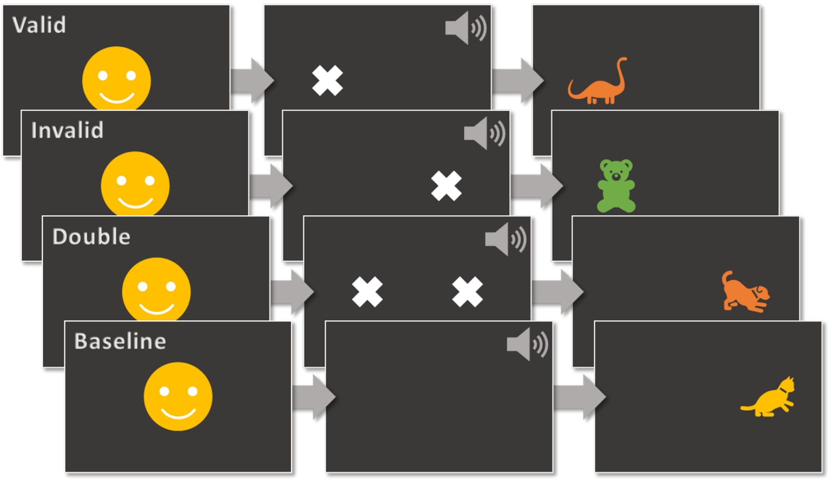

Figure 2: A schematic visualisation of the attention task.

3.3. Memory task

The memory task is inspired by [14] and it checks the predisposition of children to fixate on novel stimuli. Initially, two identical pictures are shown to the child. The screen is subsequently blanked, and then the pictures are presented again, with one of them substituted by a new, unseen image. This test is performed using three types of differences in image content to measure what the child is able to identify, namely differences in colour, shape, and faces. A summarising picture can be found in Figure 3.

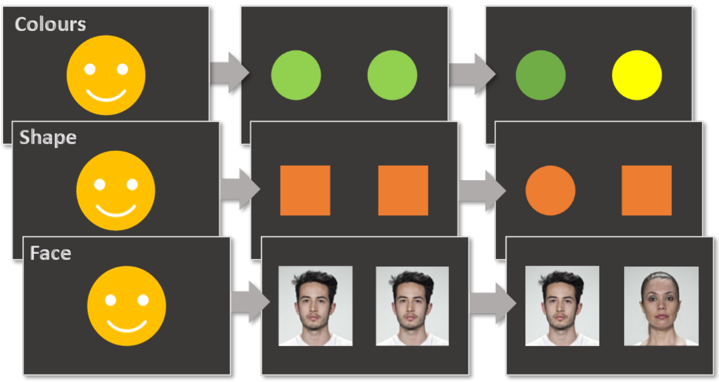

Figure 3: A schematic visualisation of the memory task. The face image dataset was taken from [15].

3.4. Social orienting task

The goal of the social orienting task is to measure if the children display responses to stimuli with social content, and it is inspired by the work of [4, 16]. The child is presented with two pictures, one depicting a person and the other depicting an inanimate object. The test measures the duration of the child's focus on the human picture. This task can be visualised in Figure 4.

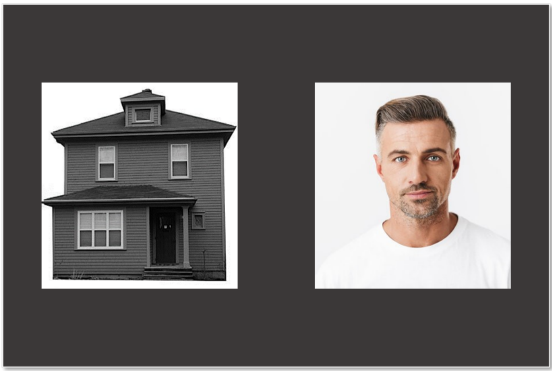

Figure 4: An illustration of the social orienting task. The face image dataset was taken from [15] and the house image dataset was extracted from [17].

4. Optimisation of the testing protocol: operational and technical aspects

4.1. Duration of the testing procedure and repetitions

The first prototype we developed was composed of the four above-mentioned tasks with 10 repetitions per task, leading to a total test length of around 8 minutes. During the initial tests involving children under 12 months of age, we noticed how none of them successfully finished the test as they started to get fussy, or were otherwise easily distracted. This allowed us to estimate the maximum attention span of the tested children to at most 3 minutes. Hence, subsequent tests consisted of a reduced testing protocol with only three repetitions of each task. In most of the cases, the length of this updated test was adequate, but a few infants were still unable to finish the test as they were too active or too nervous. These aspects were hard to control in the experimental setting and, since the tasks were always administered in the same order, this lead to an imbalance in the collected data for different tasks. We thus decided to present the tasks in batches of one task of each type and we repeated the batches for three or four times depending on the prototype. This ensured the availability of experimental data for each individual task.

4.2. Child activity level and distracting elements

In the initial testing configuration, the child was seated on their parent's lap at a distance of 60 cm from the 24-inch computer screen. For collecting the data on the child's eye-movements, we used a Tobii 4C eye-tracker (Tobii AB, Danderyd, Sweden) mounted to the computer screen. The test operators and the other parent were located behind the computer screen, inside the child's field of view. This condition led to the loss of concentration of the children as they were distracted by the people around the testing area. Thus, in the subsequent testing setup, we decided to use a white separator screen (a curtain as seen in Figure 5) to divide the testing area from the operator space. In this way, the children could not see the operators except for the parent who was holding them in their lap. We also decided to increase the room lighting after introducing the separator screen. Lastly, we covered the table in the testing area with a white blanket as the light in the room gave rise to the reflections on the table surface that distracted the child.

4.3. Technical aspects

We noticed that bigger screen sizes increased the occurrence of distractions in the children, especially in the case of the attention task. Therefore, we fixed the computer screen size to 18 inches. We also noticed that the calibration procedure initially did not catch the children's attention as the crosses on the screen were too small in size. Hence, we made the calibration more attractive in subsequent examinations by visualising larger crosses. Also, the presented stimuli during the tests were increased in size to better keep the attention of the children. Lastly, we noticed how the currently used testing protocol was only suitable for children of ages up to 15-18 months since older children are less compliant to the testing procedure and are more easily distracted (e.g. by wanting to touch the screen or wandering around the examination room). The operators repeatedly made this observation during the examinations.

4.4. Room setup

The room setup comprises two separate parts: an examination room and an operator space divided by a white separator screen. The child is inside the examination room and is seated at a distance of 60 cm from the computer screen on his parent's lap. All parents inside the examination room wear sunglasses so their gaze is not interfering with the test. The eye-tracker is held on a 24-inch monitor using magnets (not visible to the child). In this space, the child can only see the parent and not the operators, which stay behind the white separator screen. The lighting in the examination room is artificial and from the side, while there is no light in the operator space. There is a table inside the examination room to the left of the child reflecting the artificial light, and we covered it with a white blanket. The setup is shown in Figure 5.

5. Discussion and Conclusion

In this work, we briefly described our test battery exploiting screen-based eye-tracking for the early detection of neuropsychological impairments in preterm babies. We explained how the test battery evolved since the deployment of the first prototype giving the reasons behind the changes. During the development of the test battery, we managed to increase the attention span of the tested children to up to four minutes, with a progressive improvement in the amount of missing values. Also, based on our findings, the current version of the test battery allows for a

Figure 5: A representation of the final testing setup.

successful calibration of the test in more than 90% of the runs. Furthermore, we highlighted some of the shortcomings and crucial aspects which emerged during the study, such as the physical location of the examiners and operators in the room during the tests, or the need to switch between the different tasks to keep the children's concentration longer. We ultimately demonstrated the feasibility of the current test battery and how it can be successfully applied starting from the first three months of life up to a likely maximum of two years. Lastly, in contrast to our previous belief, we saw that with children older than two years, we necessarily need to change the design of the test to make it dynamic and interactive, using e.g. cartoons to spark and maintain the child's interest.

Our findings described herein pave the way for the refinement and deployment of an optimised test battery suitable for the collection of eye-tracking data in preterm children, with the ultimate goal of early detection of possible developmental impairments.

Acknowledgments

The PARENT project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie – Innovative Training Network 2020, Grant Agreement N° 956394.

References

[1] C. P. Howson, M. V. Kinney, J. E. Lawn, Born too soon: the global action report on preterm birth., volume 25, 2012 ed., March of Dimes, Partnership for Maternal, Newborn and Child Health (PMNCH), Save the Children, WHO. World Health Organization. Geneva, 2012., 2012.

[2] J. E. Lawn, R. Davidge, V. K. Paul, S. von Xylander, J. de Graft Johnson, A. Costello, M. V. Kinney, J. Segre, L. Molyneux, Born Too Soon: Care for the preterm baby, Reproductive Health 10 (2013) S5. doi:10.1186/1742-4755-10-S1-S5.

[3] R. S. Jois, Understanding long-term neurodevelopmental outcomes of very and extremely preterm infants: A clinical review, Aust. J. Gen. Pract. 48 (2019) 26–31. doi:10.31128/AJGP-04-18-4545.

[4] E. J. Telford, S. Fletcher-Watson, K. Gillespie-Smith, R. Pataky, S. Sparrow, I. C. Murray, A. O'Hare, J. P. Boardman, Preterm birth is associated with atypical social orienting in infancy detected using eye tracking, J. Child Psychol. Psychiatry 57 (2016) 861–868. doi:10.1111/jcpp.12546.

[5] S. Ross-Sheehy, S. Schneegans, J. P. Spencer, The infant orienting with attention task: Assessing the neural basis of spatial attention in infancy, Infancy : the official journal of the International Society on Infant Studies 20 (2015) 467–506. doi:10.1111/infa.12087.

[6] S. Ross-Sheehy, B. Eschman, Assessing visual stm in infants and adults: eye movements and pupil dynamics reflect memory maintenance, Visual Cognition 27 (2019) 78–92. doi:10.1080/13506285.2019.1600089.

[7] G. Gredebäck, C. von Hofsten, J. Karlsson, K. Aus, The development of two-dimensional tracking: a longitudinal study of circular pursuit, Exp Brain Res 163 (2005) 204–213.

[8] K. Strand-Brodd, U. Ewald, H. Grönqvist, G. Holmström, B. Strömberg, E. Grönqvist, C. von Hofsten, K. Rosander, Development of smooth pursuit eye movements in very preterm infants: 1. General aspects, Acta Paediatr. 100 (2011) 983–991. doi:10.1111/j.1651-2227.2011.02218.x.

[9] Y. F. Kaul, K. Rosander, C. Von Hofsten, K. S. Brodd, G. Holmström, A. Kaul, B. Böhm, L. Hellström-Westas, Visual tracking in very preterm infants at 4 mo predicts neurodevelopment at 3 y of age, Pediatr. Res. 80 (2016) 35–42. doi:10.1038/pr.2016.37.

[10] Y. F. Kaul, K. Rosander, C. von Hofsten, K. Strand Brodd, G. Holmström, L. Hellström-Westas, Visual tracking at 4 months in preterm infants predicts 6.5-year cognition and attention, Pediatr. Res. (2021) 1–8. doi:10.1038/s41390-021-01895-8.

[11] A. Gerbasi, V. Groznik, D. Georgiev, L. Sacchi, A. Sadikov, Detecting Mild Cognitive Impairment Using Smooth Pursuit and a Modified Corsi Task, in: Springer (Ed.), Int. Conf. Artif. Intell. Med., Springer International Publishing, 2021, pp. 168–172. doi:10.1007/978-3-030-77211-6_19.

[12] S. Ross-Sheehy, S. Schneegans, J. P. Spencer, The Infant Orienting With Attention Task: Assessing the Neural Basis of Spatial Attention in Infancy, Infancy 20 (2015) 467–506. doi:10.1111/infa.12087.

[13] S. Ross-Sheehy, S. Perone, K. L. Macek, B. Eschman, Visual orienting and attention deficits in 5- and 10-month-old preterm infants, Infant Behav. Dev. 46 (2017) 80–90. doi:10.1016/j.infbeh.2016.12.004.

[14] S. Ross-Sheehy, B. Eschman, Assessing visual STM in infants and adults: eye movements and pupil dynamics reflect memory maintenance, Vis. cogn. 27 (2019) 78–92. doi:10.1080/13506285.2019.1600089.

[15] J. H. Filliter, J. M. Glover, P. A. McMullen, J. P. Salmon, S. A. Johnson, The Dal-Houses: 100 new photographs of houses with ratings of typicality, familiarity, and degree of similarity to faces, Behavior Research Methods 48 (2016) 178–183. doi:10.3758/s13428-015-0561-8.

[16] B. Dean, L. Ginnell, V. Ledsham, A. Tsanas, E. Telford, S. Sparrow, S. Fletcher-Watson, J. P. Boardman, Eye-tracking for longitudinal assessment of social cognition in children born preterm, J. Child Psychol. Psychiatry 62 (2021) 470–480. doi:10.1111/jcpp.13304.

[17] L. DeBruine, B. Jones, Face Research Lab London Set (2017). doi:10.6084/m9.figshare.5047666.v5.