USING MULTI-MODAL BIO-DIGITAL TECHNOLOGIES TO SUPPORT THE ASSESSMENT OF COGNITIVE ABILITIES OF CHILDREN WITH PHYSICAL AND NEUROLOGICAL IMPAIRMENTS

Hock Chye Gan

Thesis submitted to the University of Hertfordshire in partial fulfilment of the requirements of the degree of Doctor of Philosophy (PhD)

August, 2014

Abstract

Current studies done using a learning test for children have problems as they only make evaluations of Physically and Neurologically Impaired (PNI) children who can succeed in the test and can be considered as a PASS/FAIL test. This pilot study takes a holistic view of cognitive testing of PNI children using a user-test-device triad model and provides a framework using non-PNI children and adults as controls. Comparisons using adapted off-the-shelf novel interfaces to the computer, in particular, an Electroencephalograph (EEG) head-set, an eye-tracker and a head-tracker and a common mouse were carried out. In addition, two novel multi-modal technologies were developed based on the use of brain-waves and eye-tracking as well as head-tracking technologies to support the study. The devices were used on three tests with increasing cognitive complexity. A self-developed measure based on success streaks (consecutive outcomes) was introduced to improve evaluations of PNI children. A theoretical model regarding a fit of ability to devices was initially setup and finally modified to fit the view of the empirical model that emerged from the outcomes of the study. Results suggest that while multi-modal technologies can address weaknesses of the individual component modes, a compromise is made between the user's ability for multi-tasking between the modes and the benefits of a multi-modal device but the sample size is very small. Results also show children failing a test with a mouse but passing it subsequently when direct communication is used suggesting that a device can affect a test for children who are of a developing age. This study provides a framework for a more meaningful conversation between educational psychologists as well as other professionals and PNI parents because it provides more discrimination of outcomes in cognitive tests for PNI children. The framework provides a vehicle that addresses scientifically the concerns of parents and schools.

Acknowledgements

I would like to thank the PhD supervision team, Ray Frank, Dr. Farshid Amirabdollahian, Dr. Austen Rainer for the countless hours they have put in (including weekends) to help me with my work. In particular, I would like to thank Ray Frank for his unwavering support and efforts. I would like to wish him an enjoyable retirement as he deserves it after all his hard work for numerous years. I would also like to thank Dr. Rob Sharp in his role as consultant psychologist to the work we have been doing.

I would like to express my thanks for all the people in the University of Hertfordshire who have given me help and support including Noel Taylor and Catherine Khan from the School of Psychology, people I have worked with in the School of Engineering and various members of staff from the School of Computer Science for providing me with opportunities, guidance and advice.

I am grateful to the schools that agreed to participate in this study especially the staff of the special needs schools for their time and effort, and to the participants and their care-givers/families in assisting us with this study.

I am grateful to my family for their patience, understanding and support they have provided me throughout my work.

I would like to thank Ian Glasscock and BioDigitalHealth for the considerable support they have given to this research. They have contributed resources and equipment, finances, their time and their expertise. This research is supported by the Engineering and Physical Sciences Research Council (EPSRC UK) under iCASE 09001842.

Table of Contents

1. Introduction to thesis

1.1. Research question

Children with special educational needs are often administered cognitive tests to determine cognitive disabilities (Broomfield & Dodd, 2004). The tests commonly run on a computer to facilitate easy testing but require responses made by way of an input device. One group of children that are typically administered such tests are physically and neurologically impaired (PNI) children. These children typically suffer from a combination of cognitive, motor and speech disability. The severity of the impairment may make it difficult to use standard input devices like a mouse. Rapid technological advances are making devices with other modes of engagement which do not require abilities needed to operate a mouse.

This thesis is generally about improving access to a digital cognitive test through the use of devices that can make a difference. A general statement can be made regarding the problem:

“RQ1. What device helps a child successfully attempt a cognitive test?”

1.2. Building on previous work

Education of children with disabilities that affect communication such as cognition, motor and speech is difficult. Education, which is a two-way process of disseminating and assessing knowledge, is hampered if the communication pathways are restricted or degraded. Received knowledge is distorted and assessment is inaccurate. A one-time cognitive test was developed in conjunction with the University of Hertfordshire by Rob Sharp (Sharp & Evans, 1981) called Early Concept Development Test (ECDT) (Khan, 2009) specifically designed for such children. ECDT assess for learning and the transfer of learning. It was recognised that a number of the disabled children who were severely impaired was excluded from the test because of the inability to handle a mouse. The advent of new technology has meant that new bio-modal devices that do not depend on using the traditional motor skills are becoming available, and at a relatively low cost. These bio-modal devices provide the opportunity for those PNI children to undertake ECDT. Since the prospect of combining the bio-modal devices appear to have the potential to provide more features that may help overcome disability they were considered as part of this study. A number of issues emerge in undertaking this task as described next but first a model is required to facilitate discussion.

1.3. The problem

1.3.1. The user-device-test model

The cognitive testing of children requires a set-up that has a number of components. Of these components, three logical entities will be highlighted in this thesis as they are tightly-coupled. The entities interact with each other and a change in any of the three would have an impact on the results. The three entities hence form a triad and are composed of the user, the device and the test. A number of attributes of each component of the triad can be used to form independent variables in a study. In this study, as an initial attempt, variables consisting of the user ability, the device complexity and the test load were chosen as they represent basic factors of the problem. Using these factors as an initial attempt, we develop a deeper understanding of the problem and future work would be able to extend these factors to form more sophisticated studies.

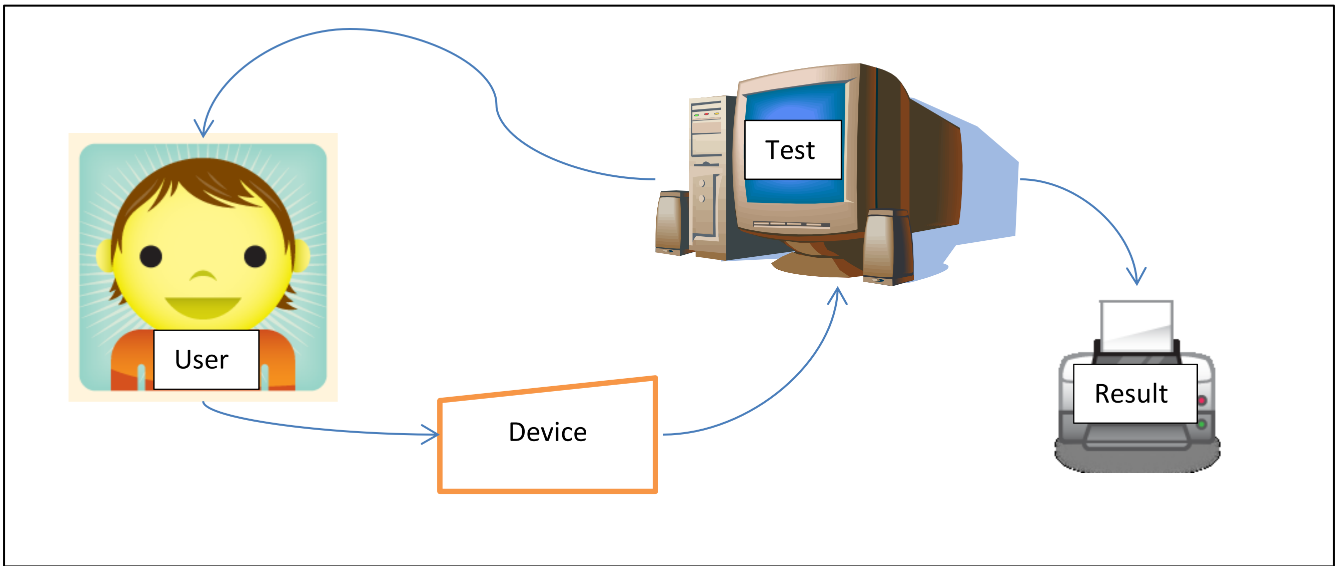


Figure 1 - User-device-test model

- RQ1 is a general question which is addressing

Figure 1 shows the relationships between user-device-test triad. The user is the participant of the test. The test which runs in a computer provides a stimulus presentation which is displayed on the screen which is perceived by the user. The output device(s) provides the test results. The user provides a response(s) to the stimulus presentation using the (input) device. The device translates the response into a format that is recognisable by the test. The test evaluates the device response and produces a result for each response. In ordinary circumstances, the device and the user are highly compatible.

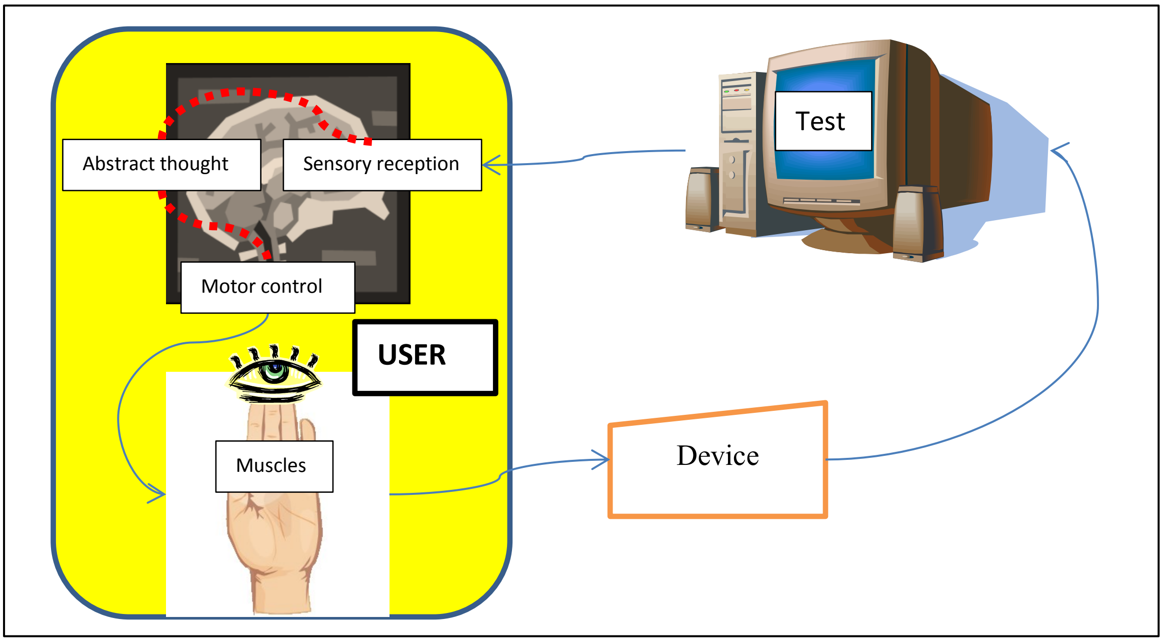


Figure 2 - User components in test interaction

Figure 2 shows a further decomposition of the user into the components that play dominant roles in the test interaction. The role of the brain is grouped into general functions that contribute to the test. The *sensory reception* function is a representation of our sense receptors where the five traditionally recognized ones are sight, hearing, taste, smell and touch (Penfield & Boldrey, 1937). The test displays images on the screen which are detected by the eyes and provides feedback in the form of visual and auditory inputs. The test stimuli are attended to using the *abstract thought* function of the brain (Grol, de Lange, Verstraten, Passingham, & Toni, 2006; Johnson et al., 2002) (Duncan & Owen, 2000; Nyhus & Barceló, 2009) which in this study includes the decision to provide a YES/NO response. The decision translates into a muscular action(s) which is brought into play by the *motor control* function of the brain (Debaere, Wenderoth, Sunaert, Van Hecke, & Swinnen, 2004). The brain sends impulses to the muscular nerve endings that causes muscle contraction and relaxation. The muscular action is generally focussed on providing a response to the test. In a simple example this would be a finger tap on a mouse button. Secondary responses may also ensue like a verbal remark. The muscular action of the user would be picked up by the bio-modal device and translated into a format which is recognized by the test running on a computer.

The breakdown above into components that take an active role in a test interaction enable a discussion of the problems associated with this study. Consideration is next given to the issues involved with each component of the user-device-test triad.

1.3.2. The user

This component of the user-device-test triad is a component that has specific requirements for the system. The target user in this study is not an ordinary user but a PNI child. Figure 2 shows the main components of the user that require consideration. The neurological impairment suffered by a PNI child would mean that all or a combination of the components identified in the figure may fail to a certain degree. A failure of the abstract thought component would be a target of the test (ECDT) but may also affect the motor control function. A failure of the relevant sensory reception function may affect participation in the test (ECDT) which must be taken into account by the test. A failure of the motor control function will prevent certain devices from being used. Impairments form a confounding factor to the undertaking of a test. In order to make a good assessment, we need to develop strategies that can effectively show how well a child has *successfully attempted* a test.

1.3.3. The device

A standard input device like a mouse (hand-held device) has been around for decades, and generations of users have grown accustomed to using the mouse and technical problems associated with the mouse have been resolved and improved upon giving an excellent combination of user and device. A proposal to replace a well-tried and tested technology such as the mouse requires the need to know the types of technologies available and whether the technology is applicable for a PNI child. A new technology will also need a period of adjustment from the user. Off-the-shelf technologies often imply that they can be used almost immediately (or else they remain permanently on-the-shelf) and be used for long periods. A large range of non-hand held device technologies are available. We need a process to determine *which of a range of devices* are appropriate for a child and be able to distinguish the effects of the device from cognitive ability.

1.3.4. The test

ECDT sets the requirements for the target test. In a normal circumstance, it is expected that the input device has no impact on the outcomes of the test. Unfortunately, a change to a new device and the fact that the target audience has specific weaknesses may invalidate those assumptions. The results obtained by the children who can now use a new range of devices may change due to issues with the device rather than the test. There is a need to be able to separate the issues of the test from the issues of the device. In order to do that *appropriate tests* need to be conceived to show device effects and cognitive effects.

1.3.5. Research question refinements

The issues discussed using the user-device-test triad introduces the following refinements to the research question and the major chapters in this thesis that answers the questions:

Number	Research question	Thesis chapter
RQ1	RQ1. What device helps a child successfully attempt a cognitive test?	10
	RQ1.1. Does a PNI child have sufficient motor control to use the device?	0
	RQ1.2. Does a PNI child have sufficient motor control to use the device and undertake a light cognitive test at the same time?	8
	RQ1.3. Does a PNI child have sufficient motor control to use the device and undertake a learning test at the same time?	9

1.4. Knowledge required

The area of work is a multi-disciplinary one and in general encompasses the following areas:

- Computer science
- Psychology
- Neuro-science

The computer science is focussed in the area of human systems interaction. This area is important because devices, tests and measures are not designed for PNI children and at the very least, some form of adaptation is required of the software. Psychology is required in the formulation of cognitive tests and understanding child development. Neuro-science is required when working with bio-modal devices such as Brain Computer Interfaces (BCI) and eye-tracking, understanding the impairments of the PNI children especially cerebral palsy and the effects of cognitive loading.

1.5. Contribution to knowledge

Before the computer revolution, psychological tests were commonly carried out using stimulus cards and motor or verbal responses. Present digital cognitive testing either assumes that input devices have no impact on the outcomes or include the ability to handle the devices as a part of passing the test. Although that is a fair assumption to make for the more well-developed participants, this thesis will show that when children of a developing age are considered, this is untrue. Probably more important is that if a participant has severe muscular impairments, it may not be possible for the child to achieve the criterion set by the test but still be able to satisfy the test intention.

In answering "RQ 1.1 Does a PNI child have sufficient motor control to use the device?" it was found that the rapid increase in technology has meant that new devices are being made available at a lower cost that uses non-hand held controls like eye movements. The devices are built with the ordinary population in mind and focus on PNI groups will not take top priority. This thesis provides a more sensitive evaluation of the child's ability that takes into account the impairments of the PNI group by using success streaks (consecutive successes) as a basis.

In answering "RQ 1.2 Does a PNI child have sufficient motor control to use the device and undertake a light cognitive test at the same time?" the effects of motor-skill in a test result were differentiated from the effects of the cognitive ability to pass the test and this thesis provides some basic procedures to separate those differences. The legacy learning test (ECDT) is a one-time test. In this work, a second test, CATTEST is used to provide an intermediate result which allows the tester the option of delaying ECDT until the PNI child is deemed ready for the more advanced test.

In answering "RQ 1.3 Does a PNI child have sufficient motor control to use the device and undertake a learning test at the same time?" the metrics used by the psychological learning test (ECDT) was reconciled with the streak measures (consecutive successes) used to provide more accurate responses from a PNI child. Thus the ECD Test was extended to be more than just a simple PASS/FAIL test and provides the vehicle for a more meaningful conversation between parents and educational psychologists as well as other professionals.

1.6.Ethical issues

Research aimed at PNI children have to take into account the lack of ability for most of these children to communicate. Hence signs of discomfort or willingness to participate are more difficult to interpret. In addition, children are used to specific protocols with adults that attribute an adult with a a greater level of authority and power. Some children tend to want to please adults. Care needs to be exercised so that the researcher become sensitive to the needs of the child.

As the child is in a lot of cases unable to understand and provide informed consent to the research, pains were taken to ensure that the circle of care-givers that the child has were fully informed. This included the parents, the teachers and special assistants that look after the child.

Consent was obtained from the parents who understood that it was possible to terminate the studies at any point in time.

Confidentiality remains a primary concern for the children as they are not aware of the hazards of personal data that can be misused. All names were therefore protected and held securely, including the names of the schools that the children attend. All discussions of the study used fictitious names. Photographs and video that contained faces of the children were not obtained even though they contained detail that may help with the analysis of the data. Such details may provide an indication of the fatigue experienced by the child during the study as well as the motivational state of the child.

1.7.iCASE Work

This PhD work was funded by an ICASE (Industrial Cooperative Awards in Science & Technology (iCASE) award from the Engineering and Physical Sciences Research Council (EPSRC). A condition for this award was that 6 months effort with the industrial partner was required. In respect of this, 3 months was spent with the Assisted Living project which was Knowledge for Business (K4B) project organised by the School of Engineering in the University of Hertfordshire. In those 3 months, control of house-hold appliances and lighting using an EEG headset and eye-tracker was integrated into the Assisted Living project. In the other 3 months of the iCASE requirement, the technical support environment for this PhD study was repackaged into a single package as a prototype for commercial delivery. The package also included a proof of concept of transferring the study data to a web server for storage by providing a secure registration function for a user of the system.

The internet version of the test introduces the possibility of a larger number of children taking the test, increasing cases available for further study although at a price of strict experimental conditions not being met.

1.8.Publications

A list of publications follow which was published with the thesis author as the main contributor:

1. Gan, H., Frank, R., Amirabdollahian, F., Sharp, R., & Rainer, A. (2014b). Development of the maximum-streak measure for evaluating the suitability of non-handheld devices in cognitive tests of Physically and Neurologically Impaired(PNI) children. *International Journal of Advances in Computer Science & Its Applications, 4*(4), 130 - 136.
2. Gan, H., Frank, R., Amirabdollahian, F., Sharp, R., & Rainer, A. (2014a). Bio-digital device impact on a constant load cognitive test of children with physical and neurological impairments. *International Journal of Advances in Computer Science & Its Applications, 4*(4), 99 - 105.
3. Gan, H., Frank, R., Amirabdollahian, F., Sharp, R., & Rainer, A. (2014c). Use of re-attempts measure for evaluating device test results of children with neurological impairments. *Human System Interactions (HSI), 2014 7th International Conference on*, 206 - 211. doi: 10.1109/HSI.2014.6860476

Paper 1 describes the development of a measure suited for PNI children as usual measures do not allow for the involuntary responses made as a result of impairments that affect the children. This paper establishes using the maximum consecutive success (maximum-streaks) as a measure and is used in the three main tests of this thesis to interpret results as described in chapters 0, 8, 9.

Paper 2 is a sub-set of chapter 8. It describes the results of a light cognitive test on PNI children and indicates that the results suggest that a PNI child that is cognitively capable of a cognitive test may have test results that are negative because of the input device used.

Paper 3 extends Paper 1 in considering a different measure based on consecutive successes. Just as with normal measures, we can choose to view an instantaneous value or an averaged value, it is possible to establish similar parallels with streak measures. Paper 3 describes the development of a streak-sum measure which provides an averaged view of streaks.

1.9.Thesis structure

In this introduction, a flavour is given of the problem and the problem model of a user-device-test triad. The user-device-test triad model will be re-used throughout the study including presentation of topics. This thesis next provides a literature review to discover the state of the art in the area in Chapter 2. A theoretical model of device handling capability is also presented to seed the basis for a device fit of PNI children.

After the literature review, the development of the operational model is presented in the order of initial exploration (Chapter 3), system implementation (Chapter 4) and pilot study (Chapter 5). The user-device-test triad is again employed in the presentation where:

- The initial exploration is one where *devices* are first explored for their capabilities
- The system implementation follows where the *tests* are set up and built to support the operational model of the study
- The pilot study verifies the system implementation with target *users* (PNI children)

After the initial set-up work has been described, the thesis continues with the description of the main study. The main study consists of a test suite of 3 tests and an overview of these tests is given in Chapter 6. Briefly here, the three tests are:

- COMP TEST (Chapter 0).

  This is a competence test with the aim of establishing simple motor-skills with the device.
- CAT TEST (Chapter 8).

  This is a cognitive test based on taxonomic categorization to establish capability with a novel device under a light cognitive load.
- ECD Test (Chapter 9).

  This is the existing learning test, now extended to include the novel devices instead of just using a mouse

A chapter follows that discusses the results across the 3 tests of the main study and re-visits the theoretical model (Chapter 10)

The thesis finishes with the conclusion in Chapter 11 where future work following the study is discussed.

Section 1.3.5 provides a road map to the answers of research questions in the thesis.

2. Literature review

2.1. Aim

This section performs the literature review to answer the research questions outlined in the previous chapter (section 1.3.5) in terms of the user-device-test model (section 1.3.1):

Number	Research question
RQ1	RQ1. What device helps a child successfully attempt a cognitive test?
	RQ1.1. Does a PNI child have sufficient motor control to use the device?
	RQ1.2. Does a PNI child have sufficient motor control to use the device and undertake a light cognitive test at the same time?
	RQ1.3. Does a PNI child have sufficient motor control to use the device and undertake a learning test at the same time?

2.2. User

2.2.1. Children with Physical and Neurological Impairment (PNI)

The World Health Organisation (WHO) (Organization, 1980) defines impairment as "any loss or abnormality of psychological, physiological, or anatomical structure or function". Children with PNI have loss of brain function which results generally in the loss of certain motor functions but other functions may also be impacted. The majority of PNI children suffer from Cerebral Palsy (CP). "CP constitutes 67% of the severe motor disabilities in childhood"(Cans et al., 2004). An understanding of CP is helpful as the problems have common ground among PNI children and it is more widely documented. Also, the majority of children in this study have CP. CP is defined as "a group of permanent disorders of the development of movement and posture, causing activity limitations that are attributed to non-progressive disturbances that occurred in the developing fetal or infant brain. The motor disorders of CP are often accompanied by disturbances of sensation, perception, cognition, communication, and behaviour, by epilepsy, and by secondary musculoskeletal problems". Effects on movement include coordinated eye movements and articulation of speech. Disturbances on sensation would include vision and hearing as primary disturbances lead to restrictions in learning. Cognition includes both global and specific cognitive processes including attention as a primary disturbance again leading to restrictions on learning. Disturbances in behaviour include psychiatric problems such as autistic spectrum disorders and ADHD (Attention deficit hyperactivity disorder). Secondary musculoskeletal problems refer to further complications developed as a result of the brain injury for example muscle/tendon contractures (tightening) (P. Rosenbaum et al., 2007). From the definition we see that the range of abilities and sensory modalities that are impacted are extensive and has huge variation in a CP population. There is also a recurring theme of initial primary disorders that lead to secondary problems.

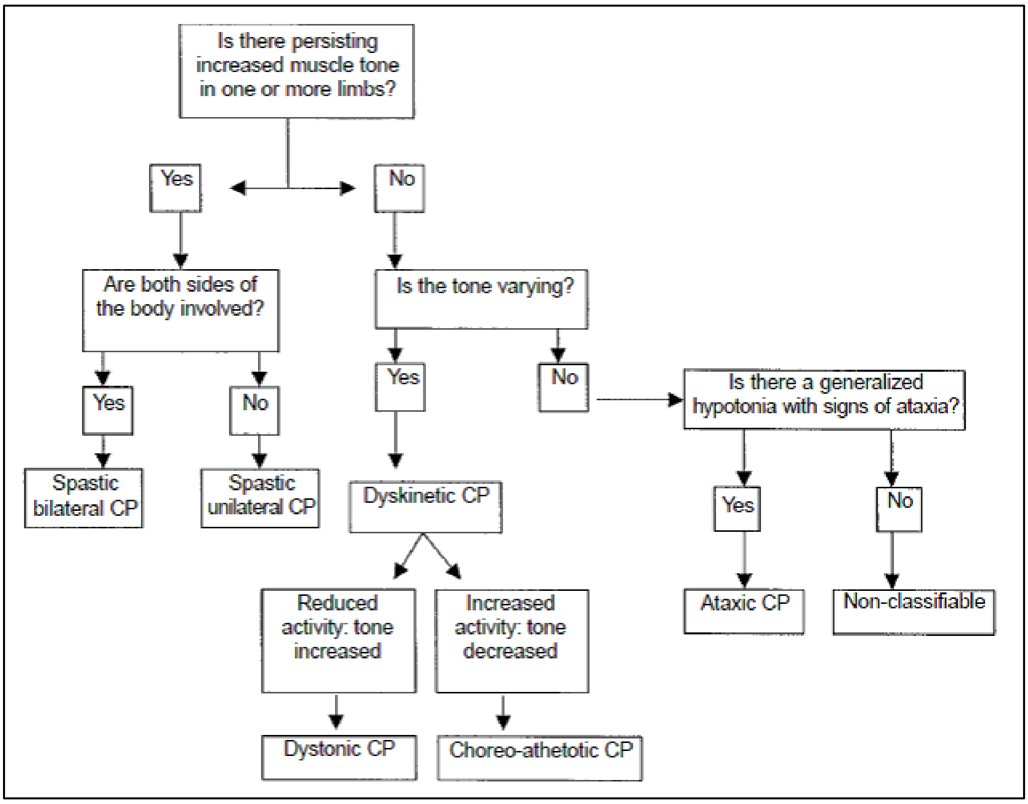


Figure 3 - Hierarchical classification tree of cerebral palsy sub-types (Cans, 2004)

Figure 3 shows a classification of CP based on motor abnormalities adopted by SCPE (Surveillance of CP in Europe), an EU funded project. The figure shows a varying degree of muscle tone control that is impacted by different manifestations of CP giving an idea of the wide variation that can be found in terms of motor control from CP children. Spasticity, involuntary movements and problems with posture are common symptoms of CP.

Current imaging techniques allow some abnormalities to be viewed. Bax carried out a study which correlated Magnetic Resonance Imaging (MRI) scans with clinical diagnoses of 351 CP children (Bax, Tydeman, & Flodmark, 2006). The scans show damage in both connective tissue (white matter in the brain) as well as damage to other parts of the brain. White matter damage of immaturity (WMDI) accounted for the major proportion of the scans (42.5%). The WMDI MRIs included damage to the periventricular regions which are heavily involved with motor control. Correlates of WMDI with clinical diagnoses ranged from a proportion of those with diplegia (71.3%), hemiplegia (34.1%) and quadriplegia (35.1%). Diplegia refers to impairments on both sides of the body (usually lower extremity), hemiplegia refers to impairments on one side of the body and quadriplegia refers to impairments on all four limbs.

Bax's participants were assessed clinically for signs of CP. Some variables of the clinical evaluation are presented in the table (Table 1). Impairments that impact device operation are abnormal hand function, involuntary movements and visual abnormalities. Abnormal hand function ranged from 54% to 87%, involuntary movements ranged from 11% to 36% and visual abnormalities ranged from 31% to 64%. Assessments of ability do not normally focus on ability to use input devices for computers.

Table 1 - Selected results from physical examination of CP children (Bax et al., 2006). The figures in brackets provide the results as a percentage of n children having a particular CP type.

Selected Results From Physical Examination			
Impairment	Type of Cerebral Palsy, No. (%)		
	Hemiplegia	Diplegia	Quadriplegia
	(n = 113)	(n = 148)	(n = 80)
Walking functional	100 (89%)	93 (63%)	7 (9%)
Sitting functional	113 (100%)	125 (85%)	28 (35%)
Abnormal hand function	98 (87%)	80 (54%)	58 (73%)
Functional severity *			
*Mild	70 (63%)	48 (34%)	6 (7%)
*Moderate	38 (34%)	66 (46%)	11 (14%)
*Severe	3 (3%)	29 (20%)	63 (79%)
Involuntary movements	12 (11%)	26 (18%)	29 (36%)
Scoliosis	2 (2%)	4 (3%)	18 (23%)
Muscle tone	96 (85%) [Mixed]	78 (53%) [Increased]	68 (85%) [Increased]
Vision	35 (31%)	56 (38%)	51 (64%)
Epilepsy	29 (26%)	21 (16%)	39 (50%)
Communication problems	44 (39%)	56 (39%)	71 (89%)

Other studies of CP like Jacobson (Jacobson, Flodmark, & Martin, 2006) reports visual field defects associated with WMDI in addition to motor impairments and introduces the observation that improved perinatal and neonatal (just before and after birth) care results in a growing population which have complex cerebral dysfunctions.

Although evidence such as these studies implies that damage to particular areas of the brain lead to the loss of particular abilities, the detailed picture is more complicated. The brain is a dynamic organ capable of reorganization even as we age (Mahncke et al., 2006). There is a chance that the brain caters for injury by using areas of the brain that are undamaged. This ability for cortical reordering is an example of a property of the brain referred to as neuroplasticity. Staudt (Staudt, 2010) shows that the developing human brain is able to reorganize itself for certain functions after pre and perinatal brain lesions. Due to the developmental process that occurs typically during the infant stages of growth, the inference was that such reorganization was more effective in infants than the adult brain.

The wide variety of impairments, secondary impacts of the impairments, and a brain that is highly adaptable, suggests that latent abilities and exhibited abilities of PNI children will vary greatly and each individual will require very specific evaluation to uncover true abilities. These abilities are important for a PNI child who is required to learn to use new devices to perform cognitively demanding tasks.

The literature suggests that in the face of impairment, the brain is versatile and evolved to a structure that avoids catastrophic failures. The question is "how much is the brain capable of compensating?". The research questions are structured to separate device skills from cognitive skills. A comparison is done with non-PNI children to see if there are specific deficiencies in performance. The question "RQ1.1 Does a PNI child have sufficient motor control to use the device?" looks for sufficient motor control abilities with a device. The question ". The question "RQ1.2 Does a PNI child have sufficient motor control to use the device and undertake a light cognitive test at the same time?" looks for sufficient muti-tasking abilities to handle both device and a light cognitive function. The question "RQ1.3 Does a PNI child have sufficient motor control to use the device and undertake a learning test at the same time?" looks for sufficient muti-tasking abilities to handle both device and a more complex cognitive function.

2.3.Device

Physically and Neurologically Impaired (PNI) children have special needs due to problems which result from brain injury. The effects of the brain impairment result in physical impairment and the need to address both problems. The involuntary actuation and inhibition of movements in PNI children mean that such children have great difficulty controlling the traditional devices used in tests, the typical traditional device being a computer mouse. New bio-modal devices that do not depend on using the traditional motor skills are becoming available, and at a relatively low cost. These bio-modal devices provide new opportunities for CP children to undertake a test, but may also require that the children develop new skills in order to control those new devices.

Devices that could help with impairments fall into two categories:

- Devices that can ignore the impairments
- Devices that make use of unimpaired functions in the body

The devices that are investigated for this study fall into the latter category. In addition, a restricted range of technologies were chosen that represent a wide enough range of alternative modalities for a PNI child to use. These technologies are subsequently described.

2.3.1. Electroencephalogram (EEG) neuro-headset

EEG signals are variations in an electric field that is generated by the brain due to processing activity that can be measured at the surface of the scalp (Pizzagali, 2007). The technology now exists that can translate EEG into commands that control electronic devices. This area of work has its beginnings around the 1970's (Vidal, 1973) under the brain-computer interface (BCI) (J. R. Wolpaw, McFarland, & Vaughan, 2000) banner. The availability of this technology means that there is a possibility for people who have disability problems to control devices (such as a mouse) that they previously would not be able to as it would have required some form of physical motor ability. The possibilities of using this technology have opened up further with the commercialization of BCI devices for the computer games market. This has meant that relatively cheap BCI devices can be used for development. These devices provide capabilities for a software developer to integrate their written software with the BCI device. Software can now be built that uses the underlying functions via an Application Programming Interface(API) provided by the BCI device manufacturer that either perform some translation on the EEG or provide the raw EEG. One such device is a neuro-headset package called the EPOC™ that is produced by Emotiv™ (Allsop et al., 2006; Emotiv, 2013).

The use of EEG falls into two categories. One involves the detection of electric field changes from the brain and the other involves detection of field changes from the head (but not the brain). EEG waveforms that are not originated by the brain but by other mechanisms are also known as EEG artefacts (Dworetzky, Herman, & Tatum IV, 2011). A purist BCI researcher only considers using EEG from the brain and not the artefacts. At this stage of the research however, artefacts represent the most reliable mechanism for exerting control via EEG. The artefacts have a magnitude of around 10 times that of the cognitive signal (200 microV vs. 20 microV). The artefacts tend to mask the brain waves and some processing is required to remove the artefacts if the interest is to use the brain waves.

The use of brain waves involves a choice of specific mechanisms to invoke the production of these waves. Three popular EEG brain signatures are P300 evoked potential, slow cortical potentials and sensorimotor rhythms (J. Wolpaw & Birbaumer, 2006). A simple procedure is the use of sensorimotor rhythms.

The use of sensorimotor rhythms has its history with Gastaut (Gastaut, 1952) finding that motor action and imagined motor action generates an attenuation of rhythmic EEG waves (called mu-rhythms) within the 8 – 12 Hz frequency range around the sensorimotor region of the head. Other rhythms also accompany the mu-rhythms and in general, there is a marked attenuation during movement and amplification when control stops (Pfurtscheller & Lopes da Silva, 1999). This is known as desynchronisation and synchronisation respectively in reference to the activity of groups of neurons whose numbers contribute to the amplitude of the rhythm. Activity tends to split a group of neurons that are synchronously idling to little groups of activity that decreases the idling population. Apart from motor action and imagined motor action, the brain also generates mu-rhythms in response to observations of motor actions performed by others triggered by a type of motor neuron known as mirror neurons (Muthukumaraswamy, Johnson, & McNair, 2004). Among BCI practitioners, the use of imagined motor action is common as research is often done for patients who have very little muscular control such as motor-neurone disease.

EEG artefacts can be generated using a wide variety of mechanisms (Reyes et al., 2012). The two common artefacts are due to eye movements and muscular movements (Nicolas-Alonso &

Gomez-Gil, 2012). More exotic forms of artefact generation such as the vestibulo-ocular (VOR) reflex involves eye movements in order to stabilise an image when the head is moving and the glosso-kinetic potentials caused by tongue movement.

The Emotiv™ neuro-headset provides the facility to interact with the machine without needing to know the underlying mechanisms. The EEG headset (EPOC™) used required an 8 s training period for the machine to learn the different mental states that are used as commands to the computer. There were no procedures specific to the training (which will be referred to as the training protocol) apart from the fact that a neutral state had to be trained and a specific mental state. This mental state would be to imagine an action performed on a displayed cube which causes the cube to exhibit some form of motion like LIFT, PUSH or RIGHT. Patents filed by Emotiv (Allsop et al., 2006) describe the use of a classifier to match mental states.

An assumption was made that it would be possible to train the generic classifier defined by Emotiv's database with signals from the user to recognise either simple motor or imagined motor actions or artifacts generated by muscular action using bites (Reyes et al., 2012).

2.3.2. Eye-tracker

The eye-tracker is a device that tracks the gaze of the eye and maps the position with a point on a computer screen. Infra-red illumination is applied to the eye which results in reflections from the cornea. The reflections are used to provide a stable reference for the eye-position. A camera is used to track the reflections and detect the position of the pupil (Duchowski, 2007). Movement of the pupil position can be used to control the location of a mouse cursor enabling a pointing function. The eye-tracker is an alternative to the mouse but requires good control of both eyes. PNI children who have impairments of the eye will not be able to use this device. However, Birbaumer (Birbaumer, 2006) indicates that eye control and external sphincter muscle control are usually the two last areas of control lost for motor-neurone patients. Thus, eye control is a fundamental control and may have uses for many groups of participants.

Two versions of eye-trackers exist. One is head mounted on a spectacle frame (ASL, 2014) and the other is remotely mounted (San Agustin et al., 2010). Remote mounting places the eye-tracker either above or below the computer screen. The head mounted version is anticipated to be a problem for many PNI children as they have involuntary movements of the head which will knock the eye-tracker. The remote eye-tracker is not intrusive and therefore is the better version to use.

The use of the eye for controlling a cursor presents several problems. The eye does not stay still when it fixates on an object. Slight deviations are observed called micro-saccades (Martinez-Conde, Macknik, Troncoso, & Hubel, 2009) (of around 10ms in duration) in addition to noise-like deviations (Duchowski, 2007). The microsaccades and noise fluctuations require smoothing before the device can be practically used. It remains for an exploratory investigation to be done to determine how much filtering has to be done to correct ocular fluctuations. In addition, Oviatt (Oviatt, 1997) indicates that a common problem using the eye for Human Control Interaction is the Midas effect (R. J. Jacob, 1991; R. J. Jacob & Karn, 2003; Majaranta & Räihä, 2002). This effect is a result of normal use of the eyes which are constantly scanning for objects of interest. A functional overlap occurs between the use of the eyes as a mouse cursor for pointing and selecting and the natural use of the eyes for vision resulting in selection errors.

Ware (Ware & Mikaelian, 1986) compares the use of the eye-tracker for selection using a “dwell-time”, with a button press, and an on-screen virtual button. Results using the eye-tracker were found to be favourable provided the button size was not too small. Although the results also show better performance for the use of selection using a direct fixation of the stimulus, in terms of setting up experimental conditions between different tests, there is a need for consistency of the non-dependent variables. In this case, the on-screen virtual button provides a way to provide for consistency in the different tests. Inconsistency in tests can account for interaction effects of independent variables. The interaction effect in this case would lie between the device and the test. If the effort required for providing the response to a test depends on the position of the stimulus, this effort will vary from test to test as the stimulus changes resulting in an inconsistent effort to manipulate the eye-tracker.

2.3.3. Head-tracker

The head-tracker is a device that tracks the movement of a nominated spot (which is translated into a square area) on the face. A video camera is used to track the motion of the spot which will be translated to the movement of a mouse cursor on a computer screen. The spot becomes a reference point for tracking. One such head-tracking system is the Camera Mouse (Betke, Gips, & Fleming, 2002) which can track body features such as the tip of a user's nose or finger. Software takes the sub image of the square area defined to match with subsequent images that are captured by the video camera. The closest match to the search is found and used as the new reference point for tracking. This head-tracker was found to compare favourably in a study by Man (Man & Wong, 2007) on adolescents with cerebral palsy.

2.3.4. Multi-modal fusion

The implementation of a multi-modal device has similarities with sensor systems set up for military applications where an array of sensor systems is set up in the combat zone. Architectures are set up to utilise data from a variety of data sources to obtain the property of an environment or object and the term data fusion has been used to describe it (Soleymani, Lichtenauer, Pun, & Pantic, 2012). Multi-modal device implementation aims to derive human intention from the data sources. Data fusion has the objective of improving the quality of the information from the different data sources and combining the data into a common representational format (H. B. Mitchell, 2007). In terms of multi-modal devices, the information derived from the combination of single mode sensors must be of a higher quality than the information obtained from each of the sensors. Another area of study which provides similar conclusions and further insight is in the area of Human Computer Interactions (HCI). Oviatt refers to multi-modal systems that integrate signals at the feature level (early fusion) as opposed to those that integrate systems at the semantic level (late fusion).

Furthermore, as a guideline, feature fusion is considered more appropriate for closely synchronised modalities (Oviatt, 1997). In the case of this study, the multi-modal devices are not closely synchronised and there is a significant amount of work to do that. Late fusion appears to be a simpler step but nevertheless is a step to learn from. At this stage another area of work provides further insight. A study was done in the area of Brain Computer Interfacing (BCI) that implemented a hybrid system using EEG produced from imagined motor movements and eye-tracking (Zander, Kothe, Jatzev, & Gaertner, 2010). This suggests the possibility of fusing the EEG headset with an eye-tracker.

2.3.5. Signing

This thesis also considers Signing as a method of input. Signing provides a contrast to bio-modal inputs and the typical physical inputs. For our research, Signing refers to a child who communicates using gestures or speech to an interpreter. The gesture and or speech acts as the child's response to a test, and this response is then entered as a mouse input via the interpreter. The interpreter may also be relaying the response via speech to a researcher.

The interpreter used is a carer who was able to pick up on subtleties of a communication with a PNI child. Such subtleties include gestures that are additionally used that are not a part of the recognised repertoire established for that child. Indications are both sent and received by the care-giver and PNI child without conscious thought which contribute to the communication (Mehrabian, 1977). In the case of a child with attention problems or less severely, moments of distraction or lack of motivation, the one-to-one communication acts as a compelled request-response protocol which demands the attention of the child.

Signing was deemed to impose the least cognitive load for manipulating a device. The child may have had years of using and developing it to use the parts of the body which were functionally capable for the purpose. Signing would be used from a very early age for communication and the participant would have had a lot of practice. The skill should have developed beyond the stage where the Signs would not require any cognitive association as indicated by Ericsson (Ericsson, Krampe, & Tesch-Römer, 1993). Once that level of skill is reached, the brain is freed up to perform other cognitively demanding tasks (Smith & Chamberlin, 1992).

As such, Signing was viewed to be one of the best (if not the best) practical devices available albeit a conceptual one.

Signing is a valuable contrasting 'input' because a child that successfully completes a test using the Signing 'device' most likely has the cognitive ability for being tested. Where that child then has difficulties completing the test successfully with a bio-modal device helps to expose the positive or negative impact of that device.

The replacement of Signing using a physical device loses some of those advantages but also associated disadvantages. Carers develop strong ties with the participant and subconsciously may provide responses that bias the response towards a more positive outcome. The subtleties of Signing would be viewed as an opportunity for further research. Another disadvantage of Signing is that appropriate carers are expensive. Signing would have been used before the tests were converted to a digital format. In progressing from a manual procedure to a more automated one, certain advantages may have been lost and others gained.

The view of Signing as a conceptual device allows for the demonstration that a physical device exists that may be more efficient and a better fit for a PNI child. The physical device implemented will depend on the communication mode used by the child. The more able-bodied children may show a preference for using voice or touch-screen devices and the less able-bodied may prefer gestures that do not require hand control (Biswas & Basu, 2011; Ren, Meng, Yuan, & Zhang, 2011).

Signing can either employ selection or pointing and selection. Normally there is an independent gesture for YES and NO but the participant may choose to point to the screen.

2.3.6. Initial theoretical model for device skills

The best fit of devices was assumed to require the least ability. We know from a body of work on attention, multi-tasking, short-term memory and chunking (Engle, 2002; Shiffrin & Nosofsky, 1994) that the human brain can only handle a limited number (around 5 to 7)(Miller, 1956) of concurrent points of attention or chunks. It is a very primitive way of expressing concurrent handling by the brain but provides a general idea. We relate the complexity of a task to a number of points of attention that has to be undertaken. We assume that the greater the cognitive complexity of the task the more ability is required.

From the above, we ranked the devices tested by the study into categories in order of complexity. We find that the devices fall into two parallel categories. One category consists of the number of bio-modal modes employed by the device. The other category consists of the number of operations required to use the device. An increase in complexity is associated with an increase in the number of bio-modal modes. An increase in cognitive complexity is also associated with the number of simultaneous operations that the brain has to deal with. Each mode or operation is an abstract representation the brain has to make. An alignment or calibration of representations takes place (Andersen, Glassman, Chen, & Cole, 1995; Fischer, 1980; van Ee, van Boxtel, Parker, & Alais, 2009). Representations of the physical world have to be associated with representations of the self. An allocentric view has to be coordinated with an egocentric view(Burgess, 2006). When we use a head-tracker or eye-tracker, we need to map between a visual and a motor space and we learn re-mappings in a piecemeal fashion rather than a general formula (Ghahramani, Wolpert, & Jordan, 1996). Sensorimotor integration is required (Ghahramani, Wolpert, & Jordan, 1995). An increase in modes or operations therefore increases the number of abstract representations that has to be made. An increase in the simultaneous number of points of attention is required as pointing has to be used to steady a cursor in a particular position and when the time is appropriate, selection has to be simultaneously carried out.

Table 2 ranks the devices based on their expected complexity, with the simplest device at the top and the most complex device at the bottom. Single mode devices are placed first since they have a lower number of modes to handle than the hybrid devices. Devices that in addition only require selection are placed at the top of the table since they require less manipulation than both pointing and selection. An ambiguity exists due to devices having both the same mode and requiring the same operations and this is resolved by assuming them to be equal in complexity. It was recognised however that further information such as latent ability (experience) with the device would make the device less complex. Signing therefore was placed at the top of the table, followed by the mouse as they would be the most easy to use device to most children.

Table 2 – Devices ranked in order of complexity of use starting with the least complex devices at the top of the table and the more more complex at the bottom. Intuitively single-mode devices are expected to be less complex than multi-modal devices. Intuitively a device that requires a selection operation is expected to be less complex than a device with both pointing and selection.

Device	Mode	Operations
Signing	Single-mode	Selection
Mouse[1] or switch		
EEG headset		
Head-tracker		Pointing and selection
Eye-tracker		
EEG headset + head-tracker	Multi-mode	
EEG headset + eye-tracker		

All other devices would be novel to most children and were regarded as equally difficult to use. Latent ability provided an indication that this initial theoretical model probably required to consider more factors. Having ranked the devices, the expectation was that results obtained from the empirical study would mostly match the theoretical model.

[1] Although the mouse can be used for pointing, this is not used in the study.

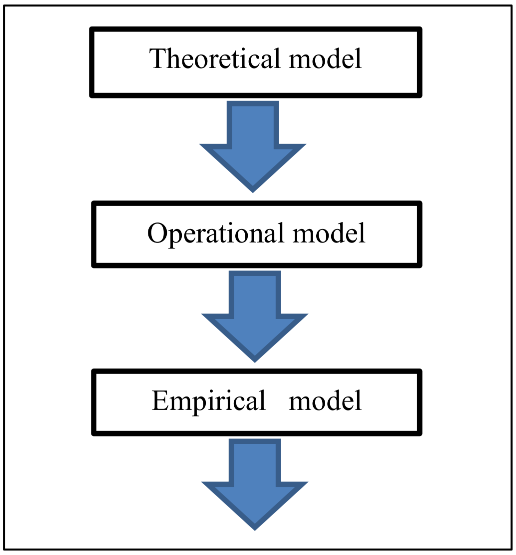


Figure 4 - Development of model for fitting device skills based on cognitive complexity required for handling device

An operational model had to be implemented to support experimental verification of the theoretical model as indicated in Chapter 5. As indicated on Figure 4, the theoretical model leads to the development of the operational model. The results obtained by the operational model produces a set of results that form the empirical model.

On the other hand, the empirical study could show short-comings with the theoretical model and a post study analysis (Chapter 10) is needed.

A number of non-hand held device types were presented in the literature and the initial theoretical model leads to the research question "RQ 1.1 Does a PNI child have sufficient motor control to use the device?" Different types of devices have to be trialled against different combinations of impairments to ascertain device suitability. In particular a comparison needs to be carried out to match performances of multi-modal devices with respect to their single modes.

2.4. Test

The introduction of novel devices to the legacy test (ECDT) requires the use of the device to have no significant impact on the test. In order to verify this situation a test additional to the legacy test such as a motor skills test is required.

2.4.1. Motor-skills test

A motor skills test checks for a user's ability with a device with an implicit aim of fitting devices to users. A previous project with a similar perspective to fitting devices to users was the "i-match" project which focused on comparing input devices to identify a best fit to individual's abilities (Amirabdollahian, Gomes, & Johnson, 2005; Amirabdollahian, Munih, et al., 2005; Bowler, Amirabdollahian, & Dautenhahn, 2011). The general principles that were extracted from that project were the motor skills assessment usually based on the use of Fitts' test (Fitts, 1954) and a multitude of benchmarks. A motor skills assessment was made in the case of the "i-match" project to evaluate the effectiveness of a number of set configurations for a single device for a sample population. In the case of this study, general configurations were found and the motor skills assessment made to determine the suitability of a number of set devices for a single individual.

Motor skill tests do not, in general, need to consider the failure to make a correct response. Devices used for such tests are often stable and tests are simple reaction tests not expected to have a high cognitive load. Participants chosen have enough developmental maturity to carry out the tests perfectly. This difference was noted by Donker who carried studies of pointing skills of children with a computer mouse (Donker & Reitsma, 2007b) who indicated "Fitts's Law only applies to error-free behaviour and may as a result not be suitable to describe the aiming behaviour of young children" (Donker & Reitsma, 2007a). The parameter of interest moves on to looking at accuracy, performance times or speed, and statistical results are generated by looking at a large enough sample (Amirabdollahian, Gomes, et al., 2005; Man & Wong, 2007). The expectation is that PNI children would make even more errors than non-PNI children because of impairments. From this point of view the impairments are unintentional inputs. The initial problem would appear to be trying to establish intention then performance. A technique to achieve this using success streaks (consecutive successes) is discussed in a subsequent section (2.4.4).

The motor-skills test in the literature indicate the need for a test with the lowest possible cognitive effort to determine the usability of devices. Fitts' law indicates that difficulty depends on size of object as well as distance to target. For the PNI children, with cognitive and sometimes visual impairments, the size of the object has to be large and the properties of the object must be simple. The research question "RQ 1.1 Does a PNI child have sufficient motor control to use the device?", depend on a test that is appropriate for the PNI child although not explicitly stated.

2.4.2. Light cognitive (Categorization) test

A motor-skills test differ from an application in that a cognitive load that is higher than that used in the motor-skills test is imposed upon the participant at the same time the participant is using a device. A further test to the motor skills test is thus useful. A Categorization test (CATTEST) appears applicable as it is a test that provides more cognitive complexity than a motor skills test, increasing the cognitive component required to pass the test and is suitable for children. According to Piaget and his 4 stages of cognitive development, classification is a feature that occurs consistently in the Pre-operational stage (ages 2 to 7) within limits (Sroufe, Cooper, DeHart, Marshall, & Bronfenbrenner, 1992). Categorization tests based on sorting cards with simple shapes and colours exist (Fisher, 2011) but a pre-condition for a later test (ECDT) that will be run (described in the next section - 2.4.3) was that pre-test training of ECDT stimuli must be avoided. In order to fulfil this requirement a categorization test that involved recognising birds and fruits and being able to work out that the stimulus consists of all birds or fruits or not can be used. This test would be within the capability of a 3 year old (Dunham & Dunham, 1995). The cognitive component that was increased in this case had no particular focus on device operation although there may be parts of that component that is also required for device operation. CATTEST can be applied on a child using his best devices. A change in results would therefore be mainly due to the increase in the cognitive component. CATTEST would not be used to test cognitive limits but to introduce a general load and therefore a ranking of cognitive ability would not be necessary.

A problem found in the previous section on motor-skills (2.4.1), the problem of discovering intentional responses would be a common problem to this test. The relevance of detecting an intentional response increases in this case as a higher cognitive function is required. The use of success streaks (consecutive successes) discussed in section (0) for the motor-skills test and in particular maximum-streaks appear to fit in this case as well. The maximum-streak as a measure can be viewed as a "ceiling" where the best effort that can be made in a test run is achieved. The test run is designed so that it is of sufficient length where the ceiling can be reached for some children. A target is set where achieving a maximum-streak of a particular threshold, say at 20, ends the test and if a child attains this, their ceiling can be higher. Introducing an additional load to a test is expected to lower the ceiling for a child. The maximum-streak measure can thus be viewed as a measure of a child's capacity to handle the demands of both device and cognitive load.

In a system which is divided into a user, device, test triad, all three components influence the generation of an outcome. The user who is influenced by their impairments (which diminish the capacity to handle a cognitive load and a device load), the input device (which generates a device specific cognitive load and modifies user responses due to its imperfections) and the test (which generates a specific cognitive load in the form of stimuli and evaluates device responses depending on design) all play a role in the final outcome which is generally regarded as a response from the user to a test. The use of two tests (the motor skills test and this test) helps to separate the cognitive demands of the device from the cognitive demands of the test. Setting the cognitive load to a specific level allows for the demands of the device to be estimated and conversely, setting the device load at a specific level allows for the demands of the cognitive load to be estimated. The motor-skills test does the former and both the motor skills test and CATTEST does the latter.

The categorisations test in the literature indicate that as a simple cognitive test, the test should be appropriate for non-PNI children around the ages of 4 to 6. The research question "RQ1.2 Does a PNI child have sufficient motor control to use the device and undertake a light cognitive test at the same time?" uses a taxonomic categorisation test as part of a progressive sequence of tests of increasing complexity which adds to the cognitive load while multi-tasking with novel device handling.

2.4.3. Psychological test (ECDT)

The Early Concept Development Test (ECDT) as briefly discussed in the introduction (section 2.4.3) is a legacy test inherited by this study. ECDT is a cognitive test that checks for learning and the transfer of learning. ECDT starts with a sequence called Initial Learning which provides card images of simple shapes and colours (referred to as dimensions) and conditions the child to learn a specific set of a dimension (e.g. the colour yellow). The child is not cued about the learning process and is told to discover the solution to the test themselves. Feedback is provided during the test to indicate whether the response of the child is correct or incorrect. Once a child demonstrates that they have learnt by getting the correct solution 10 consecutive times, another sequence is presented to check that the child is able to transfer learning to a new set in the same dimension or to a different dimension. For the purpose of this study, a different set in the same dimension (referred to as an intra-dimensional shift) is always used. The test is similar in many ways to a cognitive test used for checking flexibility when rules of enforcement are applied called Wisconsin Card Sorting Test (WCST) (D. A. B. Grant, Esta A., 1981, 2003). Due to the strong similarity and a large amount of literature (Nyhus & Barceló, 2009) devoted to WCST, the metrics used by that test can be re-used in this study. ECDT can therefore be extended to provide more indications than just a PASS/FAIL test.

In order to distinguish device impacts on the results of the test (ECDT), the results can be interpreted in two ways; a low-level interpretation which considers performances subjected to involuntary muscular contractions and interference of control due to impairments and a high-level interpretation which considers the psychology of the response (with a stronger focus on the cognitive objectives of the test). The two views need to be integrated. A low-level view using streaks as a basis was provided by Gan (Gan, Frank, Amirabdollahian, Sharp, & Rainer, 2014b) (Chapter 0) in a motor-skills study and discussed briefly in a previous section (2.4.1).

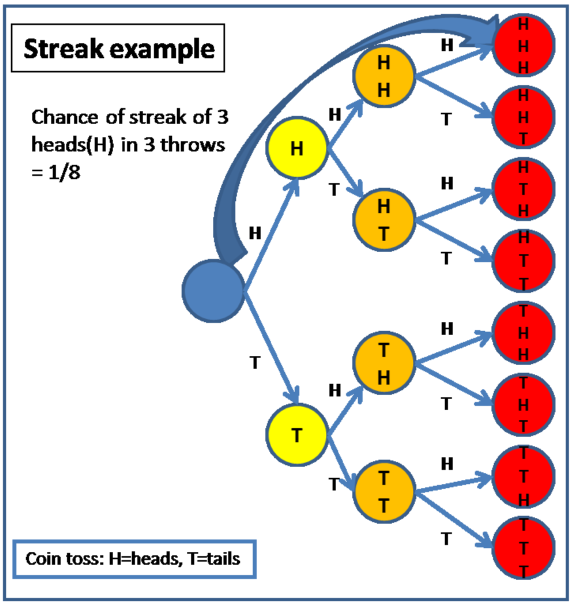
The work to extend ECDT to provide a finer granularity of indications and the integration of high and low-level interpretations is provided in a subsequent section of this thesis (section 9.1.2). The question "RQ1.3 Does a PNI child have sufficient motor control to use the device and undertake a learning test at the same time?" besides looking at the suitability of novel devices for the learning test also looks at providing a more informed test.

2.4.4. Development of new measure using *Streaks*

A previous section (2.2) which reviewed the impairments of PNI children suggested a complex range of disabilities which involve loss of control. We anticipate a higher error rate than would be found in non-PNI children. This study, when considering the problems of the severely impaired, uses success streaks instead of a simple number of successes and performance timing. The unintended inputs caused by the involuntary movements, cognitive impairment, fatigue of the child, together with mechanistic imperfections of the system, are treated as noise. Streaks are consecutive binary event sequences of the same type, and success streaks refer to a sequence of successes of two or more. Streaks have been the subject for mathematical publications (Makri, Philippou, & Psillakis, 2007). A simple coin tossing example using streaks is shown in Figure 3. A proof of the probability of obtaining a particular streak size is provided in the Appendix (Appendix 18). Streaks have previously been the subject of studies in gambling behaviour (Altmann & Burns, 2005; Ayton & Fischer, 2004). Streaks have been used to establish success in cognitive tests like the Wisconsin Card Sorting Test (WCST) (D. A. Grant & Berg, 1948) where reinforcement or learning is concluded after a certain number of consecutive successes and various other cognitive significances are drawn from other thresholds within the test. This study draws inspiration from those procedures to benchmark significant inputs from the motor skills test. So in the motor-skills test, a success threshold would exist as well as a threshold to separate intention from noise. It is notable that the term intention is used here to mean that the action is not the result of a reflex action but one arising from a conscious decision (Libet, Wright Jr, Feinstein, & Pearl, 1993).

One advantage of a parameter similar to the number of successes used in standard tests is that it is simple measure that provides a graded scale of success. It is helpful in the case of this study, to also start by considering a parameter which is able to generally provide a measure of success in order to assess suitability of different bio-modal devices. This study uses the maximum streak size (Gan et al., 2014b) in that it represents the best attempt in a block of trials. Mathematical proofs (Appendix 18) offer good supporting data but in a live environment, the probabilities are altered by the participants and the system used. Details of how the measure is developed can be found in the description of any of the major tests done in this study (Chapters 0, 8, 9). The best attempt and other significant attempts would be the result of (Oskarsson, Van Boven, McClelland, & Hastie, 2009) intention and control largely affected by practice(Ericsson et al., 1993), talent(Marcus, 2008) and noise. There is an indirect contribution to this noise caused by the motor-skills test which takes in a binary (YES/NO) response. If there are binary streaks of considerable length, the outcome would be a reflection of the algorithm that generates the sequence of stimuli presentation. This sequence is in general random but the algorithm can be improved to keep the streak noise low in those cases by ensuring that there are no more than 2 stimuli in sequence that require a YES or a NO response.

The literature suggests that the use of streaks is a better approach to interpreting responses of PNI children where significant responses are in error as a result of involuntary motor actions. The research questions "RQ1.1 Does a PNI child have sufficient motor control to use the device?" "RQ1.2 Does a PNI child have sufficient motor control to use the device and undertake a light cognitive test at the same time?" "RQ1.3 Does a PNI child have sufficient motor control to use the device and undertake a learning test at the same time?" provide a test-bed of results that the use of streaks can be applied to to obtain a more informed view.



The chance of a streak as this coin-tossing example shows gets rarer as the size of the streak gets larger and is used to classify test responses.

Figure 5 - Example of probability for streak of 3 heads

2.5.Conclusions

The literature review has opened up areas of investigations that require empirical work. We use the user-device-test triad to summarize the work required.

In terms of the user, the literature review has revealed that the main participants (PNI children) have a complex range of problems whose combination generates individual problems. Initial empirical work with such a population is required. An exploration pilot study was set up for this particular purpose (chapter 5). The use of streaks as a measure requires development which also used the pilot study.

In terms of the devices, the review has revealed options for each type of device that require hands-on exploration to determine which option is better. This work is required before more structured testing with the users is done. The next stage carried out in the PhD work flow is thus the exploration of devices (chapter 3).

In terms of the tests (section 2.4), the review has revealed that new tests are required as a result of the introduction of novel devices. In addition, the legacy test (ECDT) requires changing to accommodate the novel devices and the limitations of the test to indicate just failure. The tests have to be built into the operational model and this is described in Chapter 4.

A theoretical model of how devices would fit the PNI children was developed. The theoretical model needs an operational model to be implemented to verify the theoretical model as described in the previous paragraphs. The verification exercise is described as the main study (Chapters 6 to 10). The results of the verification exercise generate an empirical model further modifying the theoretical model. Speculation of additional factors then suggests how the modified theoretical model may advance with future work (Chapter 10).

3. DEVICE: Exploratory investigations of devices

3.1. Introduction

The literature review (chapter 2) has highlighted that all the devices chosen for the study have a range of options for configuration and use. Some exploration work has to be carried out to determine the appropriate choices and to work on prototypes of software to fit the device paradigm. An informal development cycle which consisted of an iteration of requirements, design and testing was undertaken on several prototypes. Each prototype investigated particular features of the devices to be used for the final pilot studies prior to deployment. Testing was carried out on members of the development team. Formal documentation of prototype development was not carried out.

A game called the TypeWriter which requires images shown to be spelt was initially developed to try out the EEG headset. This software was subsequently changed to work on the development of a hybrid device which involved the EEG headset and an eye-tracker. The software developed for the hybrid was another spelling game called PickNDrop.

3.2. EEG headset

The Emotiv™ neuro-headset was used in this study. Apart from the providing the EEG time-series, Emotiv™ is able to provide the following translations of EEG time series through their API:

- Expressive indications
- Affective indications
- Cognitive indications

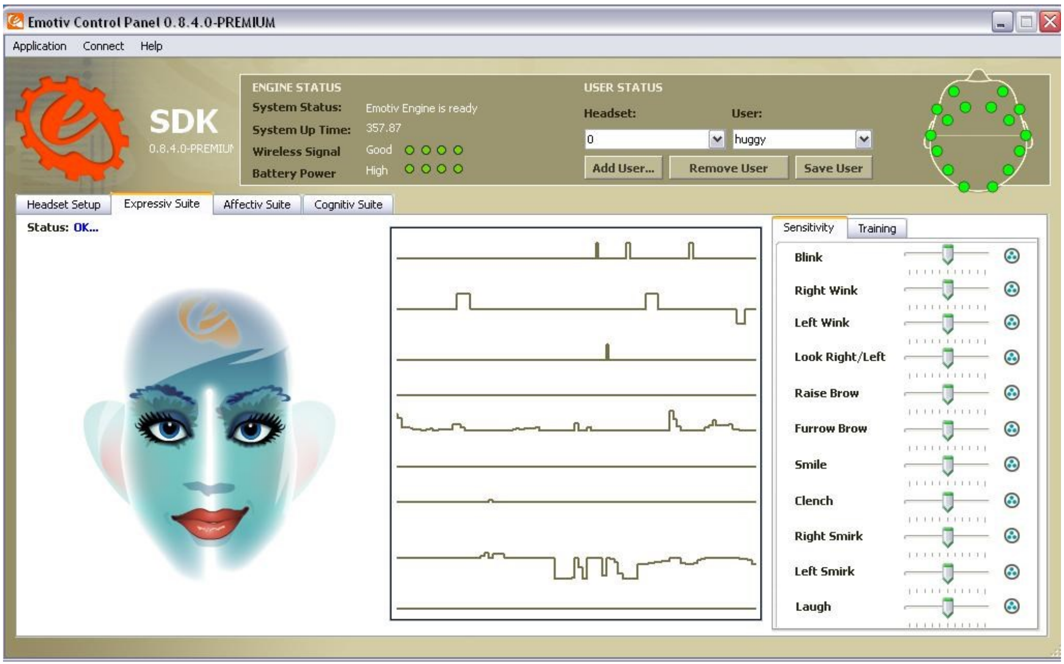


Figure 6 - Capture of facial expressions using Emotiv EEG translation

3.2.1. Expressive indications

Figure 6 show the use of the Emotiv API as applied to the capture of facial expressions. A blue avatar mimics the facial gestures of the user. Feedback is provided from a panel that shows a stepped output for the eleven gestures. False positives and negatives do occur. The sensitivity of detection can be increased or decreased. It is possible for all the sensitivities to be set to minimum and false positives to occur still. There is no standard setting that is guaranteed to produce reliable results. Each individual has to make some form of adjustment or provide some training of the machine. A blink may not be easily distinguishable in the processing from a right or left wink at the same time. Emotiv does not make the detections exclusive in that a blink cannot happen at the same time a wink happens.

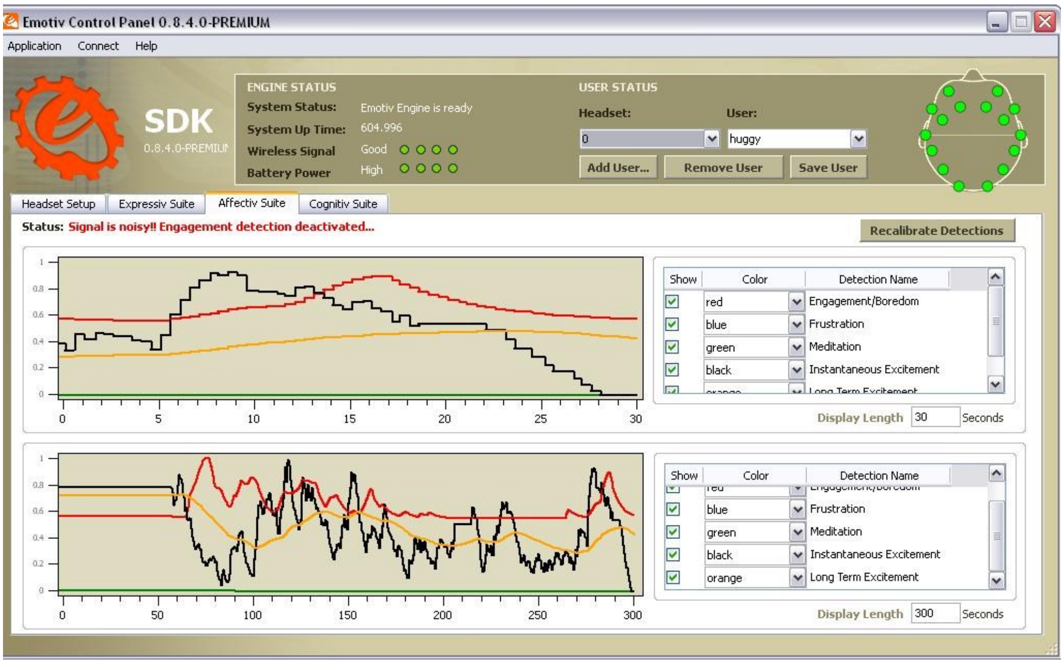


Figure 7 - Capture of emotions using Emotiv EEG translation

3.2.2. Affective indications

Figure 7 show the use of the Emotiv API as applied to the capture of emotions defined by Emotiv. The display is a monitoring tool that uses the API and encourages the development of particular mental states like excitement and meditation although it measures frustration and boredom as well. Only engagement/boredom (red), instantaneous excitement (black) and long term excitement (orange) is shown as this version of Emotiv (prototype) did not support frustration and meditation. The excitement level is the easiest to manipulate if one has a control of one's heart rhythms. Bringing excitement levels down to zero by clearing thoughts is supposed to bring meditation levels up. Frustration is difficult to achieve because it is an extreme mental state. Emotiv claims that they achieved consistent results when they ramped up the difficulty level of First Person Shooter (FPS) games so that the user was overwhelmed by the opposition (the bad guys). Emotiv guards the definition of the mental state signatures as their Intellectual Property. The company claim to tune the emotional states by making a psychological evaluation of various situations with EEG, oxygen levels, galvanic skin response and blood volume measurements. However, the Emotiv user manual does attribute the recognition of the Engagement/Boredom mental state as being associated with increased beta rhythm and alpha suppression.

3.2.3. Cognitive indications

Figure 8 show the use of the Emotiv API to capture emotions defined by Emotiv. The figure shows the training application which allows a user to train Emotiv to recognise different “mental commands” and use those commands to perform actions on a virtual cube. A Neutral Action must be trained for first to have a no-action command in the repertoire. After that a number of specific actions can be trained for, such as push, pull, lift, drop, left, right, rotate left, rotate right, rotate clockwise, rotate anti-clockwise, rotate forward, rotate backward and disappear. When a user trains Emotiv, the mental state is translated into the specific cube action. The term “mental command”

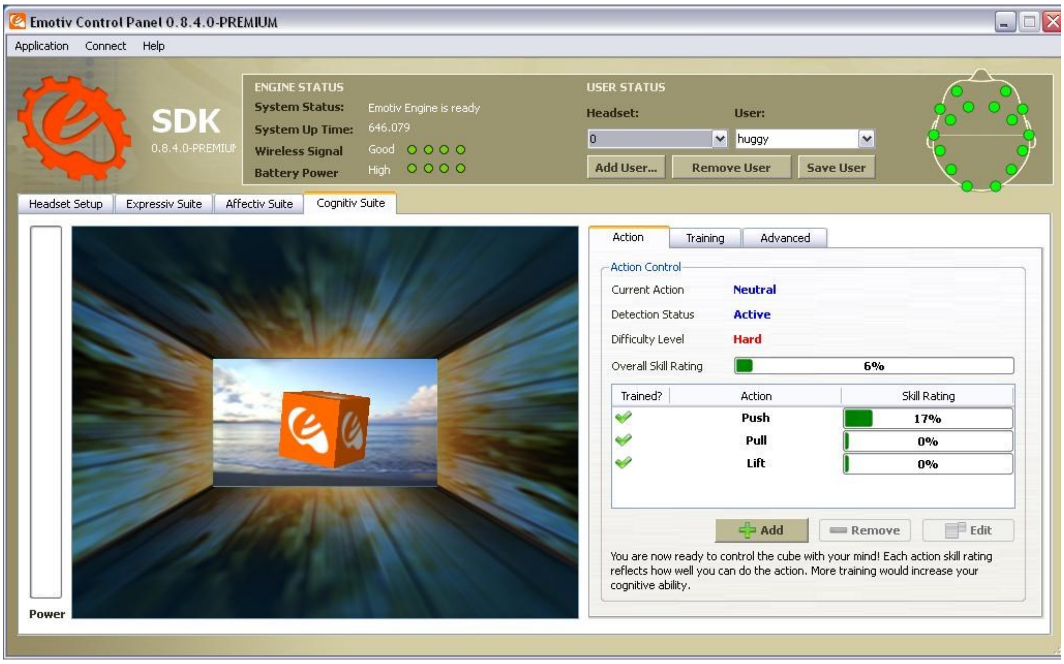


Figure 8 - Capture of trained EEG using Emotiv EEG translation

was used initially to mean the mental state as it is easier to explain using that term. The mental or EEG state generated does not have to be specific. The user can use the mental state for pushing the cube to mean lifting the cube. As long as one mental state is different from another, the system will execute the action associated with the training. The problem for the user is to remember the same mental state to use over time especially if he or she has trained for several commands. Users may find it easier to use facial expressions and gestures to associate a mental state to an action. It is possible that Emotiv will then associate the facial expression (specifically the EEG pattern generated by the facial expression) to the action. However, if there are artefacts, the signal may be rejected because it is too noisy. It is normally simpler to start with a Neutral Action and one other obvious action like Push. The user can be told to focus on the virtual horizon and imagine the cube getting smaller as it approaches the horizon. Once that is accomplished, other commands may be introduced. False positives and negatives can occur. Patents filed by Emotiv (Allsop et al., 2006) describe the use of a classifier to match mental states.

An assumption was thus made that it would be possible to train the Emotiv headset via the generic classifier to recognise either simple motor or imagined motor actions or artifacts generated by muscular action using bites (Reyes et al., 2012).

3.2.4. Aim

In order to use the EEG headset for interactions with the computer, the specific EEG patterns to be generated must be determined. The aim is therefore to test two general possibilities:

- Whether it is feasible to use mind control
- Whether it is feasible to use muscular artifacts and if so, which would be suitable

3.2.5. Method

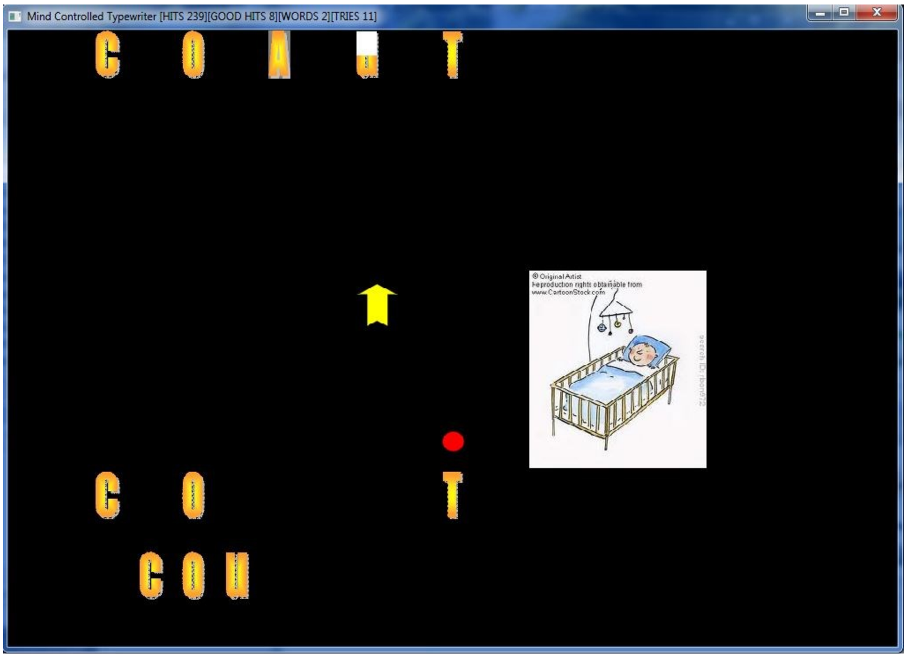


Figure 9 - TypeWriter game that uses Emotiv headset to choose letters by lifting an arrow to the letter

A software game, written in C++, called TypeWriter was developed as part of this research by the author. The game facilitated the testing of the EEG headset. The software structure of TypeWriter is similar to that shown in Figure 18 and Appendix chapter 17. Figure 9 shows the game presentation. The top of the screen displays a restricted alphabet {C, O, A, U, T}. An arrow (yellow) is used to "type" the letters. The arrow moves horizontally at the bottom of the screen constantly repeating the positions below the alphabets in a round-robin fashion. A set of characters defines the letters to be typed, in this case {C, O, and T} positioned at the bottom of the screen above the letters that have been previously entered by the player {C, O, and U}. A red dot above a letter of the definition set indicates the letter that needs to be typed in at the current instant. A picture of what needs to be typed appears on the right. The trained EEG LIFT command is used to move the arrow. A LIFT command causes the arrow to rise from the bottom (above the row formed by {C, O, and T}) to the top until it hits the character above which causes a small white rectangle to partially blot the letter hit. The arrow proceeds in discrete steps to the top (4 as default). At each step an evaluation is made if there is enough of a mental surge to trigger another step upwards. If there is not enough mental power the arrow resumes its position at the bottom. Once the player reproduces the specified set {C, O, T} a new set of characters will be specified for the player to reproduce.

The EEG LIFT command was trained using a variety of modes to be tested. In order to find out if mind control is possible, users were told to push at the table in the front of them without moving the table during the training of the LIFT. In order to investigate the feasibility of muscular artifacts, a biting action was used to train the LIFT. Subsequently, other muscular actions such as blinks and lifting of the eyebrow were also used.

3.2.6. Observations and discussion

The results of the different LIFT training modes when applied to the TypeWriter game is shown in Table 3.

Table 3 – EEG motor control results using Emotiv headset. Testing consisted of checking if the EEG signals produced by facial and body motor actions can be detected by the Emotiv interface.

EEG signals	Successful detection
Clench/grit teeth	yes
Blink	yes
Lift of eyebrow	yes
Sensorimotor control – movement	no

It was found that using mind control to train the headset did not work as the headset could not be trained to use the static push. Later in this section we show a user being told to use mind control but unreliable results were obtained. We show the possibility that a muscular artefact was used instead of cortical signals. The results of the use of different muscular actions show that the range of muscular actions chosen could all be used as LIFT control in the TypeWriter game by both the adult and the PNI child. There are a number of issues with the use of the headset by children. For users with small head sizes the headset does not fit well. Some users will feel that the headset is uncomfortable within an hour and would want to remove it.

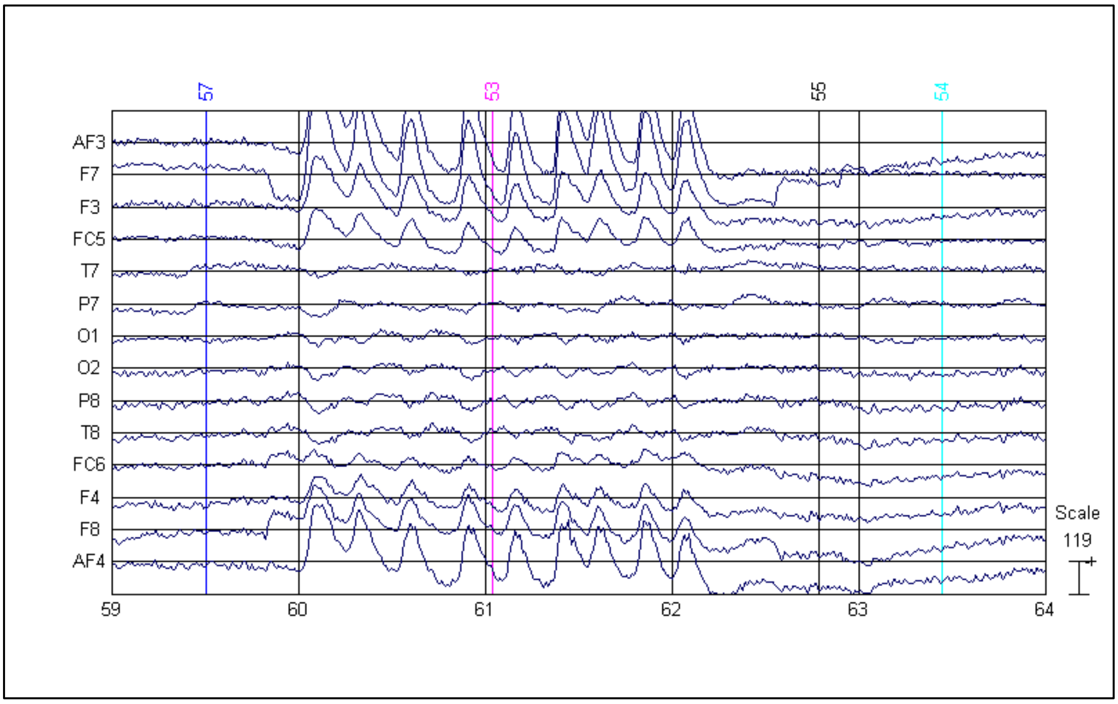


Figure 10 – EEG trace show that it is possible to use blink. Latencies of detection are shown by numbered coloured vertical lines.57 = stimulus start, 53 = Emotiv detected start of surge, 54 = Emotiv detected end of surge, 55 = Emotiv detected peak of surge. Large EEG fluctuations show user response to stimuli. Horizontal axis measures time in seconds and the vertical axis, signal voltage in microVolts.

A detailed examination of the EEG time-series during a test run with the TypeWriter show clear indications of blinks. Figure 10 is a plot of the 14 channels (sensors) of the EEG headset. The vertical axis shows the signal voltage in microVolts and the horizontal axis shows the time in seconds. The blinks are easily distinguishable as “hills” that seem to appear about 60 seconds after the trace capture and end around 62 seconds. The “hills” have a frequency of around 4 Hz (4 hills per second). Coloured vertical lines tagged with simple numbers (53, 54, 55, 57) are markers that show the onset of various events of interest. The numbers are arbitrarily chosen and are ascii encodings of simple numbers (5, 6, 7, 9). These numbers are sent by TypeWriter to Emotiv to integrate into the EEG time-series at the time of receipt.

When TypeWriter first positions the arrow under a new letter, it will send 57 (blue) to Emotiv to indicate start of stimulus. We can see latency in the reaction of the adult as the blinks (hills) occur 0.5 second after the start of the stimulus. Emotiv have to process the blinks and translate that to a LIFT indication. When the first LIFT indication is detected by TypeWriter, it sends

53 (Pink) to Emotiv. LIFT indications range as a fraction from 0 to 1 and occur as surges that peak and then drop to zero. The peak of the surge is detected by TypeWriter about 1.75 sec after detection of the start of the surge and 55 is sent to Emotiv (black). Emotiv sends a cessation of surge (zero LIFT) to TypeWriter about a sec after the blinks cease which is indicated by TypeWriter with a 54 (cyan).

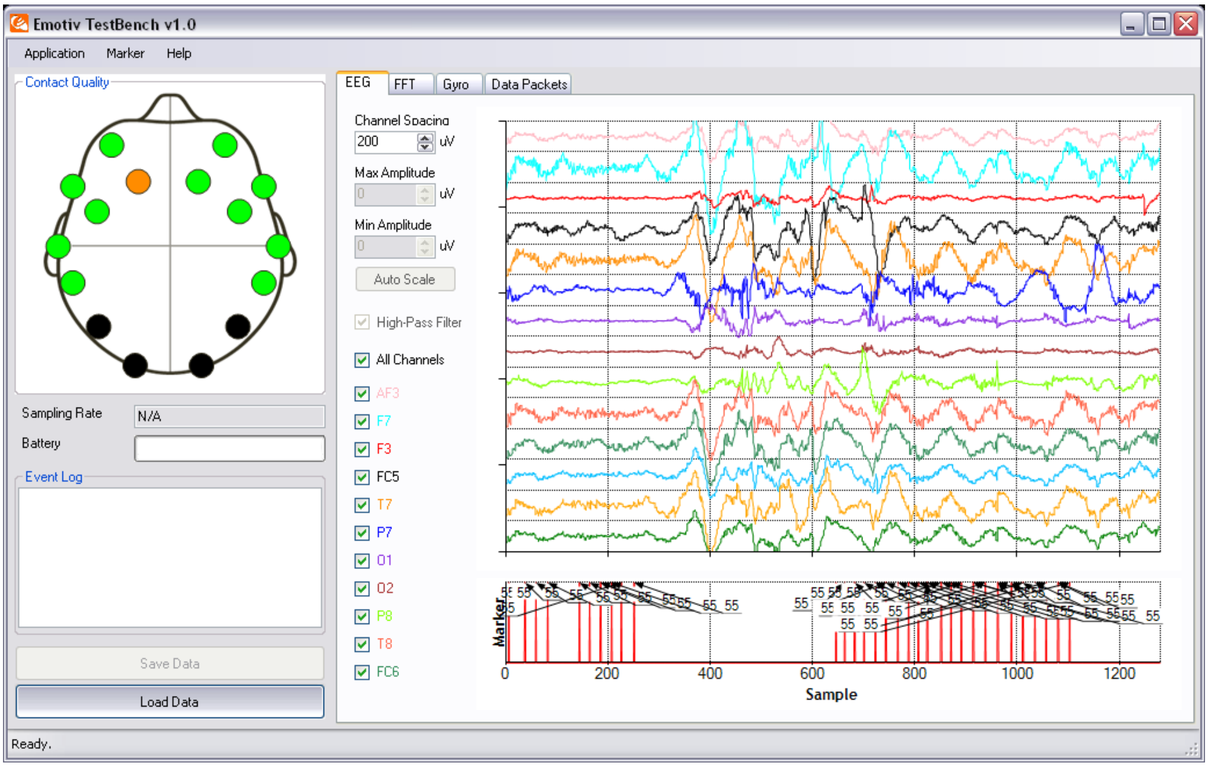


Figure 11 - User instructed to use mind control but uses facial expressions instead. 14 EEG channels (time-series) of the headset are displayed as plots with varying colours. The positions of the channels are shown by the dots on the silhouette of the head giving a scalp map. The notation used for the channel position corresponds to the clinical 10/20 standard (JASPER, 1958). The vertical axis shows the EEG signal voltage in microvolts. The horizontal axis shows time in sample numbers. A conversion to seconds has to be done using a 128 samples/sec scale. 200 samples would be roughly equivalent to 1.6 seconds. The bottom plot shows marker positions signalled to Emotiv by the test program and each red marker (55) represents a LIFT command. A strong series of coloured plots mark muscular action (e.g. F7, FC5, T7, P7) which corresponds to the region near the left side of the head above the ear. The signals on the right side (T8, FC6, F4, F8, AF4) are less prominent. The red markers are delayed by about 250 samples from the start of the muscular artefact (starting around the 380th sample) which is about 2 secs.

The results show the huge difference between the blink signals and normal EEG signals shown as little squiggles (around 10 microvolts) between the hills (can be around 200 microvolts). The blinks occur in the low frequency end (4 Hz) and swamp any cortical indications of delta (1 to 4 Hz) and theta (4 to 8 Hz) rhythms. Such muscular artifacts manifest themselves as high energy low frequency EEG signals and would probably have a strong influence on any classifier.

Figure 11 show the EEG time series of a user in a separate experiment who was instructed to imagine LIFT of the virtual cube shown in Figure 8. The same sort of LIFT command would have been used in the testing of the TypeWriter game. The observations suggest the possibility that during training of mind control the participant unconsciously uses the stronger signals that are generated by a facial expression or the widening of the eyes or a movement of the head for mind control. Similar separate experiments were carried out by adults and one tendency was to move the head up and down in order to LIFT the virtual cube. The movement of the head causes a whole series of other muscular actions including the movement of the eyeball. From literature on EEG artifacts it is also known that the friction of the eyeball against the eyelid (Schalk & Mellinger, 2010) can result in EEG artifacts. These muscular actions and eye movements are picked up as training signals by Emotiv and consequently control through the Emotiv headset can only be used with confidence when a particular artifact signal such as a bite is used. The choice Emotiv made for positioning of the sensors round the head provide more efficient detection of EEG artifacts than the detection of cortical signals.

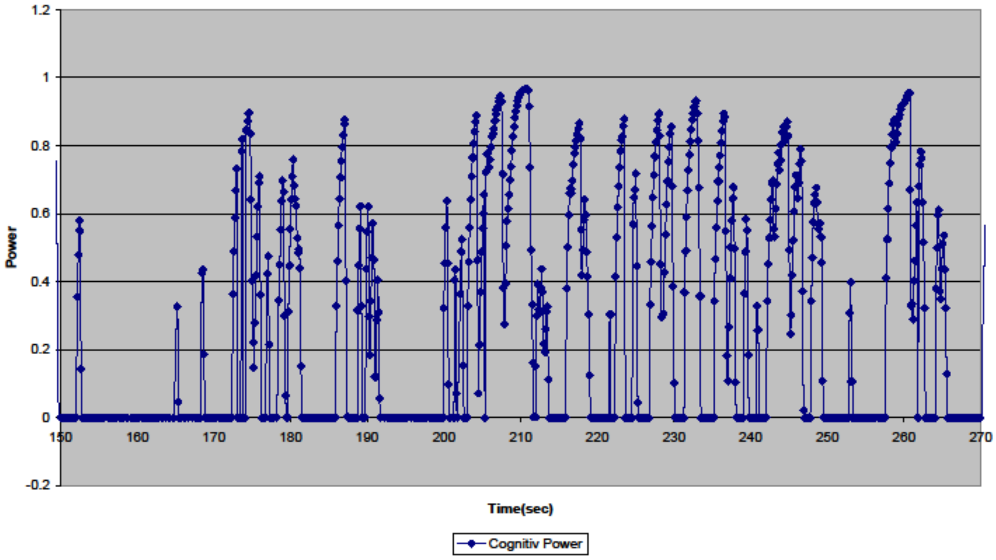


Figure 12 - Emotiv LIFT commands shown as surges of power levels as provided by the Emotiv API. A cluster of surges are required to raise the virtual cube shown in Figure 8 and maintain the cube at the peak position. An imagined lifting of the cube was used to train Emotiv.

Figure 12 gives an example of the trained Emotiv LIFT command as applied on the lift of the virtual cube shown in Figure 8. The vertical axis is a representation of the effectiveness (power) of the LIFT expressed in a range of values from 0 to 1 to 6 significant figures. The horizontal axis is a measure of time in seconds. In order to maintain the cube aloft for a long period, multiple surges are required. The height to which the cube moves is related to the power of the LIFT.

The excessive nature of the surges suggested that some form of high level filtering similar to that provided by debouncing filters for switching is useful especially when parameters can be configured. The filtering involved was implemented as a "think-time" filter where intention was signalled only if a thought was made for long enough duration and the total effort involved was sufficient.

This filter was particularly useful in the case of a hybrid implementation where an EEG headset was involved with an eye-tracker. The requirements for "think time" could be reduced because of an overlap of the selection function from the two bio-modal modes (section 3.3.4). Normally a "think-time" of 2 sec may be used but in the case of a hybrid, a 1 sec think time is sufficient.

3.2.7. Conclusion for EEG headset

The results show that the use of mind control using software provided by Emotiv would be unreliable. Better results may be obtained by proprietary software which would then have to include filters, a classifier, translator and trainer. The use of artefacts based on facial expressions and Emotiv software is reliable enough and easier. The use of the headset would be more practical for children as they tend to be restless. In order to avoid influence of muscular artifacts, the children would be required to be very still. The training for PNI children should be simpler as the instructions would not require something abstract that is difficult to verify.

The development of the TypeWriter game provides a framework for a more sophisticated game to be developed for implementing hybrids of the EEG headset and the eye-tracker.

3.3. Eye-tracker and fusion

3.3.1. Introduction

The Tobii™ remotely mounted eye-tracker was used for this study. As the Tobii API only provide raw eye-gaze data through their API, any filtering or smoothing of eye motion software has to be written. As the literature review (section 2.3.4) pointed out that early fusion was inappropriate, late fusion was used in the design of the hybrid device. In the design of hybrids for this study, there is an interest in providing a device that notionally focusses on the reduction of fatigue. A model was looked for that was based on the natural way that the body works. One such model was where the eye is constantly picking up objects of interest and then cognitively examining it. Such a model is supported by HCI studies that indicated that users tend to use complementarity in modes rather than redundancy (Oviatt, 1997). The eye-tracker was thus integrated with an EEG headset using a new game called PickNDrop described below and in the Appendix (chapter 17).

3.3.2. Aim

The aim of this test is to determine if a hybrid device using both eye-tracking and bite is able to provide the control to enable the manipulation of virtual objects on the screen.

3.3.3. Method

A software game, written in C++, called PickNDrop was developed as part of this research by the author. The game facilitated the testing of the EEG headset and the eye-tracker. Details of the software structure of PickNDrop is provided in the Appendix (Chapter 17) and Figure 18. Figure 13 shows the game presentation. The letters displayed can be selected by using the eye to point to letters and using bites with the Emotiv EEG headset. A cursor (white rectangle) that is controlled by the eye movement is used to point. As the cursor is within range of selecting a letter, an arrow appears. When the arrow rotates, bites must be used to pick up the letter. A rotation is used as a feedback mechanism because the user trains the Emotiv headset for a ROTATE-CLOCKWISE (Figure 8) command. The cursor is at times annotated with other symbols

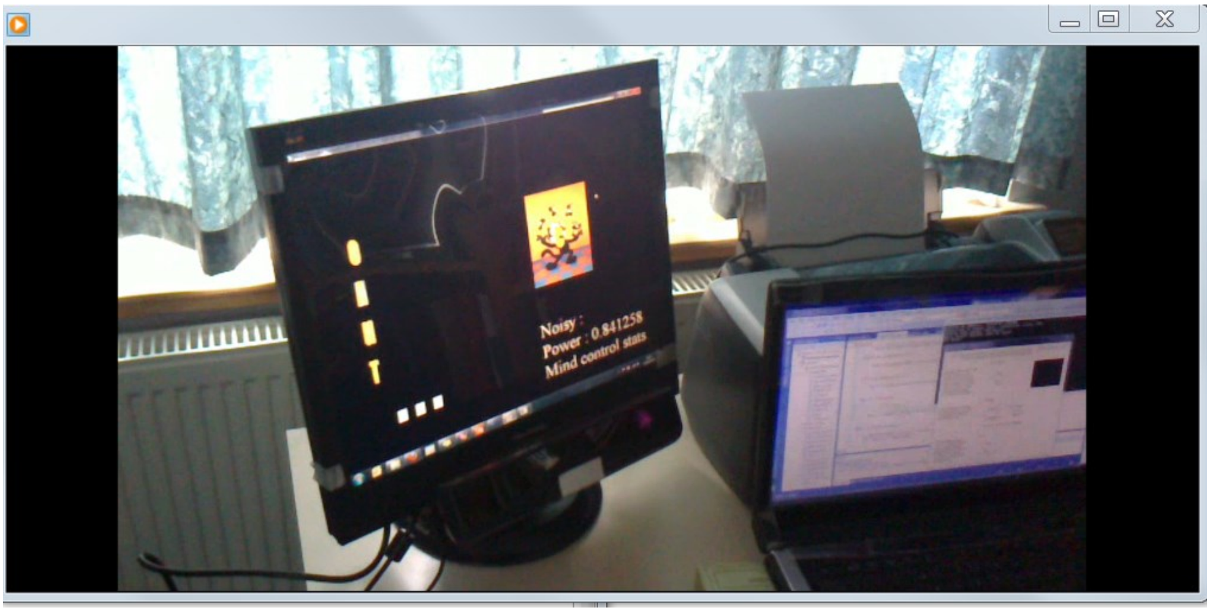


Figure 13 - PickNDrop game to test the fusion of EEG and eye-track control

to provide additional information. If the cursor is appended with a loop, it means that eye contact was lost and the head position probably needs adjustment to maintain contact. The procedure is similar to a mouse operation. The eyes are used to move the cursor and bites used to click (select).

The same action is required to drop the selected letters in the image of the cat and haircut. The game requires the user to spell the image (CAT and CUT). Once the letter is selected, it will be piggy-backed onto the cursor. (This is similar to dragging a mouse object). The eye gaze is used to move the letter to the image. When the cursor is within the image, an arrow will be displayed in the image which will start to rotate. This is the cue to bite to drop the letter. Once the letter is dropped it will appear in a row in the bottom left hand side which records the attempts at spelling the image.

If the correct characters are spelt, for either image, a smiley face will appear.

The Noisy field in the gaming window gives an indication of whether brain signals are being rejected and the Power field gives an indication of the effectiveness of the bite.

An alternative version of PickNDrop was also developed to investigate the effect of only using the eye-tracker for pointing and selection. In this case, instead of using bites to select "dwell-time" was used as the selection mechanism. The user had to fixate on a letter for certain "dwell-time" duration before it was picked up and subsequently fixate on the target location before the letter was dropped.

3.3.4. Observations and discussions

PickNDrop when tested on an adult proved successful. The adult was able to spell out both CAT and CUT words. The setup for PickNDrop was initially useful for exploring the teething problems posed by the eye-tracker when the API was used to pick up the eye-gaze positions. The API for the eye-tracker only provided raw gaze data for the position of each eye. The application software has to cater for the determination of a common gaze position for both eyes, filtering any noise, determining the rate of motion of the eye-cursor projected onto a screen and the handling of out of range gazes.

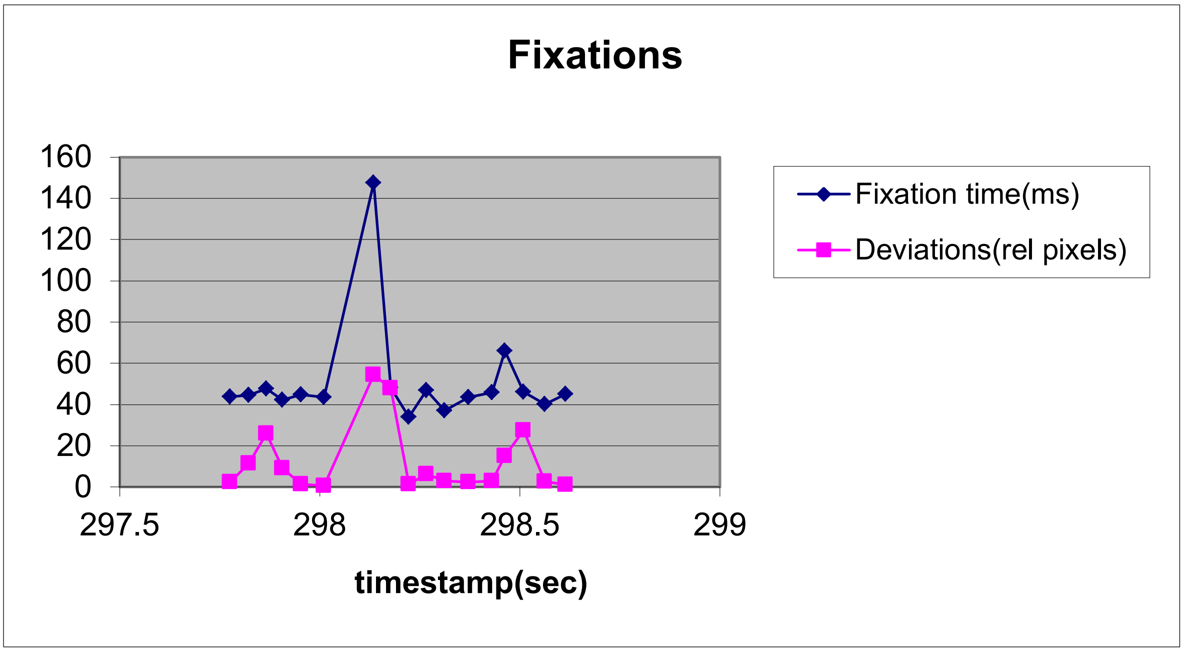


Figure 14 - Ocular fluctuations. The horizontal scale is the operating system time (in this case Windows time). Fixation is shown in blue and shows the length of time (eye-tracker-Tobii time) each gaze sample remains stationary. Deviation of gaze is shown in pink and shows the distance moved in relative pixels (x 100) from the previous position. Together they give a picture of the ocular fluctuations presented by the eye as it shows small jitters about a position before darting off to other locations with subsequent jitter about the new location.

The figure above shows a fixation behaviour record derived from the raw gaze data of a participant. From this record, we wish to get an idea of the behaviour of random jitters and normal fixation times of the eye. The eye is always scanning for items of interest and is in constant motion. Gaze positions are obtained as x,y co-ordinates for both eyes in terms of relative pixels. Relative pixels (range from 0 to 1) normalize the pixel position to a percentage of the screen width and height. The average x,y co-ordinates for both eyes are calculated to provide a single x,y coordinate. The x-axis in the figure is a record of the sample timestamp and the y-axis is the recorded difference of time spent and spatial deviation between time-stamps. The blue trace indicates the time spent (in ms) on any one gaze location before the next update. The pink trace is super-imposed on the graph plot. Each pink point represents the deviation from each gaze location for each time-stamped record. The deviation is calculated as the Euclidean distance of the new gaze position in relative pixels to the previous. All calculations of deviation are then scaled by 100 to form the pink trace thus providing a compatible scale with the fixation time characteristic. The sample rate of the eye-tracker was taken to be around 40Hz which meant a minimum of 25 ms update sample. From this characteristic, it was established that most values lie around 40 ms and the highest value was 150ms. Thus any value above 500 ms would represent a good value for dwell-time for this range of eye-trackers. The dwell time is an inordinate amount of time that the eye normally spends on an object of interest and used to signify interest in an object to an application program.

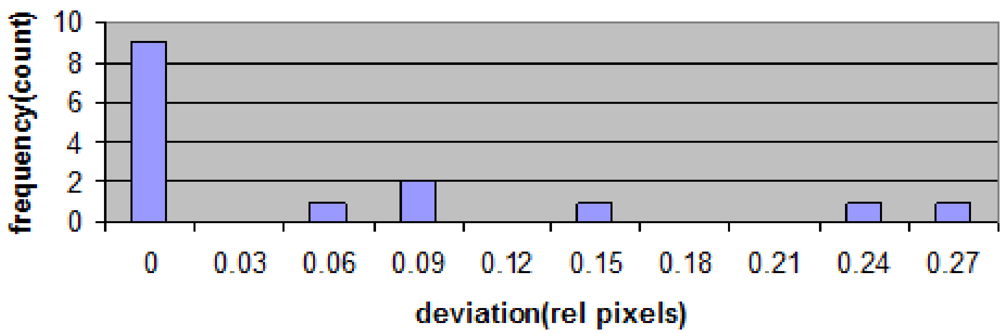


Figure 15 – Frequency distribution of eye-gaze deviation (ocular fluctuations). The horizontal axis is the distance (deviation) between two consecutive eye-gaze positions. The eye-gaze position is expressed as a fraction (relative pixels) between 0 and 1 of the screen width and height. The deviation is the Euclidean distance between the two consecutive eye-gaze positions.

Figure 15 shows the distribution of eye-gaze distances between consecutive fixations. The distribution shows the cluster of points where the count is highest for a small deviation (less than

0.03). It is assumed that the cluster represents the majority of the ocular fluctuations. The screen dimensions of the test were 1280 x 1024. The 3% jitter on such a screen would result in a 40 x 30 pixel fluctuation. The use of a filter to stop jitter by keeping the eye-cursor steady for all eye gaze less than 100 pixels apart should be adequate. This solution was used for PickNDrop and found to be acceptable.

A second stage of smoothing was found to be helpful as there were still strong jitters that are made as the eye gaze shifts from pick up to target for distances greater than 100 pixels. This stage of smoothing was implemented as a configurable parameter in addition to the 100 pixels no-movement zone. The eye gaze movement could be decelerated by a particular percentage. For example, if the deviation was 400 pixels, and a 50% deceleration was required, the eye gaze would be allowed to move 400/2 = 200 pixels. The possibility that an eye-gaze can never reach its target was discounted after tests where it was found that there was enough eye movement to discount that. There are other algorithms that can be deployed apart from the simple constant deceleration proposed. The deceleration could be made non-linear using an approximation to the Gaussian curve such as a cubic hermite spline polynomial (Lazzari, Vercher, & Buizza, 1997; D. P. Mitchell & Netravali, 1988)

The policy for error handling when the eye drifts outside the screen area is important when handling of hybrid devices involve steadying the eye cursor when a selection using a bite is required. The policy may dictate that the cursor always resets itself to a certain position on the screen (for example the top right hand corner) or that the cursor stays where it was when tracking has been lost or out of range. A policy that resets itself to another position has the disadvantage that it increases the jitter effect significantly.

In order to make fixation tasks (where “dwell-time” was required) easier on the user, an invisible field was placed around the objects that the user had to fixate for selection. This field extended the window of selection for the user. The user no longer had to be as accurate in fixating at the object but merely have to come within range of the object. The field size was a configurable parameter of the test program. Although the field was unnecessary in PickNDrop, it was found useful in the main study implementation. One problem with implementing a field of this nature was that there had to be enough space separating all the objects that require fixation.

The advantage of fusion of two bio-modal devices to form a hybrid was that each mode can compensate for the weakness of the other mode. In the case of the eye-tracker when the eyes accidently look at an object for too long, it triggers a false positive (Midas effect). The use of the EEG headset compensates for that by removing the need for a “dwell-time” to indicate selection and the “think-time” is used instead. Another compensation that could be made was that the individual requirements for operation of each mode could be relaxed. The criterion for making a selection no longer has to be for example 2 seconds of think-time. In pointing the device at the object required for selection (which required a degree of latency to selection as the gaze has to be stable), provided a certain amount of redundancy to requiring a further “think time”. The think-time could thus be reduced to a lower value (for instance, 1 second). Similarly, pointing could be relaxed as it would not be necessary to point with precision at an object of interest as selection also has to be indicated. The invisible field described in the previous paragraph could be made to relax the amount of effort required to point.

3.3.5. Conclusion for fusion of EEG headset and eye-tracker

In the exploration exercise for the eye-tracker, parameters required to implement ocular fluctuation filters based on smoothing out the fluctuations were determined.

The integration of both EEG interaction and eye-gaze proved to be a viable form of multi-modal operation using a form of pick-and-choose as a model. It has the advantage over the single eye gaze mode in that it avoids the Midas effect. It compensates for using two modes by reducing the effort for each mode that makes up the multi-modal device.

The PickNDrop game developed for this exploration exercise provides a useful prototype for developing the software for the actual study.

3.4.Head-tracker

3.4.1. Introduction

Camera Mouse™ was used as the head-tracker for this study. As Camera Mouse™ does not provide access to an API some other method for translating movements of cursor position to test responses have to be used. Since Camera Mouse already translates head movements to mouse cursor movements the cursor position can be extracted from the operating system (Windows™) and used to work out the responses made by the user.

As the head-tracker appears relatively simple to operate, the integration of this device to an application was left until the actual tests for the study was implemented.

3.4.2. Aim

The aim of the test is to decide the best body feature to use as a reference point for the head-tracker.

3.4.3. Method

A range of reference points was used to investigate good reference points that can be used for head-tracking. This included the finger, the corner of the eye and the tip of the nose. Figure 16 shows the procedure used to define the reference point. A feature of the body is selected by

pointing the mouse at the feature and clicking. A green square appears on the point signifying the area of the body that will be used as the reference point. Various movements were made to ensure that the use of the reference point was stable. The reference point was made to move horizontally and vertically across the screen to cover most of the area of the screen.

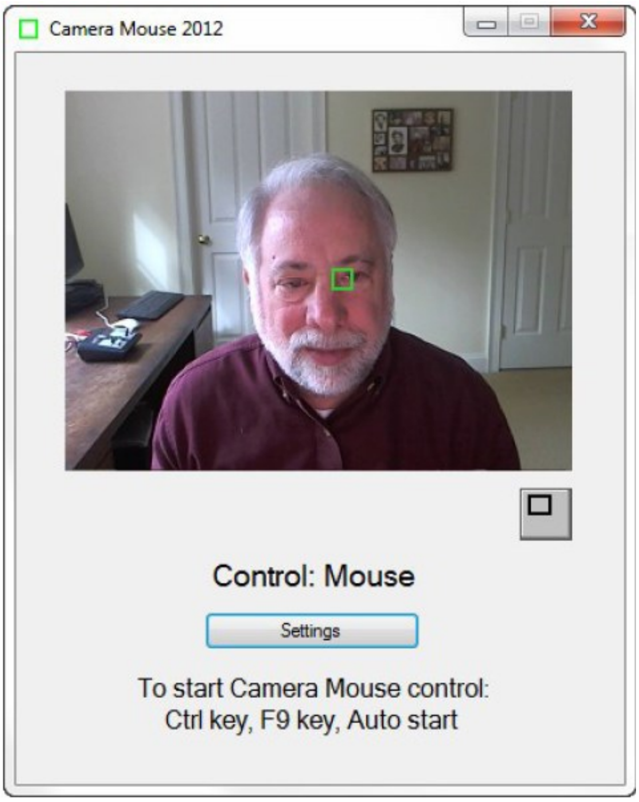


Figure 16 - The calibration procedure of the head-tracker. A point on the face (shown by a green square) is used by the head-tracker to determine the orientation of the head.

3.4.4. Observations and discussion

Table 4 - Head-tracking reference point tests. The trackability results show the optimal point for use as reference by the head-tracker.

Body feature	Trackability
Tip of finger	Lose track of reference point on the finger easily. Finger tends to move too quickly and loses reference.
Tip of nose	Relatively good results because head does not move much
Corner of eye	Relatively good results, comparable with nose tracking

The results of the reference point tests are shown in Table 4. When the tip of the finger was used as a reference point, the reference was lost too easily as the finger moves. The speed of movement of the finger appears to leave the reference point behind in a previous location.

When either the tip of the nose or the corner of the eye was used as a reference point, the results were more stable. The relatively slower speed of the head movement turning side to side and up and down allowed the reference point to be better tracked. However, the results with the fingertip expose a vulnerability of this device. If movement was to be too violent, the reference point would be shifted. With a large area like the face, the shift will mean that the reference point moves to another part of the face. It is always possible to move the reference point back with calculated movements. If the reference point moved to an area beyond the face, there is still a chance for recovery by moving the face so that it overlaps the area occupied by the new reference point.

3.4.5. Conclusion for head-tracker

The test confirmed the use of the tip of the nose as a reference point. The software use of an indirect mechanism to interact with the device due to the unavailability of an API did not show significant problems with latency. This mechanism could be used to provide a generic form of basic interaction with any device which does not have an API available.

3.5.Conclusion

A range of devices which required different types of motor capability proved to be viable for this study. The motor capability required was in the regions of:

- The hands (mouse/switch)
- The eyes (eye-tracker)
- The neck (head-tracker)
- The face (EEG headset using bites)
- Combination of the face and eyes (EEG and eye-track hybrid)
- Combination of the face and neck (EEH and head-track hybrid)

Hence, a range of devices that exploits a diverse range of motor capabilities should be made available for testing on PNI children. This exploration phase is important for the EEG headset because an important decision of not proceeding with using mind control due to the inherent noisy properties of the EEG headset system was found.

The next stage is to set up the system to be used in the main study which requires the strategies for test implementation to be resolved. The implementation of both TypeWriter and PickNDrop allows for incremental development where elements of the previous implementation can be re-used in the main study implementation. This is described in the next chapter (4).

4. TEST: Design and implementation of the test system

4.1.Introduction

The previous section explored the devices that were chosen for the study and we gained knowledge of how the devices could be made to work. This section describes the implementation of the system that supports the study. The implementation needs to consider the operational model of the study, especially the tests that need to be implemented as they have been put in the background so far. The operational model of the study is the supporting environment and processes that is needed to support verification of the theoretical model introduced in section 2.3.6 which describes a hypothesis for how devices would fit the capabilities of PNI children.

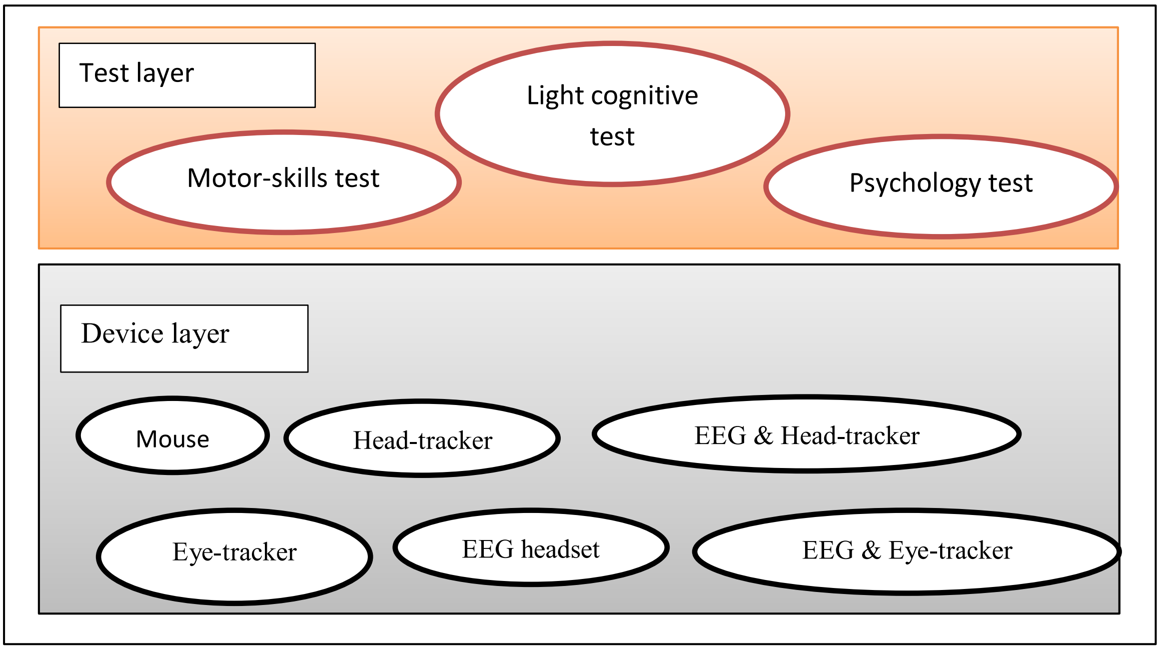


Figure 17 - Psyborg logical layers

A system consisting of both hardware and software is required to handle both tests and devices. The system is called Psyborg (Figure 17). Psyborg had two developmental phases. The first (Psyborg 1) was constructed to support the operational model of the study. The second (Psyborg 2) was developed as a requirement for the ICASE funding (section 1.6). Psyborg 2 took 3 months to build and consisted of packaging Psyborg 1 as the first stage of commercialized product. Psyborg 2 also included the capability to connect to a web-server and established a prototype capability to provide a secure registration facility for users and to transport the data collected in a study for storage in a remote server. Psyborg 2 consists of a single installable package. The screen shots shown in this chapter are taken from Psyborg 2.

This section is a high level guide to the system used for the study. The section does not provide a detailed software design description but serves to provide basic functional views of the system.

4.2. Aim

The aim of this section is to describe the design and implementation of the hardware and software system to support the operational model.

4.3. System overview

Figure 18 provides a basic overview of the system that was used to support the study. The hardware setup consists of input devices that provide user responses to the laptop or pc and an external screen monitor. The laptop is connected to the monitor (using a VGA connector) to output results and experimental stimuli. The input devices consist of the EEG headset which communicates with the laptop using a wireless (Bluetooth) dongle. The eye-tracker is mounted below the screen monitor and consists of infra-red illuminators and cameras. The eye-tracker is connected to the laptop using a USB connector. The head-tracker uses an externally mounted web-camera that is connected to the laptop (with a USB connector) to track segments of the head image.

The software used in the system is split into four general components:

- Applications that are specific to each input device and interfaces with each device using their API
- A multi-modal interface that knows how to integrate single or several modes to the system

- A test program that interfaces with the multi-modal interface to obtain test responses. Test results are stored on the laptop. Test stimuli are generated onto the externally mounted screen
- A system manager (Figure 20) that controls specific test programs and options for the study.

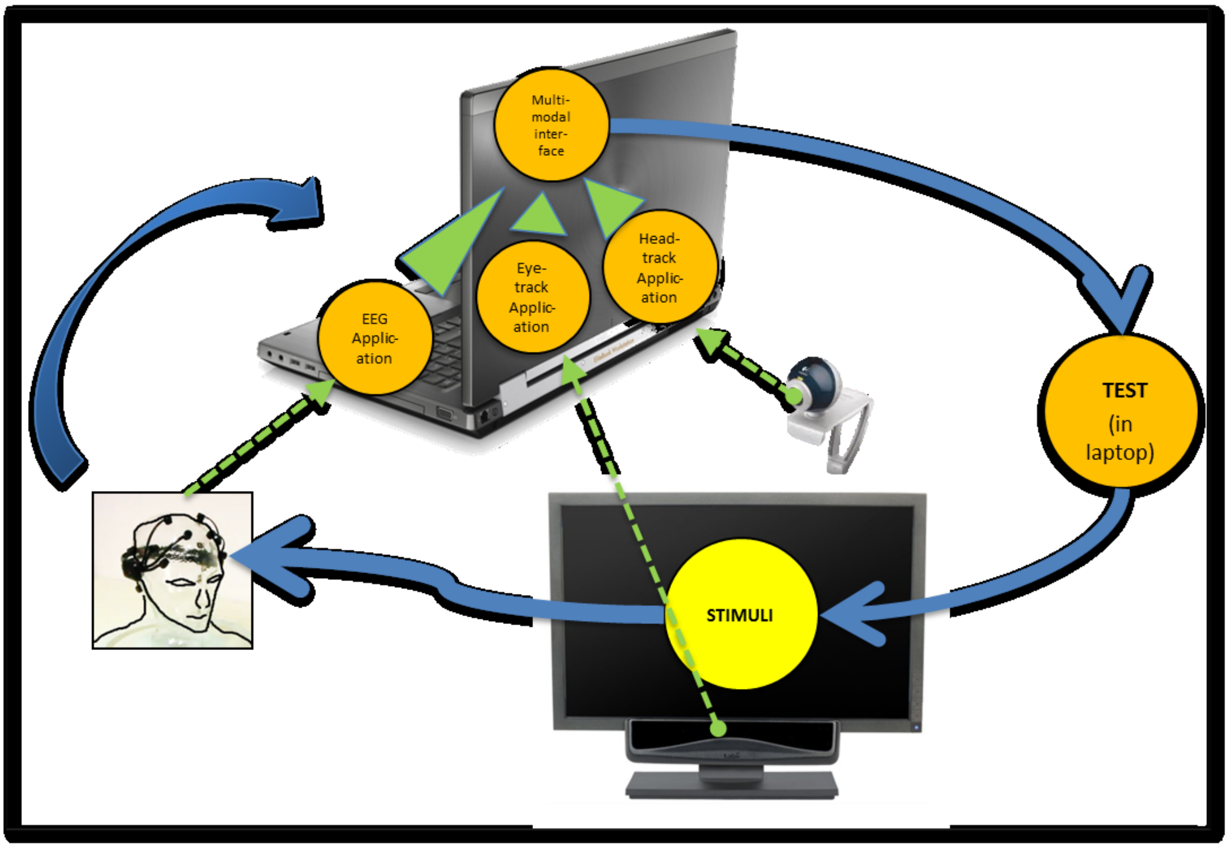


Figure 18 – The system used to support the operational model of the study. The participant uses an EEG-headset, an eye-tracker and a head-tracker to respond to the stimuli on the screen. The devices pick up signals from the participant and send them to their respective applications on the laptop. A multi-modal inter-face arranges for the appropriate combination of signals to form the response to the test program which generates the next set of stimuli.

4.4. User interface overview

Figure 19 shows the user's entrance point into Psyborg which is the Control panel. The Control panel is a Psyborg 2 implementation and not all features require description for this study. The features provide a general overview of the system functions of Psyborg which will be subsequently described.

4.4.1. Test programs

Psyborg is able to launch three different tests (section 6.1.4). The tests vary in cognitive complexity. The *Motor Test* (COMPTEST) option is used to check that the device fits the user. The

Cognitive Test (CATTEST) option is a light cognitive test to check for basic cognitive problems. The *Psychological Test* (ECDT) option is a learning test. Each test requires a device (section 4.4.2) to be chosen.

There are configuration parameters that can be set for each test as shown by the tabs in the figure. *Common Test Options* apply to all tests. *Motor test options*, *Cognitive test options* and *Psychology test options* apply to the respective tests.

Test results may be obtained from the *Maintenance* menu. *Read file* shows the directory where all the results are kept and allows the user to choose the appropriate log file which holds results for a particular test or device. The results are displayed on the *User Notifications* window.

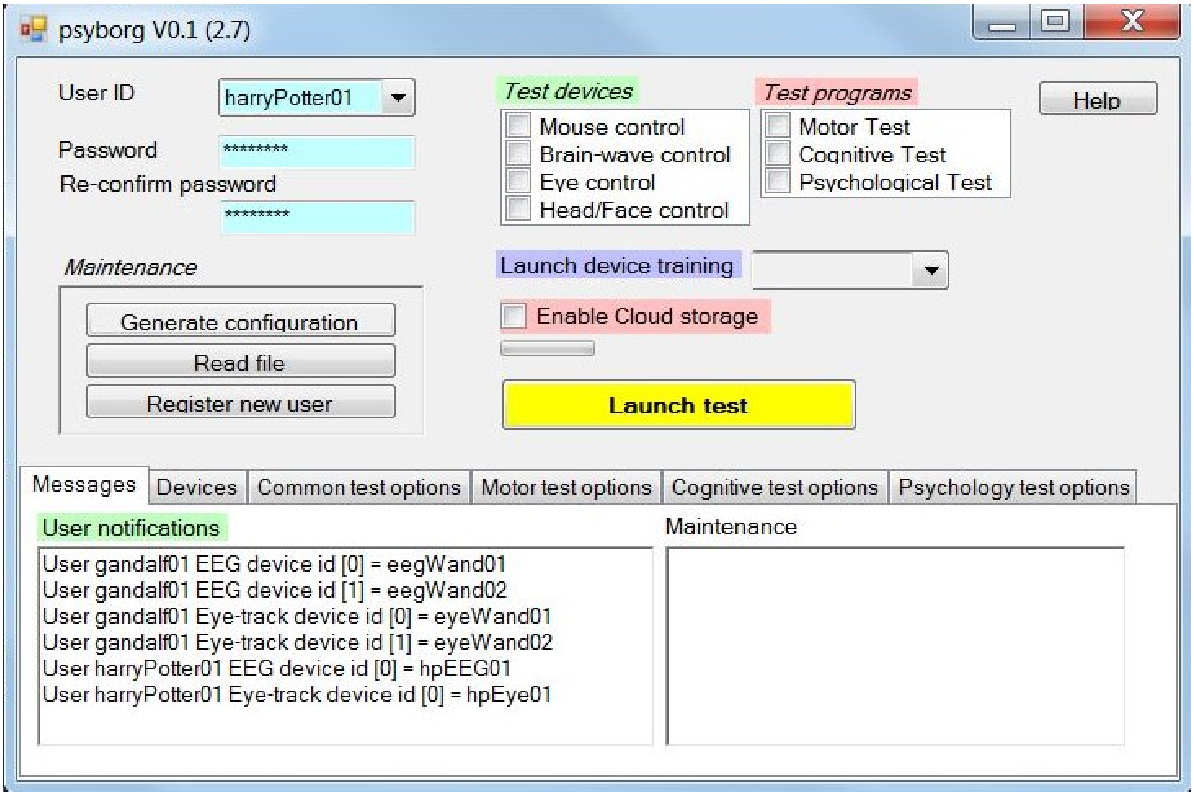


Figure 19 - Psyborg control panel showing general functions provided by Psyborg

Each test may launch a familiarisation program that allows the user the opportunity to get used to the test environment such as the device to be used for the test. The launch of the familiarisation program is one of the test options.

Execution progress can be tracked by an expert user using the *Maintenance* window.

4.4.2. Test devices

Psyborg allows one of 6 devices to be used with a test:

1. Mouse/switch
2. Eye-tracker
3. Head-tracker
4. EEG headset
5. EEG headset and eye-tracker hybrid
6. EEG headset and head-tracker hybrid

The devices are chosen by ticking one or more of the *Test devices* option.

Devices may require training or calibration. There are separate programs that have to be launched to perform the training or calibration. These can be done from the *Launch device training* drop down list. The EEG headset requires training with a virtual cube (Figure 8). The eye-tracker requires a calibration (section 6.1.3.3) program to be launched. The head-tracker requires a spot on the face to be specified for tracking (Figure 16).

4.4.3. Test parameters

Test parameters for stimulus control are found in the *Common test options* tab. These parameters include the control for the length of time that a particular image (stimulus) has to be presented (*Stimulus duration*) and the duration for the pause between images (*Inter-stimulus pause*).

4.4.4. Device parameters

Device parameters that affect the test display are also found in the *Common test options* tab. These parameters include device debounce-time, device deceleration rate and image field expansion sizes.

Device debounce time include the “think-time” (section 3.2.6) that is used by the EEG Application for an EEG headset, the “dwell-time” (section 3.3.4) that is used by the Eye-track Application and Head-track Application.

Device deceleration rate is used by the Eye-track Application for smoothing the transition path of the eye-gaze (section 3.3.4).

Image field expansion size is used by the Eye-track Application for expanding the object of interest with an invisible field (section 3.3.4).

In addition, a switch is also provided to the EEG Application to enable marker synchronisation with the EEG headset so that markers can be integrated with the EEG time-series as shown in Figure 10 and Figure 11.

4.5. Software functional components

A system overview of the Psyborg software components can be seen in Figure 20. The system components which set the context for the running of Psyborg are shown as blue ellipses. The red rectangles show the components responsible for handling the specific tests and the black rectangles show the components for handling the specific devices.

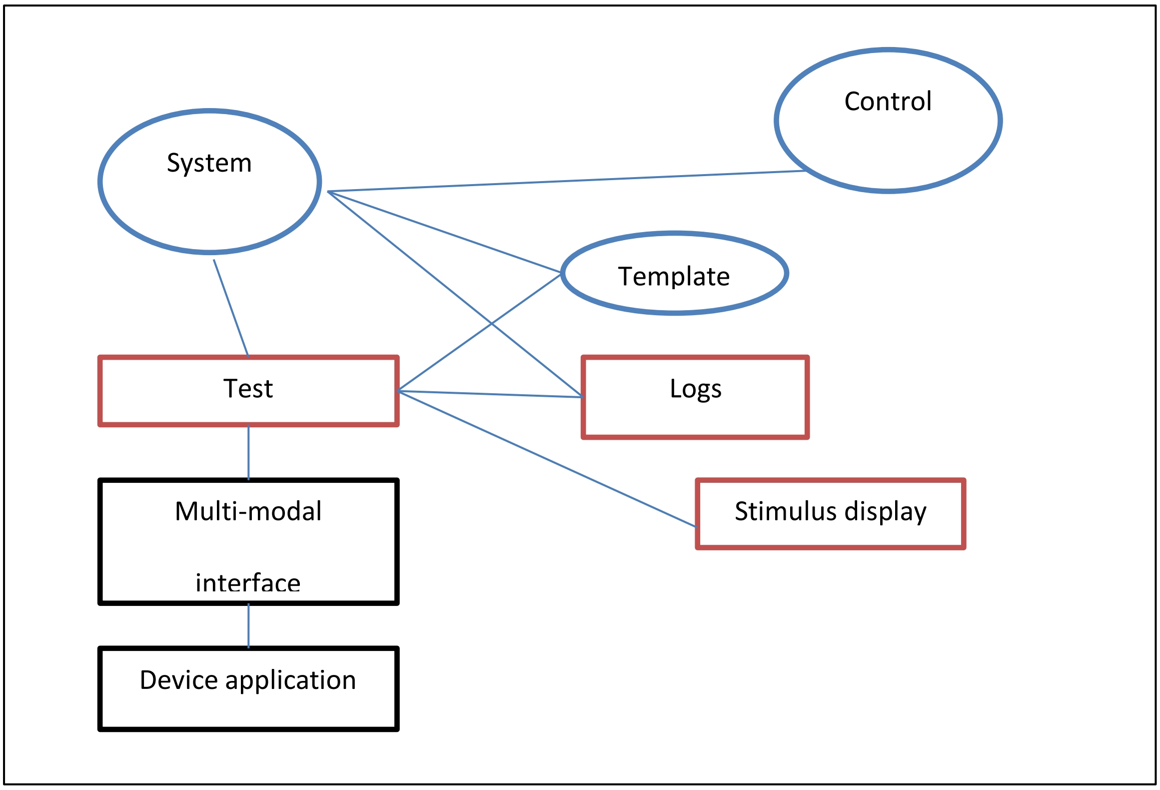


Figure 20 - Psyborg software components. The blue ellipses represent system components, the red rectangles represent test components and the black rectangles represent device components.

4.5.1. System components

The System components include the *System manager*, the *Control panel* and the *Template* components.

The *System manager* act as a staging function that takes in the parameters required to run the tests and launches the appropriate test programs. The *System manager* takes in the parameters from both the *Control Panel* for run-time configurations and the *Template* for default configuration. The *System manager* merges the parameters from both sources to form a set of parameters for the current test run. When the test results are required by the *Control Panel*, the *System manager* accesses the *Log* files which contain the results for display.

The *Control panel* provides the display view of the panel and the collection of the data from the panel to be passed to the *System manager*. Processed parameter values by the *System manager* are received by the *Control panel* and displayed.

The *Template* manages a system of files (XML) which contain default and run-time parameters. The default parameter file can be filled in off-line using a text editor. The run-time parameter file is set up by the *System manager* who copies the data from the default parameter file and overwrites the data with any updates from the *Control Panel*.

4.5.2. Test components

The Test components consist of the specific *Test* program (variant of Motor test, Cognitive test and Psychological test), the *Log* component that handles the storage and retrieval of the results and the *Stimulus display* component that handles the presentation of the test stimuli (images). The *Test* program needs to access its configuration parameters from *Template*.

4.5.3. Device components

The Device components consist of the specific *Device application* (EEG Application, Eye-track Application and Head-track Application) and the *Multi-modal interface* component that manages the integration of the data from the single mode devices to form a common representation that is used to present user responses to the *Test* program.

4.6. Scenario for general successful test cycle

This section describes a walkthrough of the execution of a specific *Test* program.

4.6.1. Launch of Psyborg User Interface

When the Psyborg program is activated, the user interface (*Control panel*) is launched (Figure 19). The program can run with the default configuration which is the *Motor test* using a mouse. Usually the user will fill in the specific *Test program* and *Test device* that is required. The *System manager* will set up the run-time configuration using *Template* whenever the *Control panel* fields are changed. Non-hand held devices normally require device training or calibration and the user gets a prompt when the device is selected.

4.6.2. Launch of device training or calibration program

The *Launch device training* (Figure 19) step has two purposes. Some devices require vendor software to be run to operate the device. The first purpose is to ensure that all vendor software that is required is activated. All software activated in this phase is normally left running for the entire test session. The second purpose is that devices that require training or calibration for the test session is carried out (section 4.4.2).

The *System manager* will launch the specific vendor or proprietary code required for training or calibration. This step may also be carried out manually by an expert user.

4.6.3. Launch of familiarisation program for test

When a test session starts, the participant is normally not familiar with the test environment and requires some warm up. Some *Test programs* (CATTEST, ECDT) have special familiarisation variants. These variants can be specified during configuration of the test (section 4.6.1) and launched when the *Launch test* button is pressed. After familiarisation, the program is terminated and the variant changed to the variant of the *Test program* used for the study.

The *System manager* will launch the *Test program* indicated by the user. The *Test program* will load the run-time configuration file from *Template* into the running test program to set the required parameters. The test is presented to the *Stimulus display* to format displays on the screen monitor. The *Test program* is terminated when the familiarisation exercise ends.

4.6.4. Launch of test program used for study

After familiarisation, the *Test program* used for the study is chosen for execution in the *Control panel* and launched using the *Launch test* button. Procedures used by the *Test program* are similar to that used for the familiarisation program in the previous section (4.6.3).

The *Test program* will set up the *Multi-modal* interface to constantly accept inputs from all the devices. The *Test program* enters into specific phases of control (states) which is centred on the display of stimuli and the reception of responses from the participants. The manipulation of device inputs depends on the states of the program. Device input and stimuli presentation events are time-stamped and logged (using *Log*) for subsequent display. The participant enters some streaks and errors and eventually meets the target criterion which terminates the *Test program*.

4.6.5. Display of results

Once the test is completed, the results can be viewed in the Psyborg control panel using the *Read file* button found in the *Maintenance* panel (Figure 19). A window that displays all the Psyborg result summary files allows the user to select the appropriate file (usually the most recent file) for viewing. The contents will be displayed in the *User notifications* window of the Psyborg control panel. The result summary window has a filter that selects only the result summary for display. In order to view the detailed results, the filter has to be switched off to view all files. A log file (in Microsoft Excel™ format) exists for the test summary results, the EEG headset test data, the eye-track test data, and the head-track results. As the result summary window also provides the name of the results folder, it is also possible to display the files manually as a spreadsheet for analysis.

4.7. Conclusions

One important and subtle finding that is made in the development of the test suite is that part of device property is embedded in tests. The tests always contain mechanisms such as a switch image that are in-built that require a non-hand held device to interact. This means that changes to the test which affect for example the positioning of the switch image may influence the test results by making the switch easier or more difficult to activate.

The Psyborg system was demonstrated to be capable of supporting the operational model of the study. The system next requires an exploratory pilot trial with some target participants. The exploratory trial will be described in the next chapter.

5. USER: Exploratory Pilot study with PNI users using the operational model

5.1. Introduction

The previous chapters explored the devices that will be used in this study (chapter 3) and the setup of the tests and equipment to support the operational model of the study (chapter 4). In this chapter we include the users in the user-device-test triad to verify the operational model.

The version of Psyborg used was Psyborg 1 (section 4.1) which did not have all the enhancements of Psyborg 2. Some of the Psyborg 2 enhancements were implemented as a result of lessons learnt from this pilot.

5.2. Aim

This section establishes the protocol for testing the operational model and represents the first instance where PNI users are tried out on the operational model.

5.3. Method

5.3.1. Participants

The pilot study undertaken includes 2 PNI children (Alex, Bob) with Cerebral Palsy (CP)(Cans, 2000) and 1 non-PNI child (Charlie)[2]. A small number of children were chosen in this pilot to establish viability of the experimental set-up. A couple of PNI children dropped out because of personal problems on the days of testing. This is not unusual due to the severe nature of impairments of the children which tend to demand more flexibility than can be accommodated during the test windows. Approval from both school and parents were sought under the University of Hertfordshire's ethics protocol aCOM/PG/UH/00006 for the children. Approval for the adults was

[2] Names of participants are replaced by fictitious names.

obtained under protocol aCOM/PG/UH/00023. Fictitious names have been used for all the participants in all references.

Table 5 - Rationalized ages of participants. The developmental age of the children were established using a picure vocabulary test; the results of which was expressed in years and months. Children were deemed not severely impaired if they were not wheel-chair bound and had enough motor ability to use a mouse.

Name	Gender	Actual age	Developmental age	Severely impaired
PNI				
Alex	Male	13	15:00	N
Bob	Male	12	5:04	N
Non-PNI				
Alice	Female	5	5:04	N
Charlie	Male	5	5:11	N

The ages of the participants were rationalized using the British Picture Vocabulary Scale III (BPVS III) (Dunn, 2009) to provide the Developmental age as shown in Table 5.

Requests for applications were made to the heads of a number of schools. Participants were recruited from two schools which agreed to the studies. PNI children were recruited from a special needs school for PNI children and non-PNI children from a junior infant school with pre-schoolers and children within the age of 5 to 12. Participants were chosen with the help of teachers in charge of classes delivering school curricula for pupils of the mental age of 4 to 6. One PNI participant (Alex) with good developmental attributes but still quadriplegic with speech slur was chosen to represent the more capable group of PNI children. Bob was another PNI child chosen who had developmental attributes around the age of 4 to 6. Alice and Charlie were non-PNI represented typical active children with developmental age of 4 to 6.

5.3.2. Procedure

The participants are tested in a room (located in a school) equipped with a laptop, separate screen monitor and hand-held and non-hand held devices for interaction with the software tests running on the laptop. Each participant was tested on all devices using the COMPTEST and two devices using the CATTEST (section 4.4.1). The screen monitor is arranged side to side with the laptop so that the participants with a view to the screen monitor are seated beside the researcher who has a view to the laptop. The eye-tracker is mounted below the screen of the monitor using magnetic mounts. The face-tracker uses a remote web camera enabling tracking from the monitor. The order of devices tested was arranged to minimize anticipated boredom and fatigue so that the easiest would be done first.

For both the COMPTEST and CATTEST a set of stimulus is produced from which the participant must provide a positive or negative response. The positive response is an active response which involves the actuation of a physical or virtual device. The negative response is passive requiring no action. Success terminates the trials for each device determined by 20 consecutive correct responses. Otherwise, the trials terminate after a block of 32 trials. Sessions of several blocks of trials involving different devices and tests are conducted once a week and all tests are run within 2 sessions of an hour each. Informal feedback was sought regarding fatigue, comfort levels, preferences after tests from carers and participants. Devices were run in order of increasing complexity. In general, this would involve running tests with the mouse first, followed by single mode devices and then hybrid multi-modal devices.

When the EEG headset is used, participants provide a positive response by gritting their teeth for a period identified as the "bite-time". When an eye-tracker is used, participants have to move a mouse cursor with their eyes to an image of a switch and "dwell" the cursor over the virtual switch for a period identified as the "dwell-time". When a head-tracker is used, participants have to move a mouse cursor using movement of their head to the virtual switch for the dwell-time. The hybrid EEG and eye-track device uses the eyes to move the mouse cursor but instead of a dwell-time, participants have to grit their teeth for a bite-time to indicate a positive response. In a similar way, the EEG and head-tracker used the movement of the head to move the mouse cursor and the teeth-grit to actuate the virtual switch.

The stimulus for the COMPTEST consists of either an image of a cuddly dog or scraggy cat (Figure 25). Participants have to provide a positive response when they see an image of the dog and negative response when they see and image of the cat. Participants are tested on all devices on the COMPTEST with each device forming a block with a maximum of 32 trials. Feedback is provided in the form of the sound of a bell tinkle when the virtual switch is actuated but not whether the response is correct or incorrect. Participants are familiarised with a few runs of the test to ensure they understand the test and are able to engage the devices.

The stimulus for the CATTEST consists of images of birds and fruits (Figure 35)(Quinn, Oates, & Grayson, 2004). Participants are expected to provide a positive response when they see a bird go into a bird cage consisting of 2 different birds or a fruit placed onto a fruit bowl holding a few different fruits. A negative response is required when a bird is placed in the fruit bowl or a fruit is placed in the bird cage. No feedback is provided to indicate whether the response is correct or incorrect but actuation of the virtual switch produces a click. Participants are familiarised with a different set of birds and fruits which does provide feedback if the response is correct or incorrect.

Only 2 devices were used for the CATTEST; the usual device that the participant is most familiar with (typically a mouse) and one other device which testing with COMPTEST showed as appropriate, especially if the device provided better results than the usual device.

Children are fetched from their classrooms, by the carers of the PNI children and the researcher for non-PNI children, and taken to the room assigned by the schools for the tests. The time before testing begins is taken as an opportunity for small talk to engage with the participant. The experimental protocol consisted of the participants first answering some personal details regarding name, age and specific impairments. They were then told about the tests they had to perform, what they had to do and what they would experience. They would be told that the research is testing out novel devices and that as it is new, the devices may not work properly. The devices would then be calibrated with the participant to provide optimum performance. The participant would then undergo the familiarisation exercise. Both the familiarisation exercise and the proper tests contain written instructions that would be diplayed on the screen for about 7 sec on what is expected in the test. After the test session completes, the children are asked how they thought the test session went and then taken back to their classrooms. Notes of significant observations or feedback are drafted after each test session.

5.3.3. Design

The experiment is a [(6 x 32) + (2 x 32)] within-subjects design for the maximum number of trials. 6 devices are used for COMPTEST and 2 devices are used for CATTEST. The test terminates after 20 consecutive successes. For each participant the following are components of the test (Figure 21):

- Device { mouse/switch, eye-track, head-track, EEG, EEG and eye-track, EEG and head-track }
- Block { 1 to 20 or 32 }
- Tests { COMPTEST, CATTEST }

There are 4 participants giving a total number of 4 x [(6 x 32) + (2 x 32)] trials maximum.

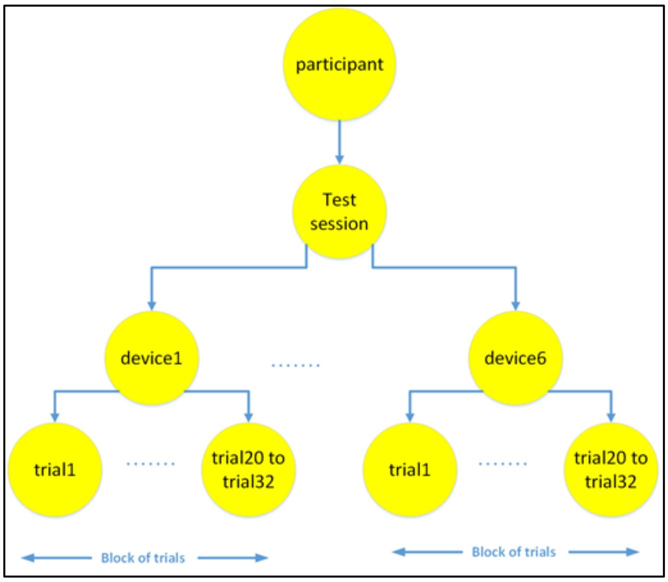


Figure 21 - Test design

5.4.Results and discussion

The tables below (Table 6, Table 7) summarises the results of the pilot study with the 4 children.

The results of each child (Alex, Bob, Alice and Charlie) are provided in separate tables; one for each test (COMPTEST and CATTEST). The results a block of trials for each device are recorded in a row. Each block of trials provide 3 outcomes; the Score, the Event Sequence and the Response time. The Score provides an average number of successes, the Event Sequence (described in detail below) is a history of successes and failures and the Response time (R. Time) is the duration before a positive response is perceived from the device.

An asterisk means that the test was not performed due to lack of time. NA indicates that the outcome was Not Applicable as the participant could not engage with the device.

5.4.1. Notation

The evaluation for this study has a requirement to finding a device user-pairing that guarantees the ability to reliably translate the decisions of the user for consecutive runs that exceeds 20 trials following the criteria used in ECDT. In this case, it was found that it was more instructive to have an overview of the events (referred to as the Event Sequence) within the block of trials. The Event Sequence is represented compactly in the form of a list notation demonstrated by the table below (Table 8). The notation can be graphically represented in any number of ways as desired. One representation using a bar chart is shown below (Figure 22) where the negative scale is used to indicate the number of errors. In general, the better device is indicated by the higher maximum value.

Table 6 - Competence Test (COMPTEST). Developmental ages of participants are given in years (brackets). Each block of trials provide 3 outcomes; the Score, the Event Sequence(Events) and the Response time(R.Time). The Score provides an average number of successes, the Event Sequence (described in Table 8) is a history of successes and failures and the Response time (R.Time) is the duration before a positive response is perceived from the device. An asterisk means that the test was not performed due to lack of time. NA indicates that the outcome was Not Applicable as the participant could not engage with the device.

Name	Alex (15 years)			Bob (5 years)			Alice (5 years)			Charlie (5 years)		
Device	*Score*	*R.Time*	*Events*	*Score*	*R.Time*	*Events*	*Score*	*R.Time*	*Events*	*Score*	*R.Time*	*Events*
Mouse	100%	0.62s	{20}	100%	1.79s	{20}	100%	1.82	{20}	100%	1.15s	{20}
EEG	100%	4.05s	{20}	78%	6.70s	B1	81%	1.87	C1	81%	1.82s	D1
Eye	100%	1.69s	{20}	NA	NA	{0}	91%	1.71	C2	NA	NA	{0}
Head	100%	2.21s	{20}	NA	NA	{0}	100%	3.50	{20}	NA	NA	{0}
EEG+E	90%	4.54s	A1	NA	NA	{0}	*	*	*	NA	NA	{0}
EEG+H	*	*	*	NA	NA	{0}	NA	NA	NA	84%	10.14s	D2

Name	*Device*	*Events*	*Event sequence*
Alex	EEG+E	A1	{6,+x,14,+x,6,+x,3}
Bob	EEG	B1	{4,-x,2,-x,7,+x2,2,+x,7,-x,3,-x}
Alice	EEG	C1	{ 1, -x1, 15, -x2, 1, -x1, 6, -x2, 3}
Alice	Eye	C2	{ 13, +x2, 1, -x1, 15 }
Charlie	EEG	D1	{8,-x,1,+x,9,+x,7,-x,+x2,1}
Charlie	EEG+H	D2	{9,+x,1,+x,4,+x,7,+x,5,+x}

Table 7 - Categorization Test (CATTEST). Two devices were chosen from the competence test (COMPTEST) trials to be used in the CATTEST. The ones with the best results were usually chosen. However for Alex and Alice, they appeared to be good at all the devices and devices that they did not have the opportunity to use in the COMPTEST due to problems with setting up the equipment were used instead.

Alex (15 years)				Bob (5 years)				Alice (5 years)				Charlie (5 years)			
Device	*Score*	*RTime*	*Events*	*Device*	*Score*	*RTime*	*Events*	*Device*	*Score*	*RTime*	*Events*	*Device*	*Score*	*RTime*	*Events*
Mouse	96%	0.95s	A1	Mouse	78%	2.20s	B1	Mouse	96%	2.93	C1	Mouse	93%	1.37s	D1
EEG+H	100%	3.10s	{20}	EEG	*	*	*	EEG+E	100%	4.46	{20}	EEG+H	90%	4.28s	D2

Name	*Device*	*Events*	*Event sequence*
Alex	Mouse	A1	{18,-x,12}
Bob	Mouse	B1	{4,-x,2,-x,7,+x2,2,+x,7,-x,3,-x }
Alice	Mouse	C1	{18,+x,12}
Charlie	Mouse	D1	{11,+x,18 +x,1}
Charlie	EEG+H	D2	{-x,6,-x,16,-x,7}

Table 8 - Notation used to capture the scores in a block of trials

Key	**Notation description**
0	Unable to engage device
<number>	Number of consecutive successes in a block of trials.
<+/->x<number>	x indicates an error in a block of trials. <number> indicates the number of consecutive errors in a block of trials. <+/-> indicates if the error is a false response to a positive or negative stimulus

5.4.2. Description

Results from the Event Sequence can be grouped roughly into 4 patterns: high confidence in the user-device pairing, potential for improvement in the use of the device, low engagement with the device, and unable to engage with the device as shown in the table below (Table 9). This grouping is not a completely accurate representation but is acceptable as a general rule.

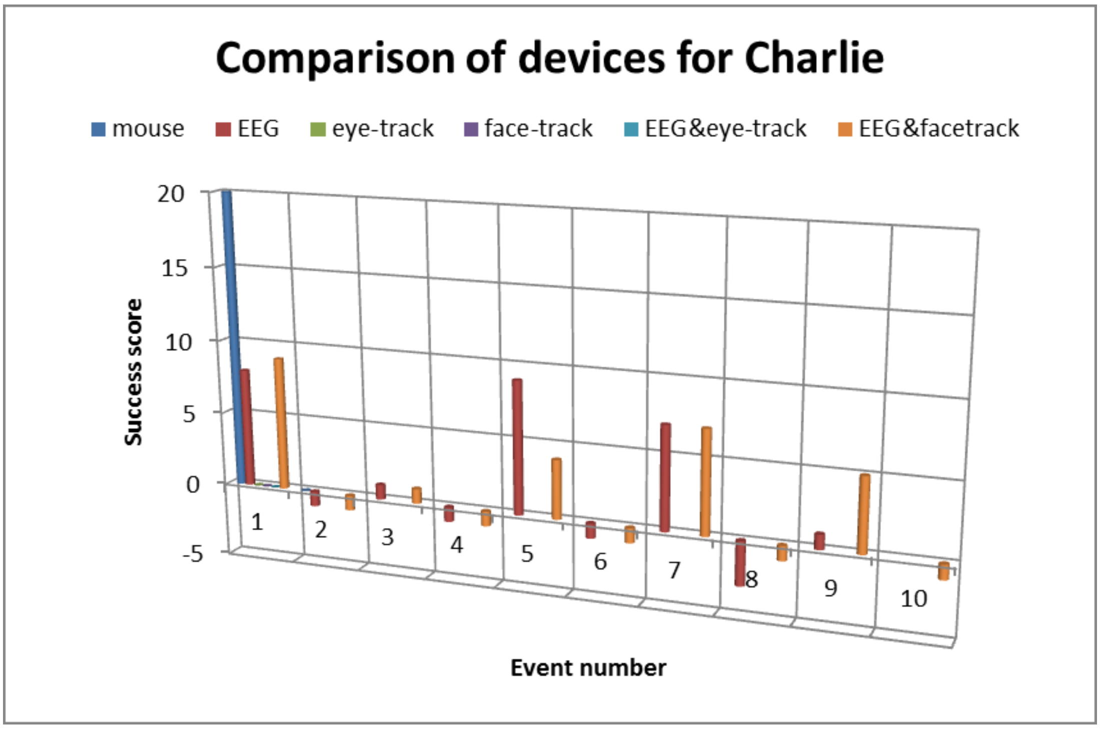


Figure 22 - Graphical representation of Event Sequence

Table 9 - Patterns of Event Sequences

Event Sequence pattern	Interpretation
20	High confidence in the user-device pairing. Contains 20 consecutive successes.
5 or more consecutive runs in 32	Potential for improvement in the use of the device.
Less than 5 consecutive runs in 32	Low engagement with the device.
0	Unable to engage with the device
5 or more consecutive errors in 32	Possible misinterpretation of instructions

The results of each child based on the tables at the beginning of this section (Table 6, Table 7) are discussed below.

Alex

This is a child with CP. Alex has all 4 limbs affected but is able to use a mouse with his left hand efficiently. Alex was selected to enable a comparison with a participant with cognitive impairment but is developed enough (15 years) to engage easily with all the devices and also offer reliable feedback. Looking at his Event Sequence, Alex had perfect scores with most devices and very strong levels of engagement with others. The score for the EEG device was initially poor. It was noticed that the EEG headset was rejecting his positive responses as noise due to the exaggerated muscular movements he had to make. The software was modified to accept the noisy responses and software developed to debounce spurious noise signals had the debounce-time extended to cater for spurious signals. A re-test was done where Alex subsequently got a perfect score. The perfect scores obtained by all the children are indicative that the user-device pairing is a reliable one and that the criterion is a reasonable grouping.

Alex was unavailable for testing using the EEG and head-tracker hybrid on the Competence Test but achieved a perfect score using the same device on the Categorization test.

Alex indicated that between responses (using bites) for the EEG headset, he had to take time to swallow which may have caused spurious responses. Excessive salivation is sometimes a problem in CP (Erasmus et al., 2009) probably due to muscular impairments. This and observations from Bacchus prompted software changes to provide configurable changes to the inter-stimulus pause per participant.

Bob

This is a child with CP with one side of the body affected (right hemiplegia) but is able to use a mouse adequately with his left hand. Bob was selected to enable comparison with a pre-school child (5 years). Children at that age were observed to be still developing cognitively. In order to engage the new devices, it is required to have the specific cognitive ability necessary for new motor skills and a degree of focussed engagement. The Event Sequence shows that Bob had a perfect score in the Competence Test with a mouse. The Response time in comparison with the others was fairly long. This starts to demonstrate that with a good user-device pairing a perfect score on COMPTEST can be expected. COMPTEST had a low enough cognitive load to demonstrate just motor skill.

Bob was unable to engage with most other devices. The need to focus long enough to get a good eye calibration was one problem. The need to coordinate eye or head movement to control the movement of a mouse cursor was another. The only other device that Bob could engage was the EEG headset. After receiving positive feedback for a successful response to a stimulus, Bob would express his jubilation for a period which sometimes extended into the next stimulus display and provide a false response. In addition, a stimulus which required a negative response would cause Bob to rest and look down as a ritual which at times also provided a false response. An extension to the inter-stimulus duration had to be made. Bob was however able to make a couple of 7 consecutive successful responses which demonstrates a fair amount of success by the user-device pairing.

In comparison, Bob was uncertain about responses to the CATTEST and did not achieve a high number of consecutive successes using the mouse. The same Event Sequence pattern was found when Bob used the EEG head-set on the motor-skills test. In this case, the problem was not the device but the cognitive task posed by the test.

There was not enough time to test Bob on the CATTEST using the EEG headset.

Alice

This is a non-PNI child. Although Alice was a pre-schooler of a few months she exhibited extremely good skills in coordination with the novel devices.

Alice did not use the EEG and eye-tracker in the competence test (COMPTEST) because of problems setting up the equipment but had the opportunity to do so for the CATTEST where eshe showed good coordination skills by doing the test perfectly.

Alice showed that using a 6s presentation time for the EEG and head-tracker multi-modal device was too short and the presentation time was adjusted to 15s.

Charlie

This is a non-PNI child. Charlie was chosen to enable comparison with a PNI child of similar developmental age (5 years). Charlie was an active pre-schooler of a few months. Children at that age are just getting used to sitting at a task with an adult providing instruction. Charlie had a perfect score using a mouse on the COMPTEST. On the CATTEST, Charlie had a strong score using the mouse.

Charlie was initially unable to engage with the head-tracker because the skill required to co-ordinate the movement of the head and the mouse cursor was a challenge. Charlie provided verbal feedback that the task was too hard. Charlie was able to overcome this barrier in a subsequent block of trials with the hybrid EEG and head-tracker device. There is a problem with the head-tracker being unable to track the nominated point on the head as children tend to move their heads too quickly.

Charlie was initially unable to engage with the eye-tracker because the eye tracker could not track his eyes. Later trials on other tests (not described) indicated that the erratic movements of the cursor due to ocular fluctuations of the eye was confusing Charlie. The software was changed to reduce the fluctuations of the cursor. The movement of the cursor was retarded so that it would only move a fraction of the way to its destination. This acts like a filter that reduces the jitter. However, this would mean that when the target is approached, the cursor would slow down drastically and appear to be unable to reach the destination. The software was then changed to allow less retardation for small movements of the cursor by using larger fractions. Charlie was able to engage the eye-tracker after these changes.

The number of trials was initially set to 64 for each test session with a device (Figure 21). This found to be too long as the Charlie kept enquiring for test completion. If the test was easy, completion would be early but if errors interrupted the consecutive successes required, each device would take 64 trial runs. The number of trials was acceptable when it was reduced to 32.

Some exploration was also done for the legacy test (ECDT). In legacy testing, no limit on number of trials was imposed. It was found that 64 trials was also too long for ECDT especially for novel devices as the participant would get frustrated as they know the solution but could not get the device to get the past the target number of consecutive successes. It was found that 32 test trials was too short for ECDT as the participant may have been on the verge of getting the right solution. A compromise of 48 trials were used which was the mean number of trials for completion in the legacy testing.

5.5.Conclusion

The pilot study suggests that cognitive tests can be influenced by the input device (especially new devices), and also that it is not possible to determine whether the results from the cognitive tests are due to a motor-skill error, or due to problem-solving cognitive error. This difficulty occurs for both PNI and non-PNI children, but is more relevant for PNI children. There are impairments for some PNI children that are severe enough such that current solutions for computer access are still unreliable. For participants with cognitive impairments, some cognitive tests will produce results that are indistinguishable whether they originate as a motor-skill error or a problem-solving cognitive error. It would be valuable for future cognitive tests to incorporate a distinct motor-skills test as a simple pre-cognitive test to help separate out the motor skill and cognitive errors in the cognitive tests itself.

The pilot study was also a means to explore configuration of parameters to allow continuation to the main study described in the next chapter. For the EEG device the “think-time” filter described in section 3.2.6 was extended to allow for “noisy” rejections by the Emotiv software.

The configurable inter-stimulus pause indicated in section 4.4.3 was introduced as a result of feedback and testing from PNI children that indicated a pause was required for excessive salivation and uncontrolled emotional displays. For the use of an eye-tracker, a configurable smoothing parameter was introduced to control the smoothing function (section 3.3.4) for the eye cursor control as a result of testing with Charlie. The number of trials was shortened for each test session as a result of indications of fatigue from the children.

6. MAIN PILOT STUDY: Overview

6.1. Introduction

This chapter describes the common elements and overall research design for the three main tests that have been conducted for this thesis and acts as an overview to the tests. The three main tests are then described in detail in subsequent chapters (Chapter 0, 8, 9). The tests investigate the performance of PNI children using novel devices to undertake the cognitive tests. Consolidation of the findings then takes place in Chapter 10. A cross-test analysis of the findings and the integration to initial models of understanding is provided.

This chapter starts with an initial introduction which includes an outline of the research questions, definitions and terms used, a summary of the devices used in the tests, a summary of the tests discussing the relationship between the main study and the three tests in terms of a user-device-test triad and a description of the participants used in all three tests.

Next this chapter provides an overview to the methods used in the research design by describing methods for ranking, comparing and detecting interactions between devices (section 6.2).

Finally, the structure of the test descriptions is represented by a standard template that has been developed to describe each test (section 6.4).

6.1.1. Research questions

The main study has the aim of answering all the research questions:

Number	Research question
RQ1	RQ1. What device helps a child successfully attempt a cognitive test?
	RQ 1.1. Does a PNI child have sufficient motor control to use the device?
	RQ 1.2. Does a PNI child have sufficient motor control to use the device and undertake a light cognitive test at the same time?
	RQ 1.3.Does a PNI child have sufficient motor control to use the device and undertake a learning test at the same time?

6.1.2. Definitions

Some terms are often used in the main study and as a point of convenience they are grouped into areas and explained here. The definitions are grouped into areas discussing study and test design, references used to develop a new measure based on streaks, terms used to discuss cognitive concepts and conceptual terms formulated as a result of this study.

6.1.2.1. Terms relating to study design

Theoretical model	The study starts off with an initial theoretical model (section 2.3.6) which projects the complexity of the devices in order to fit the users of the device.
Operational model	The operational model of the study is a reference to the study implementation which supports the verification of the theoretical model. The model consists of experimental procedures and infra-structure set up (section 2.3.6).
Empirical model	The empirical model is the empirical results from an execution of the operational model that proves, disproves, modifies or presents anomalies to the theoretical model (section 10.5).

6.1.2.2. Terms relating to test design

Test suite	This study is composed of three tests (COMPTEST, CATTEST, ECDT). The target test is ECDT. The other two tests (COMPTEST and CATTEST) are pre-tests for the target test.
Test session	Each test in the test *suite* of three will have a particular organization of a test session (Figure 23) in which the test session for each test participant is sub-divided. A test session may either be sub-divided into *blocks* or *sequences* of *blocks*. Each *block* consists of a number of *trials*.
Sequence	A sequence in the context of a test session is a sub-division of the test *session*

	which consists of a separate test objective. Each sequence will comprise of a *block* of *trials*. Test sequences are only used in ECDT and there are only two sequences; one for assessment of learning and the other for assessment of transfer of learning.
Block	A block in the context of a test session refers to a number of *trials* that makes up the test session. A block is a sub-division of a test *session* or part of a *sequence*. Each block may be run on a per device basis or on a test *sequence* basis. The number of *trials* for each block depends on the specific test requirements.
Trial	A test *session* requires the execution of a *block* of trials, each of which requires a YES/NO response.

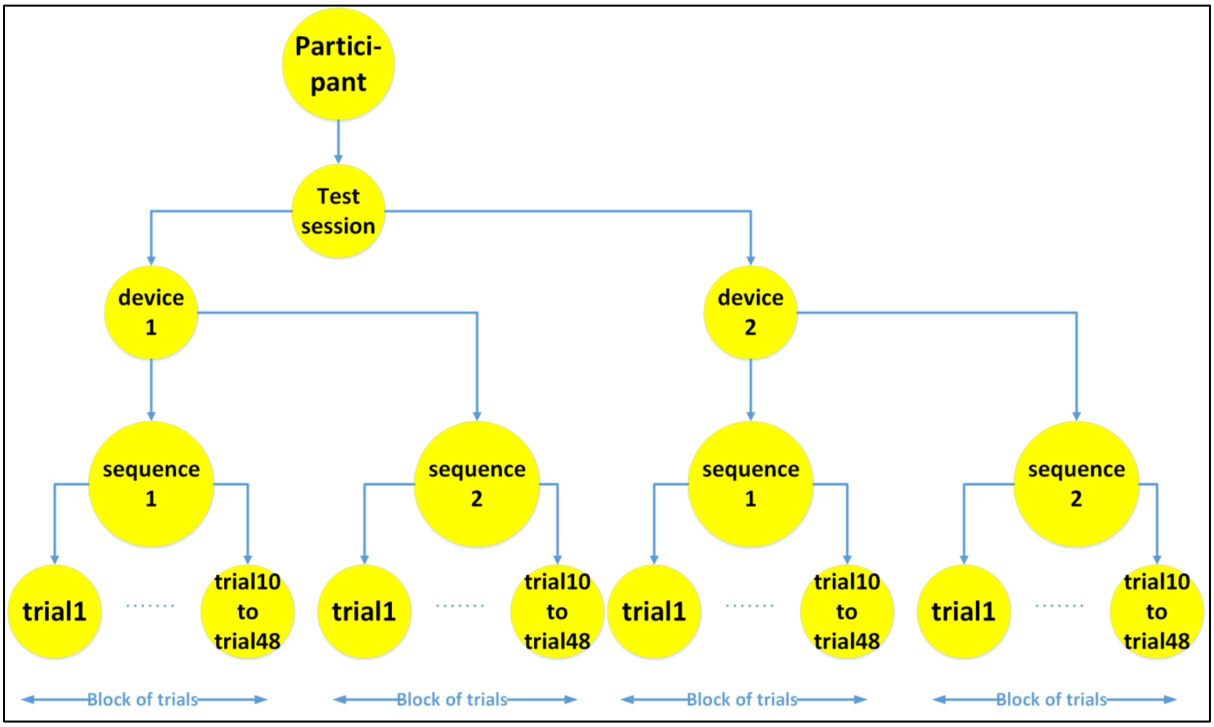


Figure 23 – Example of a test design hierarchy

6.1.2.3. Terms relating to measures used in study

Streak	Streaks are consecutive binary event sequences of the same type. Success streaks refer to a consecutive sequence of success events, of two or more. Maximum-streak is the highest streak size attained in a test *session*, each requiring a binary (YES/NO) response. Streak sum is the sum of all *intentional* streaks in a test session.
Intention	The term intention is used in this study to mean that the response made in a test is not *noise* or a reflex action but one arising from a conscious decision (Libet et al., 1993; Ouellette & Wood, 1998).
Noise	The term noise has been used in this study to mean the results from tests that provide no information due to the high probability that the result is contaminated by the effects of impairments such as involuntary motor actions. Results that are valid are deemed *intentional*.

6.1.2.4. Terms relating to cognitive concepts used in study

Cognitive complexity	Cognitive complexity (Kieras & Santoro, 2004) is used to provide an indication of different levels of sophistication of a task. There are in general at least two tasks in every test in this study. One task is to handle the requirements of the test and the other to handle the device. Both device and test have varying degrees of complexity. Each task requires the brain to generate some *cognitive effort* to respond to the *cognitive load* for specific areas of *attention*.
Cognitive load	Cognitive load (Paas, Tuovinen, Tabbers, & Van Gerven, 2003) is used to provide an indication of the amount of work the specific functions of the brain has to do to perform a task. The cognitive load is related to the

	cognitive *complexity*. The higher the *complexity* of a task the higher the load generated by the task. The *effort* devoted to handle the different functions of a load is referred to as *cognitive effort*. The *effort* provides the *attention* devoted by those specific functions.
Cognitive effort	Cognitive effort (Longo & Barrett, 2010) is used to provide an indication of the energy devoted by the brain to handle the *cognitive load*. Cognitive effort is dedicated to provide *attention* to elements of a task. The brain may devote a lower amount of effort than that required to handle specific *attention* areas of the task which results in task incompletion.
Attention	A task requires a certain amount of *cognitive effort*. The *effort* is devoted to specific areas of attention for a particular length of time. Attention (Deco & Rolls, 2005; Folk, Remington, & Johnston, 1992; Folk, Remington, & Wright, 1994) is a psychological state provided by functions of the brain for a certain length of time that requires *cognitive effort*.

6.1.2.5. Terms relating to other conceptual elements used in the study

Signing	Signing refers to a conceptual device implemented for this study. The device consists of a participant providing direct communication using gestures or voice to an interpreter who uses a mouse to provide a response to a test. One of the main objectives of this process is to establish that the participant has poor results from an actual device as a consequence of the device and not the test.
User, device, test triad	The user, device and test triad are mentioned in the thesis to indicate that the three entities are strongly related. They interface with each other and each entity is an independent variable. Changing the properties of any one

entity affects the outcome of the system and possibly affects the other two entities.

6.1.3. Summary of devices used in the tests

Conceptually, two types of devices are employed in this study. The first type is selection. This device basically provides a binary type YES/NO responses. The second type is pointing. A pointing device provides spatial information about the position of the device allowing a mapping to be made with positions on a screen. A pointing device also has a selection capability.

Devices can be fitted to PNI children using two strategies: (1) the device is able to mitigate impairment (2) the device uses a part of the body that has little or no impairment. A range of devices were chosen based on the latter strategy although adaptations are made to reduce the impacts of impairments. The number of types of devices used in any study will be limited by the available time. Also, Birbaumer (Birbaumer, 2006) indicates that eye control and external sphincter muscle control are usually the two last areas of control lost for motor-neurone patients. In this case, a representative of eye control, head control and face control was thus chosen. Other technologies are currently available for further choice which presents opportunities for widening the study.

The following devices were chosen for this study:

1) Switch/mouse
2) A neuro-headset (Emotiv, 2013) based on EEG used to detect facial artefacts (Barreto, Scargle, & Adjouadi, 2000; Reyes et al., 2012)
3) An eye-tracker (Tobii Technology, 2013; Ware & Mikaelian, 1986)
4) A face/head-tracker (Betke et al., 2002; Boston College, 2013)
5) A hybrid device consisting of both EEG and eye-tracking (implemented for evaluation)
6) A hybrid device consisting of both EEG and head-tracking (implemented for evaluation)
7) Signing (conceptual device based on direct communication)

6.1.3.1. Switch/mouse

In this study, the input device that is most commonly used by the participant is included in the tests as a comparison. In most cases, this would be a switch or a mouse. Although the mouse has a capability as a pointing device, only its selection capability is used.

The switch is basically a big button for PNI children with good enough control to hit the switch but unable to use the fingers to click a gripped mouse.

A variant of the mouse exists called the finger-on-palm mouse that only has selection capability. Some PNI children may be able to curl their fingers to form a fist but are unable to use the fingers to click a gripped mouse. The finger-on-palm mouse uses a contact strapped to the palm and finger so that a click can be produced when the finger touches the palm.

6.1.3.2. EEG neuro-headset

The EEG neuro-headset is a device that is trained to map certain EEG patterns to responses. In this study, the EEG headset is trained to recognise bites. To be more precise the EEG headset is trained to map the electrical interference (artefacts) to EEG caused by muscular activities that correspond to bites. Any muscular activity that uses muscles common to that used in a bite will also map to the same response. It is possible that the EEG headset will also respond to the flicker of an eye-brow in addition to a bite when the former action generates a strong enough indication in some bite muscles (e.g. temporalis). Training is carried out for each participant at the start of each test run.

The EEG headset is used as a selection device only. The headset is trained to recognise a bite as a YES and a neutral facial expression as a NO.

In this study, selection is only valid after a fixed duration (bite-time) and if the bites exceed a particular magnitude (unit proprietary to vendor).

6.1.3.3. Eye-tracker

The eye-tracker is calibrated to map the fixation gaze of both eyes to a spot on the screen monitor. Movement of the eye gaze is used to move the screen cursor. If the eye-tracker loses track of the eyes, the screen cursor holds its last known position.

The calibration process requires the eyes to be fixated on various points on the screen monitor. A spot is used to attract the attention of the gaze. The spot moves to the point for calibration and a shrinking black spot appears in the background of the initial spot to focus attention on the calibration point. The spot can be configured to move at three different speeds. The slowest speed is designed for infant calibration and uses a teddy bear instead of a spot. The teddy bear appears at the different fixation points under manual control. Once in position, the teddy bear shakes and rattles to attract the infant's attention. All calibration is done using 9 points. The eye-tracker is affected by the use of some glasses and participants are asked to remove them unless it affects the tests. Calibration is carried out for each participant when they start the test runs.

The eye-tracker is a point and select tool. In this study, pointing is required on an image of a switch. Selection is based on a fixed fixation time (dwell-time) which can be configured.

6.1.3.4. Head-tracker

The head-tracker is calibrated to map the position of a reference point on the face (the nose) to a spot on the screen monitor. Movement of the nose is used to move the screen cursor. If the movements are too violent, the head-tracker loses the reference point on the face and the reference point is moved to some other location on the head. A reference point is always present and violent movements can at times improve on a reference point which is not in a good position.

The head-tracker is a point and select tool. In this study, pointing is required on an image of a switch. Selection is based on a fixed fixation time (dwell-time) which can be configured.

6.1.3.5. EEG and eye-tracking hybrid device

The EEG and eye-tracking hybrid device is a proprietary implementation for this study. The eye-tracker is used only as a pointing device and the EEG neuro-headset is used as a selection device. In this study, pointing is required on an image of a switch. Selection is based on using the EEG headset to detect bites. Selection is only valid after a fixed duration (bite-time) and if the bites exceed a particular magnitude (unit proprietary to vendor).

6.1.3.6. EEG and head-tracking hybrid device

The EEG and head-tracking hybrid device is a proprietary implementation for this study. The head-tracker is used only as a pointing device and the EEG neuro-headset is used as a selection device. In this study, pointing is required on an image of a switch. Selection is based on using the EEG headset to detect bites. Selection is only valid after a fixed duration (bite-time) and if the bites exceed a particular magnitude (unit proprietary to vendor).

6.1.3.7. Signing

This study also considers Signing as a method of input. Signing provides a contrast to bio-modal inputs and the typical physical inputs. For our research, Signing refers to a child who communicates using gestures or speech to an interpreter. The gesture and or speech acts as the child's response to a test, and this response is then entered as a mouse input via the interpreter. The interpreter may also be relaying the response via speech to a researcher.

The interpreter used is a carer who was able to pick up on subtleties of a communication with a PNI child. Such subtleties include gestures that are additionally used that are not a part of the recognised repertoire established for that child.

Signing was deemed to impose the least cognitive load for manipulating a device. The child may have had years of using and developing it to use the parts of the body which were functionally capable for the purpose.

As such, Signing was viewed to be one of the best practical devices available albeit a conceptual one.

Signing is a valuable contrasting ‘input’ because a child that successfully completes a test using the Signing ‘device’ most likely has the cognitive ability for being tested. Where that child then has difficulties completing the test successfully with a bio-modal device helps to expose the positive or negative impact of that device.

The replacement of Signing using a physical device loses some of those advantages but also has the associated disadvantages. Carers develop strong ties with the participant and subconsciously may provide responses that bias the response towards a more positive outcome. The subtleties of Signing would be viewed as an opportunity for further research.

The view of Signing as a conceptual device allows for the demonstration that a physical device exists that may be more efficient and a better fit for a PNI child (section 2.3.5).The physical device implemented will depend on the communication mode used by the child. The more able-bodied children may show a preference for using voice or touch-screen devices and the less able-bodied may prefer gestures that do not require hand control (Biswas & Basu, 2011; Ren et al., 2011).

Signing can either employ selection or pointing and selection. Normally there is an independent gesture for YES and NO but the participant may choose to point to the screen.

6.1.4. Summary of tests

6.1.4.1. Competence test (COMPTEST)

This is the first of a series of three tests in a test suite. This test can be described as a very simple motor skills test. The test consists of checking the ability of a PNI child to provide basic YES/NO responses from various input devices. The devices are ranked according to the maximum-streak size.

This test is set up to answer the research question “RQ1.1. Does a PNI child have sufficient motor control to use the device?”

6.1.4.2. *Categorisation Test (CATTEST)*

This is the second of a series of three tests in a test suite. This test is a categorisation test which imposes a light cognitive load. This test checks the ability of a PNI child to provide basic YES/NO responses from an input device such as a mouse under a light cognitive load.

This test is set up to answer the research question “RQ1.2. Does a PNI child have sufficient motor control to use the device and undertake a light cognitive test at the same time?”

6.1.4.3. *Early Concept Development Test (ECDT)*

This is the third of a series of three tests in a test suite. This test is a legacy test for testing if a PNI child has the ability to learn and to transfer that learning. The implementation was extended to provide the capability to recognise inputs from novel devices other than the mouse.

This test is set up to answer the research question “RQ1.3. Does a PNI child have sufficient motor control to use the device and undertake a learning test at the same time?”

6.1.5. The user, device, test triad

6.1.5.1. *Order of tests*

An order of execution was established for the three tests in the test suite because of a dependency chain set up between the three tests. The first test (COMPTEST) evaluates various input devices that are candidates to be used for the last test. Out of this evaluation, two devices will be chosen to be carried forwards for the execution of the other two tests. The second test (CATTEST) verifies that the two devices from the first test are still valid under a light cognitive load. The final test verifies that the two devices are appropriate for the target learning test. Each test will have a chapter dedicated to its description in this thesis (Chapters 0, 8, 9).

6.1.5.2. Cognitive load variance due to device demands

The left vertical axis in Figure 24 shows the device as an independent variable. Each device imposes a cognitive load on the user. The test result demonstrates the device ability in the form of maximum streak sizes if both the user and the test are kept constant. In general we expect the demonstrated device ability to be high if cognitive load is low and vice versa. The cognitive load imposed on the user depends on the latent skill of the user. The higher the skills level of the user with a particular device, the lower the imposed cognitive load.

The devices used for the test are depicted on the right in Figure 24 and arranged in order of complexity determined by the initial theoretical model for device ability (section 2.3.6).

6.1.5.3. Cognitive load variance due to test demands

The horizontal axis in Figure 24 shows the test complexity as an independent variable. Each test imposes a cognitive load on the user. The tests are arranged in order of increasing cognitive load from left to right (COMPTEST, CATTEST, ECDT). The test result demonstrates the test ability in the form of maximum streak sizes if both the user and the device are kept constant.

6.1.5.4. Variance of user ability

The inward axis in Figure 24 shows the user ability as an independent variable. Each group of users has a general level of ability. The user groups are arranged in general order of increasing ability starting from the bottom (PNI children, non-PNI children, adults). The mean test result demonstrates the user ability in the form of maximum streak sizes if both the device and the test are kept constant. For the purpose of comparisons between children, the mental age was normalized using a picture vocabulary test. A similar form of normalization is envisaged if PNI adults were to be considered.

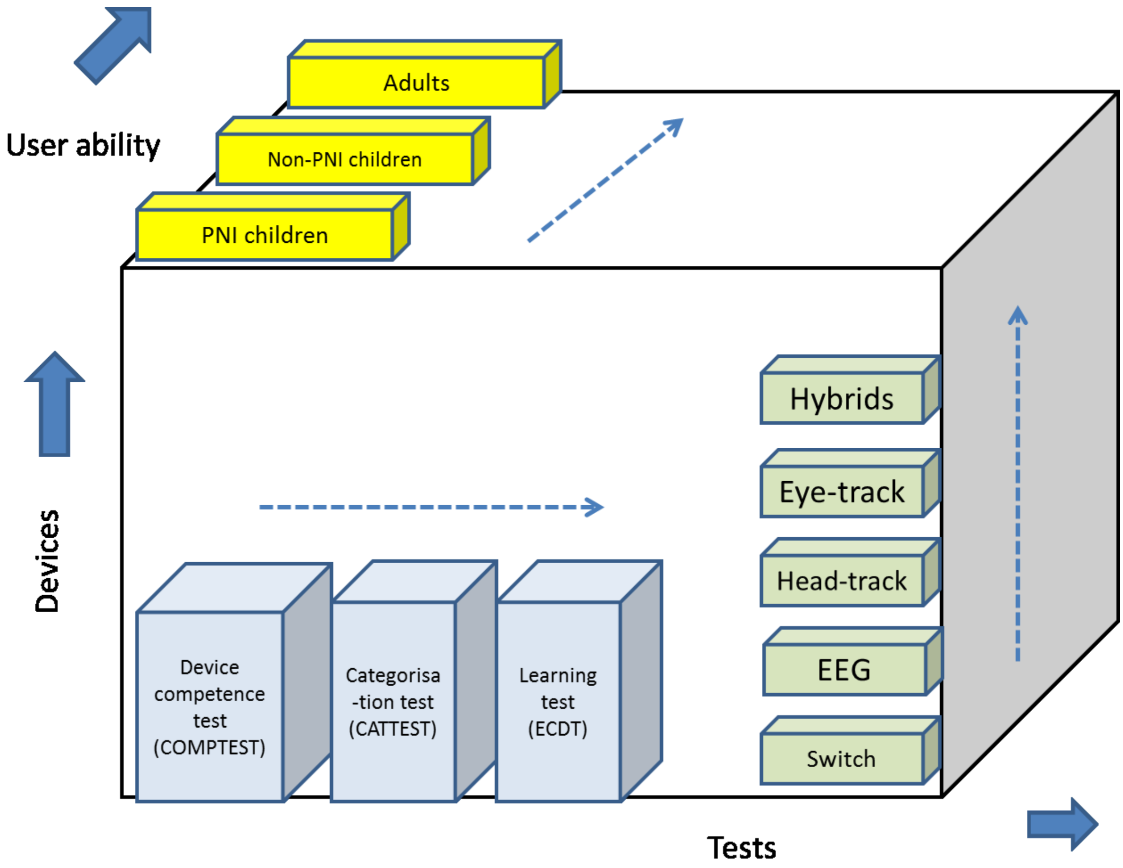


Figure 24 - Study design showing user ability, device, and tests as independent variables

6.1.6. Participants

Seven PNI children, 6 non-PNI children and 5 adults were tested as shown in Table 10. The number of participants that recruitment achieved for children were fairly small and thus no attempt was made to balance the numbers between groups of participants. Although it was possible to recruit larger numbers of adults, it was anticipated that there would be little variation in results from initial formative testing. The plan would be to provide more of a hand lens approach to the data then a statistical approach due to the small participant size. Requests for applications were made to the heads of a number of schools. Participants were recruited from two schools which agreed to the studies. PNI children were recruited from a special needs school for PNI children and non-PNI children from a junior infant school with pre-schoolers and children within the age of 5 to 12. Participants were chosen with the help of teachers in charge of classes delivering school curricula for pupils of the mental age of 4 to 6. One PNI participant (Geronimo) with good developmental attributes but still quadriplegic with speech slur was chosen to represent the more capable group of PNI children. Other PNI children chosen had developmental attributes around the age of 4 to 6. Approval from both school and parents were sought under the University of Hertfordshire's ethics protocol aCOM/PG/UH/00006 for the over 5 children and adults. Approval for the under 5 children was obtained under protocol aCOM/PG/UH/00023. Fictitious names have been used for all the participants in all references.

Table 10 - Participants for main study. Mental age was established using the British Picture Vocabulary Scale (BPVS III). This scale rates the results in years and months.

PNI children					Non-PNI children				Adults		
Name	***Gender***	***Severely impaired***	***Age***	***Mental age(years: months)***	**Name**	***Gender***	***Age***	***Mental age(years: months)***	**Name**	***Gender***	***Age***
Apollo	Male	Yes	14	04:10	Achilles	Male	5	05:11	Aladdin	Male	23
Bacchus	Male	No	12	04:07	Clementine	Female	6	06:05	Anubis	Male	26
Baldr	Male	No	15	< 04:00	Isis	Female	5	05:11	Atlas	Male	31
Geronimo	Male	No	13	11:03	Lactose	Female	6	05:08	Cupid	Male	25
Lavender	Female	Yes	12	< 04:00	Mercury	Male	6	06:06	Damocles	Male	31
Nimrod	Male	No	13	04:05	Sinbad	Male	4	04:10			
Thor	Male	Yes	12	07:03							

Table 10 shows the details of the participants. The table is divided into three sections; the one on the left hand side detail the PNI children, the one in the middle the non-PNI children, and the one on the right hand side, the adults. The adults and the non-PNI children are controls. Five of the PNI children have various form of cerebral palsy (CP) (section 2.2), one has methotrexate leucoencephalopathy (Ch'ien et al., 1981) and one has septo-optic dysplasia with autistic spectrum disorder (Webb & Dattani, 2010). The non-PNI children are either in the reception class or have just started Year 1. The adults were recruited from the university and are all non-PNI.

The table indicates for the PNI children, three participants (Apollo, Lavender and Thor) who are severely impaired in that they are wheelchair bound, have almost no speech, and have involuntary muscular problems and weak muscular control. The remaining four suffer some impairment to a lesser degree and are not wheelchair bound.

The ages of the children were rationalized using the British Picture Vocabulary Scale III (BPVS III) (Dunn, 2009) to provide a mental age that range between 4 to 6 years old to allow meaningful comparison between the two groups of children. The rationalization was based on verbal maturity. Both Baldr and Lavender could not be assessed using BPVS as their scope exceeded the lower limits. As a result of impairments of the children, some BPVS results were best effort results as the case with Apollo and Thor. Geronimo was picked as a control participant that represented a person with CP and with mature developmental capability (mental age of 11 years old).

PNI: Apollo

Apollo is a 14 year old who has athethoid cerebral palsy. He is quadriplegic and has no ability for speech and depend on Signing to communicate. Apollo raises his right hand to indicate YES and places his hand on his left chest to indicate NO. Signing was confounded by involuntary muscular activity. He is severely impaired being wheel-chair bound and does not have good motor control. His best attainment level was Key Stage 1c in English. His communication impairments were too severe for the researcher to establish his developmental age and the results from the school taken from the summer of 2013 was used. His development age was recorded as between 3:07 to 4:10 (BPVS).

PNI: Bacchus

Bacchus is a 12 year old who has cerebral palsy. He is hemiplegic and tends to be quiet. Bacchus can get around without any problems and is weak on his right and is therefor left handed. His best attainment level was Key Stage 1a in Information and Communications Technology. He uses glasses and has no known specific visual impairments. His developmental age was 4:07 (BPVS) at the time of testing.

PNI: Baldr

Baldr is a 15 year old who has septo-optic dysplasia and displays an autistic spectrum. He is paraplegic and tends to be quiet. Baldr tends to be distant as a result of his impairments. He can get around without any problems but limps on one side. Baldr is left handed. His best attainment level was Key Stage 1a in Mathematics. He uses glasses and the researcher was not made aware of any specific visual impairments. His developmental age was less than 4:00 (BPVS) at the time of testing.

PNI: Geronimo

Geronimo is a 13 year old who has athethoid cerebral palsy. He is quadriplegic and has delayed speech but communicates readily. He has a limp but manages movement well. Geronimo is left-handed and requires glasses. His best attainment level was Key Stage 4 in Sceince. His development age was 11:03 (BPVS) at the time of testing. Geronimo is highly capable and was chosen to represent that group.

PNI: Lavender

Lavender is a 12 year old who has a form of cerebral palsy (hypoplasia). She is quadriplegic and has no ability for speech and depend on Signing to communicate. Lavender folds her arm to form a wing and flaps it to indicate YES and turns her head to indicate NO. She has a very short attention span and can be highly excitable. She is severely impaired being wheel-chair bound and does not have good motor control. Her best attainment level was P8 in English. Her developmental age was below 4:00 (BPVS).

PNI: Nimrod

Nimrod is a 13 year old who has A form of cerebral palsy (Worster drought syndrome). He has good motor control but has no ability for speech and depend on Signing to communicate. Nimrod has a rich Signing repertoire. He predominantly uses a thumbs up to indicate YES and wags his fingers to indicate NO. However, other signs may also be used. His best attainment level was P8 in English. His development age was 4:05 (BPVS). Nimrod becomes very reactive when he is out of his comfort zone which makes him unsuitable for novel devices that have to be worn until he can be persuaded to try them out.

PNI: Thor

Thor is a 12 year old who has methrotrexate leukoencephalopathy. He is quadriplegic and has no ability for speech and depend on Signing to communicate. Thor looks left to indicate YES and looks right to indicate NO. Signing was confounded by involuntary muscular activity. He sways his upper body irregularly. He is severely impaired being wheel-chair bound and does not have good motor control. His best attainment level was Key Stage 1b in English. His communication impairments were too severe for the researcher to establish his developmental age and the results from the school taken from the summer of 2013 was used. His development age was recorded as between 5:11 to 7:03 (BPVS).

non-PNI: Achilles

Achilles is an active child who gets bored easily. He is keen to explore the environment and can be impatient. He has a developmental age of 5:11 (BPVS).

non-PNI: Clementine

Clementine has good maturity. She has a developmental age of 6:05 (BPVS).

non-PNI: Isis

Isis is an active child who is quite restless and tends to fidget. She has a developmental age of 5:11 (BPVS).

non-PNI: Lactose

Lactose has good maturity. She has a developmental age of 5:08 (BPVS).

non-PNI: Mercury

Mercury is an active child who gets bored easily and can be impatient. He has a developmental age of 6:06 (BPVS).

non-PNI Sinbad

Sinbad is an active child who can be impatient. He tends to be playful. He has a developmental age of 4:10 (BPVS)

Adults

All adults tested were post-graduates from the university and tend to be very focused on their activities.

6.2.Experimental protocol

Children are fetched from their classrooms, by the carers of PNI children and by the researcher for non-PNI children, and taken to the school assigned test room. The time before testing begins is taken as an opportunity for small talk to engage with the participant.

The experimental protocol consisted of the participants first answering some personal details regarding name, age and specific impairments. They were then told about the tests they had to perform, what they had to do and what they would experience. They would be told that the research is testing out novel devices and that as it is new, the devices may not work properly. The devices would then be calibrated with the participant to provide optimum performance. The participant would then undergo the familiarisation exercise. Both the familiarisation exercise and the proper tests contain written instructions that would be diplayed on the screen for about 7 sec on what is expected in the test. After the test session completes, the children are asked how they thought the test session went and then taken back to their classrooms. Notes of significant observations or feedback are drafted after each test session.

Adults were tested in a room in the university used for development work.

6.3.Overview to the design

6.3.1. Method for ranking devices

In the overview of the three tests in the test suite (Section 6.1.4.1) it was indicated that COMPTEST was used for ranking the devices. In this section, the elements of choosing appropriate devices using COMPTEST are further discussed. The discussion covers how COMPTEST rank devices in order of difficulty.

It was noted that common measures are not designed for children with impairments. The other way of looking at it is that it is possible to design measures that would ignore the effects on the results caused by impairments. The measures are based on consecutive events (streak events (Section 6.1.2.3)). The use of streaks for ranking makes available two types of measures. The first is where the maximum streak value (Section 6.1.2.3) is used. This measure provides an indication of best attainment. A particular target is set for success which terminates the test. If the target is not reached, the value represents the best effort reached within a set window of tries. If a particular maximum is reached more than once, a ceiling is implied. The ceiling has two complementary causes. One is related to the capability of the child. The point has been reached where the child at that moment can do no better. This could be due to natural ability or the presence of impairment. The other cause is related to the difficulty presented by the device. The device has imperfections relative to the child that limit the possibility to achieve better results. The second measure is an average measure. Streaks above a particular threshold are summed and referred to as streak-sums (Section 6.1.2.3). The streak-sum can be averaged over the number of tries to provide an indication of general capability with a device.

In this study results are presented using maximum streaks as a first candidate measure. The measure is also easily traceable in an event history of a sequence of tries in a test when it is of interest to get more information in anomalous cases. The measure provides a means of ranking as it is an ordinal measure.

6.3.2. Method for comparing performance of devices

The performance of devices is viewed in terms of minimal effort required to drive the device to achieve the best possible result in the test. We estimate the performance of a device or devices in a number of ways:

- The device has continued success with the test criterion streak size of 20 in both COMPTEST and CATTEST and succeeded in ECDT.
- The device result compares with a trusted benchmark (Signing)
- Pairs of devices that have passed the COMPTEST criterion of a streak-size of 20 have similar results in CATTEST

The estimation provides devices that have the capability to be as good as the trusted bench mark or reach criterion. In such cases, the device has not prevented the user from reaching their test target or targets. The user has a good enough motor skill with the device.

A device that is able to succeed in all the tests in the test suite is a trivial case of a device with good performance.

Signing (6.1.3.7) is seen as a reliable device to be used for comparison. Signing should provide the best result in any test (within a given variance for statistical error). Signing should either outperform or equal any other device. If a device has the same results as Signing, there is high confidence that a much better result will not be achieved. The device has reached the quality of the trusted benchmark.

Devices that have succeeded in COMPTEST can be used as benchmarks for each other. If those devices produce the same results in another test like CATTEST, there is high confidence that the results are good.

6.3.3. Method for detection of interaction effects involving devices

An interaction effect is an effect that is caused by not only one independent variable but a result of a combination of independent variables only. An interaction effect which involves devices has the device acting in combination with some other independent variable (for example the user or the test) or variables to establish the result of the test.

One way of detecting an interaction effect which involves a device is to detect a significant discrepancy in results with a trusted benchmark (Signing) between two tests. The trusted benchmark should provide the best performance. If the device cannot keep up with the trusted benchmark, the device has been influenced by some other factor. The device is first checked to be able to provide the same result as Signing in COMPTEST. This may also be a device that reaches the target criterion of a streak size of 20 in COMPTEST. Next the device is checked if it can also achieve as good a result as Signing in CATTEST. (It is assumed that the device cannot outperform Signing). If the device then achieves a significantly poorer result in CATTEST than Signing, the conclusion is drawn that an interaction effect with the device has taken place.

6.4.Standard template for test description

This section provides the chapter template for subsequent chapters that describe the three tests of the main study (Chapters 0 - COMPTEST, 8 - CATTEST, 9 - ECDT). The only material that is common to all three chapters falls in the Participants section and is placed in a section in this chapter (6.1.6). Table 11 shows the table of contents of a chapter describing a test in the test suite of the main study.

Table 11 - Template for chapters covering specific tests in the main study

Section	Title	Description
1	Introduction	Provides overview of the test.
2	Aim	Points to the research questions
3	Method	Describes test method
3.1	Participants	Describes choice of participants
3.2	Procedure	Describes procedure followed during testing
3.3	Design	Describes the test structure
3.4	Data capture	Describes the notation used to capture results
4	Results	Show results and describes trends observed
5	Discussion	Discusses the trends observed
6	Conclusion	Provides summary of test

6.5.Conclusion

Having established an overview of the main study in terms of the research questions that the study answers, the general structure of the study, common elements of the study, the rationale of methods used in the study and the template for the description of the study, a description of the actual study follows with the first test in the test suite of 3.

7. MAIN PILOT STUDY: Motor-skills test (COMPTEST)

7.1. Introduction

One of the major difficulties with physically and neurologically impaired (PNI) children, particularly those children that are severely impaired, is that physical impairment makes it difficult to assess cognitive ability. For example, tests of cognitive ability usually require and assume that the child has some level of physical ability such as the ability to control a mouse in a computer-based test. The introduction of technology that may assist these children in the form of non-hand held devices brings out the issues of whether the children use those new devices to interact adequately with a cognitive test and whether those new devices have an additional impact on the test outcomes that are unrelated to the intentions of the test again. As a start, there is a requirement for a motor-skills test which checks the competency of the child with the device. This would provide a first step towards answering both issues. The study in this section uses an implemented motor-skills test to evaluate the competency of the child.

Conversely, the test could be thought of as a test to find out which device is best suited for the child. This test as part of a three test program has also the objective of being able to determine if a given device is suitable for use for a later (ECDT) test in the program. The tests will also provide intermediate levels of ability where more analysis had to be done to verify if this is due to a ceiling (Ericsson, 2006) set by device competency (Hill & Schneider, 2006) or a cognitive limitation. The differentiation is made by this study to separate specific motor effects from cognitive effects which affect both device and test handling. Cognitive competency typically involves the use of frontal lobes of the brain (Miyake et al., 2000; Nyhus & Barceló, 2009). Device competency requires in addition other area of the brain such as the motor cortex (Penfield & Boldrey, 1937; Ungerleider, 1995). Device competency can be overcome with better devices. A cognitive limitation referred to here is related to the test and cannot be resolved by a better device since activation of the frontal lobe is required (Duncan & Owen, 2000). If the cognitive limitation is due to for instance frontal lobe impairment, a better device would not produce better test results.

The measures used in this test give focus to the impairments of PNI children by filtering out the noise cause by muscular instability. The measures are primarily based on consecutive (streaks) patterns which are a better indication of intention than singular responses. Using this approach, this test also reduces the level of noise that is a consequence of the pattern of stimulus presentations in ensuring that it is not possible to attain success streak sizes above a certain bounded value. The use of streaks is particularly relevant for a motor-skills test like this as the process addresses noise caused by physical random fluctuations (Shannon, 1948). However, the use of streaks (Altmann & Burns, 2005) is not precluded from higher level cognitive tests as the principle of looking for intention applies whether it is from inputs caused by a jittery device (either due to mechanical imperfections of the device or muscular impairments of the child) (Oskarsson et al., 2009) or inputs due to uncertainty of choice (psychological inputs)(D. A. B. Grant, Esta A., 1981, 2003; Rhodes, 2004).

We see from the above a differentiation of responses due to mechanical, muscular and psychological origins. There is a different distinction we can make in looking at the muscular and psychological responses which are both from the user. Although the intention of a motor-skills test is to evaluate the competence level of a child with respect to purely using a device, it is recognised that effort is also required to understand and respond to the test. This thesis refers to the effort applied by the brain to control the device as motor control and the effort to understand the test as abstract thought. The effort is visualized as a specific cognitive load. The test is made as simple as possible so that the effort required is kept to a minimum but at the same time still demanding a level of effort to be successful in the target test (ECDT). Within the set of three tests selected for the main study, this test is deemed to have the lowest cognitive load. The test is described as a constant cognitive load test as the test requires a mechanistic procedure for the response to each trial or stimulus presentation. This is to distinguish from later tests where there is no mechanistic procedure for the responses initially until a solution is deduced and thereafter, the mechanistic procedure applies.

As the devices will be novel to the children and demand skills that they may not have used often, it is expected that a learning time is required to operate the device. The assumption is made that it will not take a long time to learn to operate the device to engage with the test albeit that the children will start off with different levels of competencies from each other. A short familiarisation exercise that lasts no longer than the actual test is given to the children. It is also assumed that some devices will be more familiar and intuitive than others so that not only will competency vary between children but devices will also vary in affordances.

As a major cause of variation for this study is cognitive impairment, the controls set up for this test are groups that do not suffer from cognitive problems. The groups chosen for controls are non-PNI children and non-PNI adults. This would validate the devices, tests and assumptions made about learning time. The confidence has also been provided to a certain degree by the pilot study.

This chapter is structured as follows. Section 7.2 describes the aims and objectives of the assessment. Section 7.3 describes the experimental method used to achieve the aims. Section 7.4 describes the results obtained by the assessment. Section 7.5 is a discussion of the trends found in the assessment. Section 7.6 concludes this chapter.

7.2.Aim

The purpose of this study is to test a participant for their ability to handle devices that will each demand different specific motor skills. A ranking is developed for the competency with devices used by each participant. The test was designed in-house on a computer and is a simple competence test (COMPTEST) which requires a binary (YES/NO) response. This test mainly addresses research question RQ1.1 highlighted below.

Research question
RQ 1.1. Does a PNI child have sufficient motor control to use the device?

7.3. Method

This section describes the method used to achieve the assessment of device competency for a future test (ECDT) for PNI children.

7.3.1. Participants

Participants used in this study are common to all main study tests and are documented in section 6.1.6.

7.3.2. Procedure

The participants are tested in a room equipped with a laptop, separate screen monitor and hand-held and non-hand held devices for interaction with the software tests running on the laptop. The screen monitor is arranged side to side with the laptop so that the participants with a view to the screen monitor are seated beside the researcher who has a view to the laptop. The eye-tracker is mounted below the screen of the monitor using magnetic mounts. The head-tracker uses a remote web camera enabling tracking from the monitor.

A set of stimulus is produced to which the participant must provide a positive or negative response. The positive response is an active response which involves the actuation of a device. The negative response is passive requiring no action. Success terminates the trials for each device determined by 20 consecutive correct responses. Otherwise, the trials terminate after a block of 32 trials. Sessions of several blocks of trials involving different devices and tests are conducted once a day and all tests are run within 2 sessions of an hour each. During the tests observations were made by the researcher regarding test response behaviour and notes compiled after testing. Informal feedback was sought regarding fatigue, comfort levels and preferences after tests from adults only as children found making comparisons difficult. Devices were run in order of increasing complexity as observed in a pilot testing. In general, this would involve running tests with the mouse/switch first, followed by single mode devices and then hybrid multi-modal devices. Signing is used when it was determined that there were no other devices that seem to be reliable enough to be used.

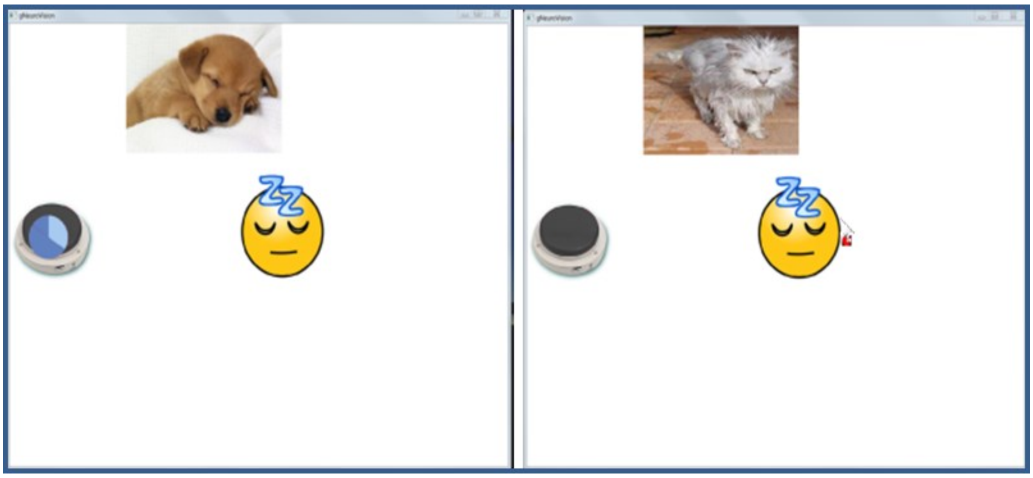


Figure 25 - Stimuli for COMPTEST. The left screen dump shows the stimulus (dog) where a positive response (push the button) is required. The right screen dump shows the stimulus (scraggy cat) where a negative response (do nothing) is required. An eye-tracker was used to "push the virtual button". When the participant gazes at the button a pie appears which provides feedback that fixation on the button is required with a light blue slice that increases with fixation. When the entire pie is light blue, the button is pushed. For a negative response, the participant has to fixate on the sleepy man.

When the EEG headset is used, participants provide a positive response by gritting their teeth for a period identified as the "bite-time". The use of bites is unusual with an EEG headset and the expected norm is to capture cortical signals. The strategy employed was to use whatever proved to be the best means of interfacing that was afforded to the PNI children which provided the option of cortical signals or facial artefacts. However, it was found that with the headset and software combination used, facial artefacts were easier to detect, amd hence more reliable, than cortical signals. When an eye-tracker is used, participants have to move a mouse cursor with their eyes to an image of a switch and "dwell" the cursor over the virtual switch for a period identified as the "dwell-time". When a head-tracker is used, participants have to move a mouse cursor using movement of their head to the virtual switch for the dwell-time. The hybrid EEG and eye-track device uses the eyes to move the mouse cursor but instead of a dwell-time, participants have to grit their teeth for a bite-time to indicate a positive response. In a similar way, the EEG and head-tracker used the movement of the head to move the mouse cursor and the teeth-grit to actuate the virtual switch. When Signing is used, the participant is asked "Is this a dog?" to which they would use gestures that are specific to them to respond with YES or NO to a carer who will then interpret their responses. A complete list of all the devices used and the general mode of usage is given in section 6.1.3.

The stimulus (Figure 25) for the COMPTEST consists of either an image of a cuddly dog or scraggy cat. Participants have to provide a positive response when they see an image of the dog and negative response, do nothing, when they see the image of the cat. A cute dog was chosen for a positive response to encourage the children not to respond to the cat because they like cats. Participants are tested on all devices on the COMPTEST with each device forming a block with a maximum of 32 trials. Feedback is provided in the form of the sound of a bell tinkle when the virtual switch is actuated but not whether the response is correct or incorrect. Participants are familiarized with a few runs of the test to ensure they understand the test and are able to engage the devices.

7.3.3. Design

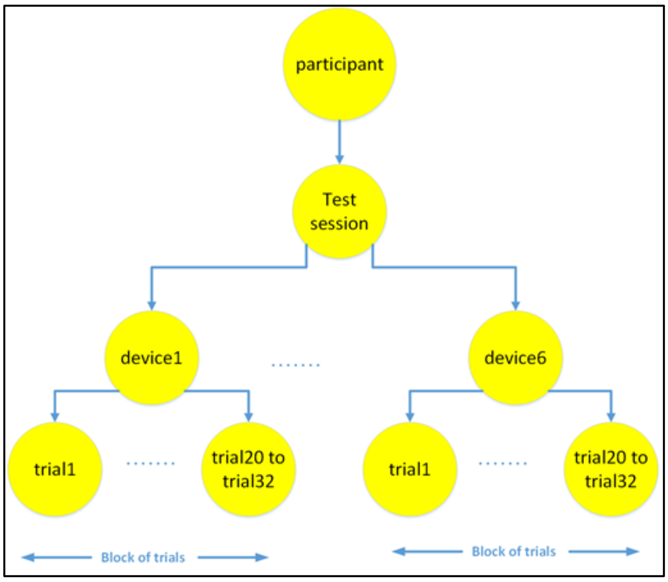


Figure 26 - Test design for main study

Figure 26 shows the test design. The experiment is a 6 x 32 within-subjects design for the maximum number of trials. For each participant the following are components of the test:

- Device {(Signing), mouse/switch, eye-track, head-track, EEG, EEG and eye-track, EEG and head-track}. Signing is only used when no other devices can be used and does not increase the maximum count of devices for trials.
- Block {1 to a range between 20 and 32}
- Tests {COMPTEST}

There are 18 participants giving a maximum total number of 18 x (6 x 32) trials.

7.3.4. Data capture

COMPTEST results are represented as a 32-bit field, each bit representing an OK/NOK (not OK) outcome for a particular trial. The field can be represented by a list of success and failure streaks. Failure streaks are suffixed with x. For example, for a list of 17 successes followed by 3 failures, 10 successes and 2 failures, the list is represented as {17, 3x, 10, 2x}. The consecutive successes and failure are referred to as success and failure streaks respectively. The entire list which captures an entire block of trials is referred to as an outcome event sequence. The maximum number of consecutive successes in the example list is 17.

7.4.Results

This section describes the results obtained for an assessment of device competency.

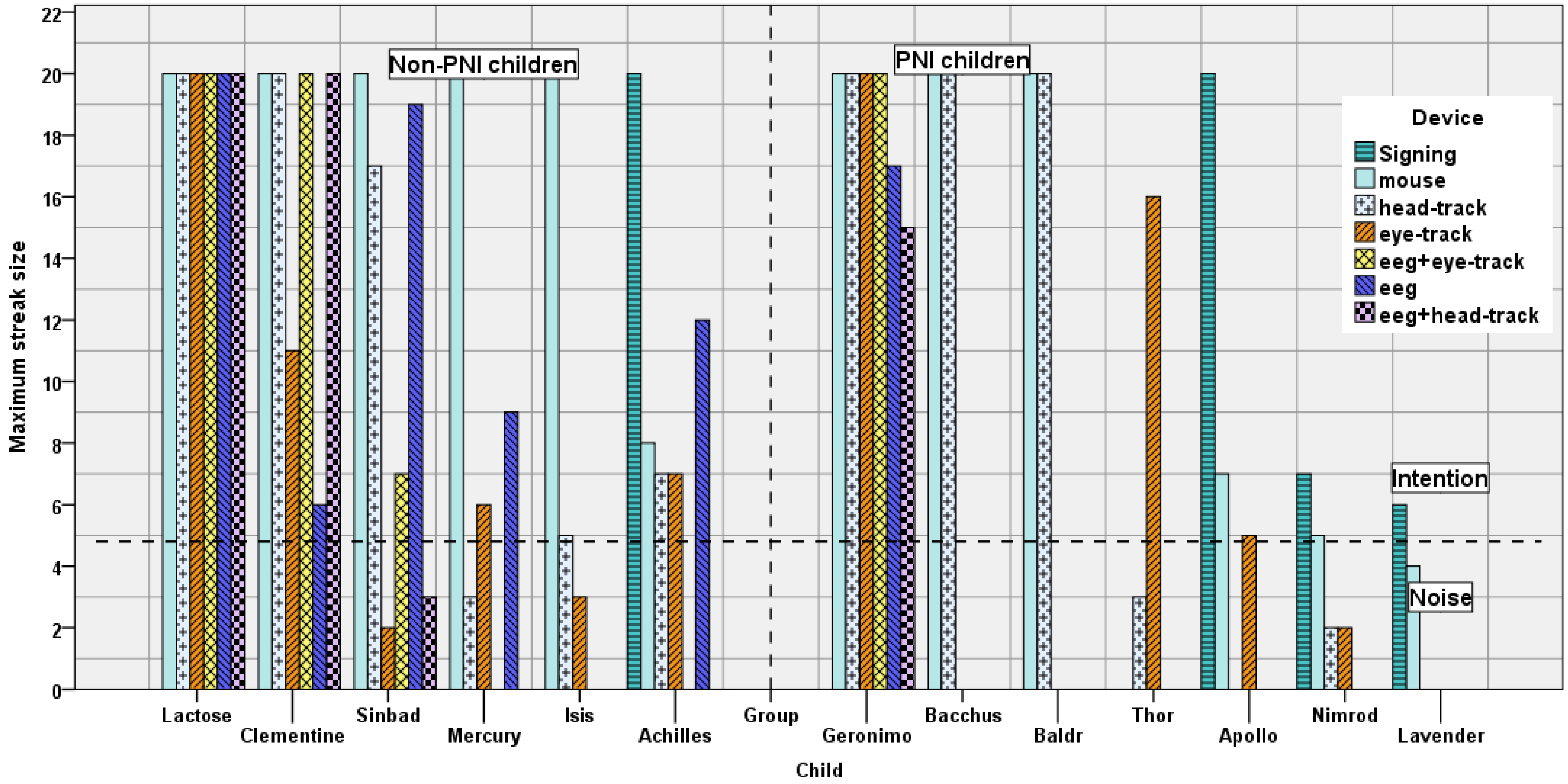


Figure 27 - Device competency of children evaluated using maximum streak as a measure in COMPTEST. The vertical dotted line separates non-PNI children (left) from PNI children (right). The dotted horizontal line represents a threshold that separates intention outcomes from possibly random outcomes. This figure shows that a mixture of abilities that are present within both the non-PNI and PNI group although in general results tend to be poorer with PNI children. Apollo demonstrates that the mouse he is using does not afford him the opportunity to exhibit his cognitive capability compared with the use of Signing. Clementine and Geronimo demonstrate better ability with the eeg+eye-track multi-modal device than the individual single modes of eeg and eye-track.

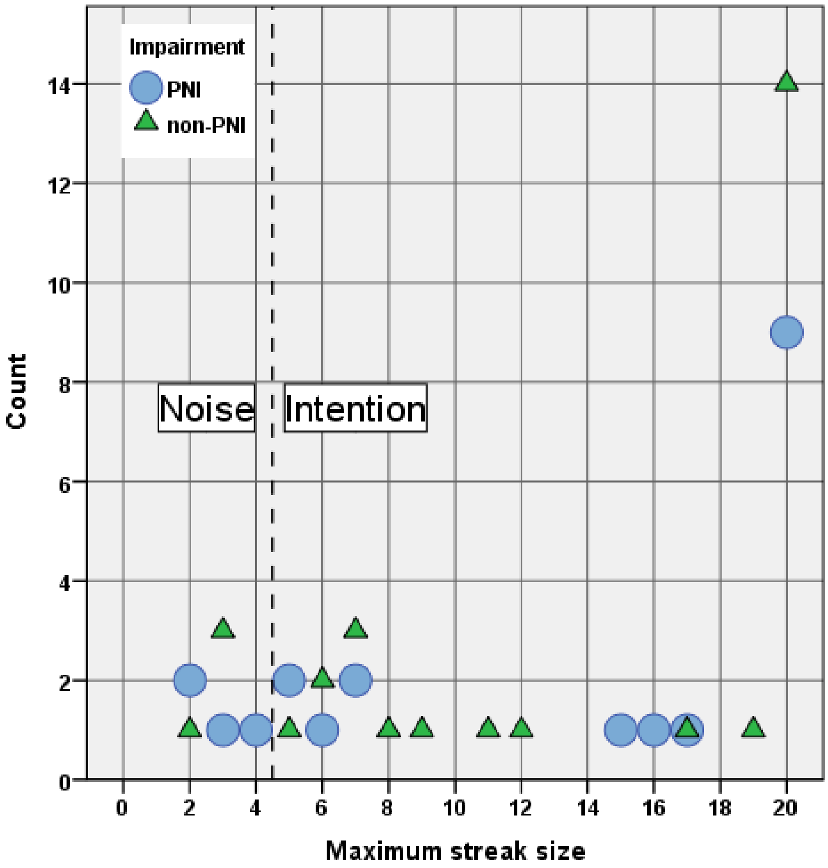


Figure 28 – Maximum streak size distribution for COMPTEST. Results for PNI children are shown with circles and non-PNI children with triangles. Each count is a test run with a different device. Streak sizes of 0 are ignored as the participants did not engage with the test. The dotted vertical line represents a threshold that separates intention outcomes from possibly random outcomes.

7.4.1. Results from the experiment with PNI Children

Figure 28 provides an overview of the maximum streak size data for success streaks. Each maximum streak size represents a block of trials that a PNI child (circles) has had with a specific device. The count therefore shows the number of blocks that achieved a specific maximum. We see that performances are clustered into three groups of achievements (2-7, 15-17, 20).

The lowest maximum streaks hint that there is a level of low streak sizes which is noise naturally generated by the system from which no useful conclusion can be derived. At this point it is helpful to

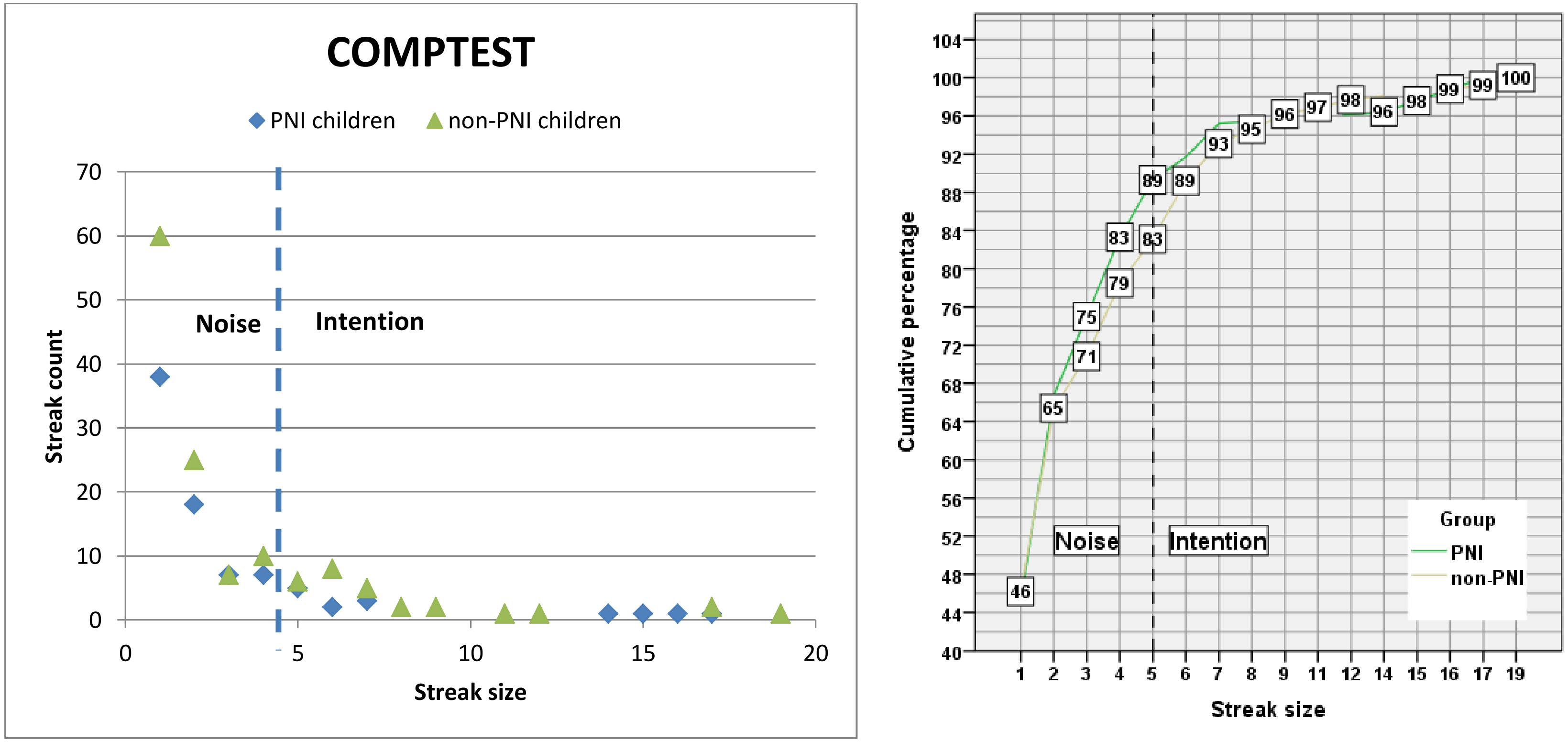


Figure 29 – Distribution of success streak size (0 < streak size < 20). The left plot is a scatter plot of two series (PNI children, non-PNI children). The right plot is the corresponding cumulative distribution function of the streak size for the two datasets (PNI children, non-PNI children). The main cumulative percentages are highlighted in the square boxes. In both left and right plots, the dotted vertical line represents a threshold that separates intention outcomes from possibly random outcomes

examine the distribution of streaks from the 2 to 19 maximum streak size groups to explore the noise threshold. The test runs that obtain 0 (not able to use the device at all) and 20 are ignored because they all terminated earlier than 32 trials. In this way, the lower attainment levels are given more focus and are further sub-divided to provide more information.

Figure 29 (left) shows the number of occurrences of a particular streak size (taken for all the tests) for the streak sizes ranging between 1 and 19. PNI children are represented using a blue diamond on the left and a green line on the right. Streak sizes of 1 and 2 are very common because the test program generates that noise for random responses. Streak sizes between 3 and 19 are much less likely. A reasonable point to choose the noise threshold starts from the "knee" of the data as the data flattens out. In order to be compatible with generic cognitive tests (eg. Wisconsin Card Sorting Test) the noise threshold was chosen at 5 so that streak sizes of below 5 would be treated as noise and for sizes of 5 and above as intention for control. The term intention is used to mean that the action is not the result of a reflex action but arising from a conscious decision (Libet et al., 1993). Figure 29 (right) provides the confidence levels for the streak sizes obtained for the left plot. A streak size of 4 and below would lie in a confidence interval of 83% (PNI children). The higher the confidence interval, the more common the data would be and the less information it would contain (Shannon, 1948).

Figure 27 emphasizes the child's ability with different physical devices using maximum streak size as a measure. The right side plot shows results for PNI children. The children are sorted to show a general trend of ability starting from the left. As the comparison is done for ability with physical devices, we can examine the trend discounting Signing (blue). Signing (section 6.1.3.7) was only carried out in cases where some validation was required due to the low scores with their best physical device to separate cognitive inadequacy from other problems (for example device problems). The dotted horizontal line identifies useful (intention) and non-useful (noise) results. The results that are above the line are useful and those that are below are not. Confidence that a child would be able to undertake the target cognitive test (ECDT) is only provided if a child manages to achieve the test target streak size of 20 with the device. The results show a trend of mixed ability both in terms of using devices as well as engaging the test. Geronimo was chosen as a control participant with a greater developmental age than the rest. None of the PNI children besides the Geronimo could use all the devices. Three children (Geronimo, Bacchus, and Baldr) have definite choice of devices that they can use for ECDT as they achieved the streak size of 20 with at least two devices. The others have to rely on a best effort fit (Thor, Apollo, and Nimrod) upon achieving maximum streaks of between 5 and 19 with physical devices. One child had problems engaging with the test (Lavender) and had a maximum streak below 5.

The problems encountered by the children (Table 12) can be roughly grouped into the following categories:

- General motor and coordination problems
- Problems with the device
- Cognitive problems

One child had attentional problems in the test and thus appeared as an entry in all rows of the table. With a mouse, the general problem is the loss of hand control in both hands. Although 2 children are cited, one child could manage a finger-on-palm mouse but input is interfered with as the child

Table 12 - Observed device problems with COMPTEST for PNI and non-PNI children. The pink table fields represent PNI children, the blue fields represent non-PNI children and the white fields are common to both PNI and non-PNI children.

Device	Child group	Problem category	Problem description	PNI count	Non-PNI count
Mouse	PNI	Cognitive	Attentional problems	1	
		Motor	Loss of hand control for both hands	2	
		Motor	Cannot close palm when wanted. Closed palm when unwanted	1	
	Both	Motor	Uncoordinated grasp of mouse	1	1
	Non-PNI	Play	Distracted by moving mouse to other monitor		1
Eye-track	PNI	Cognitive	Did not understand how to use device	1	
		Cognitive	Attentional problems	1	
		Motor	Cannot support head with neck	2	
		Motor	Torso moves regularly and eye-tracker loses track	1	
		Motor	Squints when trying to control eyes to move cursor	1	
	Both	Motor	Move head instead of eyes to move cursor	1	3
	Non-PNI	Motor	Jerks head to engage eye-tracker		1
		Conditioned	Unable to stop looking at switch (Midas effect)		4
		Habit	Obstruction due to need to support head with hand		1
		Play	Use only one eye to try device out		1
Head-track	PNI	Cognitive	Did not understand how to use device	1	
		Cognitive	Attentional problems	1	
		Motor	Torso moves regularly and eye-tracker loses track	1	
		Motor	Cannot support head with neck	2	
	Non-PNI	Motor	Head movements too forceful		4
		Habit	Restless movements causes loss of tracking		1
		Habit	Obstruction due to hands fidgeting or supporting head		2
EEG	PNI	Cognitive	Uncomfortable with any contact with device	1	
		Cognitive	Attentional problems	1	
		Device	Head-rest prevents headset from being worn. Cannot support head with neck	2	
	Both	Device	Bites do not register or unintentional triggering	2	2
	Non-PNI	Device	EEG headset uncomfortable		3
		Habit	Restless movements cause interference of signals		2
EEG + eye-track	PNI	Device	Could not use either EEG or eye-tracker	6	
	Non-PNI	Device	Too complex		4
EEG + head-track	PNI	Device	Could not use either EEG or head-tracker	6	
	Non-PNI	Device	Too complex		4

cannot predictably control the mouse clicks. With an eye-tracker, the lack of neck support results in regular motion of the head causing the eye-tracker to lose track (2 children) and when they can engage with the eye-tracker, coordination of eye control with cursor control was the next problem with one child blinking and another moving the head instead in the process (2 children). One child did not understand that eye movement is used to control the cursor. The head-tracker had similar problems as the eye-tracker. With the EEG headset, the head-rest acts as an obstacle as the rest supports the neck of those who cannot hold their neck up and sensors at the back of the head tend to be knocked hard.

Apollo

There were no physical devices that suited Apollo. Apollo successfully completed the test using Signing demonstrating the he cognitively knew what had to be done. The finger-on-palm mouse that he normally used provided poor results due to involuntary muscular activity. The eye-tracker and head-tracker was ineffective as his neck was too weak to hold his head steady. The EEG headset could not be fitted as he tends to knock his head against the head-rest of his chair.

Bacchus

Two devices suited Bacchus; the mouse and the head-tracker. Bacchus normally used a mouse. The head-tracker was successful as Bacchus was patient and had the coordination required to use the device. Bacchus could not engage the eye-tracker as he found the co-ordination required to be confusing. He tended to squint his eyes to try to control the eye-tracker but that made the tracker lose contact with his eyes. During calibration of the eye-tracker where he had to track the motion of a moving spot on the screen, the spot speed had to be reduced. The EEG headset was unsuccessful because Bacchus tried too hard and became too tense during periods when no reponses were required. The tensing of the facial muscles caused undesired triggering of the EEG headset which provided unwanted YES responses.

Baldr

Two devices suited Baldr; the mouse and the head-tracker. Baldr normally used a mouse. Initially, Baldr did not understand how to use the head-tracker but noticed that he could move the cursor with an eye-tracker by moving his head. He was subsequently put on the head-tracker again where he was quite successful. Once he worked out how to control the head-tracker, he derived great joy in its use. The eye-tracker was not effective because the head movements made the eye-tracker lose track. The EEG headset was ineffective because Baldr could not produce bites that were effectively distinguishable.

Geronimo

Geronimo did not appear to have problems with any of the devices he used. The result of interest was that he managed better with the EEG and eye-track hybrid device than the use of the individual single mode device, in particular the EEG headset.

Lavender

Lavender could use a big button switch but she had a short attention span. She did not have the attention span to learn how to use the other devices. The EEG headset could not be fitted because she tended to knock her head against the head-rest of her chair. It appeared that her poor results in the test was the result of her poor attention span. When Signing was used, she achieved a better result than the use of the switch because she was interactively prompted for an answer. Even so, the results were poor.

Nimrod

Nimrod could use a mouse but was a bit clumsy with it. Nimrod had cognitive problems with the test. He had problems with both the eye-tracker and head-tracker because he did not seem to be aware that his head or eyes were controlling the cursor on the screen. He did not want the EEG headset to be put on his head as he was wary of situations outside his comfort zone. When Signing was used, he achieved a better result than the use of a mouse but the results were still poor. Nimrod seemed to be confused after a short time which seems to be the cause of the poor results.

Thor

Thor could only use an eye-tracker due to his severe impairments. The device appear to be limiting his performance in the test. Thor could not use a mouse because he does not have the motor control for it. He could not use a head-tracker because his neck is weak and he does not have the fine control required. He could not be fitted with an EEG headset because he keeps knocking his head against the headrest of his wheel-chair. The fact that he could use an eye-tracker was quite an achievement because he has involuntary movement of his upperbody which causes the eye-tracker to lose track. A great amount of effort goes into re-acquiring track and manipulating the eyes as well as responding to the demands of the test.

7.4.2. Results from the experiment with non-PNI children

The distribution of a particular streak size for non-PNI children (shown by green triangles in Figure 29 (left)) shows a steep drop with a “knee” that resides close to a streak size of 5. This provides the cumulative distribution plot shown by the light brown line for non-PNI children in Figure 29 (right). A streak size of 4 and below would lie in a confidence interval of 79% (Figure 29 (right)).

Device competency results (Figure 27, left) show a trend of mixed ability both in terms of using devices and all could engage the test. Almost all the children except one (Achilles) can be definitely represented by one physical device having achieved the target streak size of 20 for the test. Two children (Lactose and Clementine) have multiple devices to choose from. Achilles was tested using Signing (section 6.1.3.7) which showed that he had no cognitive problems with the test.

The observation notes taken during the test indicate the problems the children had with the devices. The problems can be roughly grouped into the following categories:

- General motor and coordination problems
- Problems with the device
- Conditioned responses
- Childhood habits that interfere with device operation
- Motivation/Sense of play in exploring limits of device
- Fatigue

Table 12 gives a summary of the problems the children had with the test. Different devices tend to have different types of problems. Some of these problems are driven by the device and others by human make up. Certain problems are more prominent than others. Generally a mouse has no problems but a motor problem observed is a lack of refined motor control which shakes the cursor out of the control area (1 child). An eye-tracker has two popular problems; one of which is the lack of coordination required to associate the eye movement with movement of a cursor on a screen (3 children), and the inability to avoid fixation on the virtual switch (4 children). The head-tracker has a similar problem of lack of coordination as the eye-tracker (4 children). The response of the head-tracker is fairly slow and a lot of children tend to want to "throw" the cursor by jerking their heads instead of using smooth and slow control (4 children). The EEG headset requires constant contact of saline with the scalp and tends to be a source of distraction (3 children). Another problem is that bites that do not generate the appropriate signals are not recognised (2 children). With the hybrid devices that are multi-modal tend to be a problem for most children because of problems with the single modes anyway or the assembly then becomes too complex to handle due to the additional cognitive load generated by handling two devices (4 children). However, Clementine who was not able to handle the single modes of the EEG head-set and the eye-tracker well but had no problems with the multi-modal EEG head-set and eye-tracker hybrid because the "Midas-effect"(R. Jacob, Leggett, Myers, & Pausch, 1993) problems with the single mode eye-tracker and the longer bite-time required by the single mode EEG headset were compensated for by the hybrid. Fatigue is listed as a possible factor that cannot be ignored but is outside the scope of this study.

Achilles

Achilles had poor results for all devices. The implication here was that there was no focus to achieve good results rather than ability to use any device. When the mouse was used, Achilles was clumsy with the device and tended to lose the cursor. When the eye-tracker was used, Achilles tried to move his head instead of his eyes. When the head-tracker was used, he moved his head too violently to provoke a reaction which made the head-tracker lose track. When the EEG headset was used he would move it to get comfortable. When Signing was used, he was successful in the test demonstrating that he knew what had to be done.

Clementine

Clementine could use all of the devices. An interesting result was that she did well with the EEG and eye-track hybrid device but had poorer results using the individual single mode devices.When she was using the eye-tracker, she could not help looking at the switch which triggered false activation for single mode operation during periods where no responses were required.

Isis

Only the mouse was suitable for Isis. She fiddled with her headband frequently and was quite restless. When using the eye-tracker, she moved her head to move the cursor causing the eye-tracker to lose tracking. When she used the head-tracker, her head movements were too forceful causing the head-tracker to lose track. When using the EEG headset she was too restless and kept brushing her hair probably not being used to having a headset on. The multi-modal devices were not used as she had basic problems with the single mode devices.

Lactose

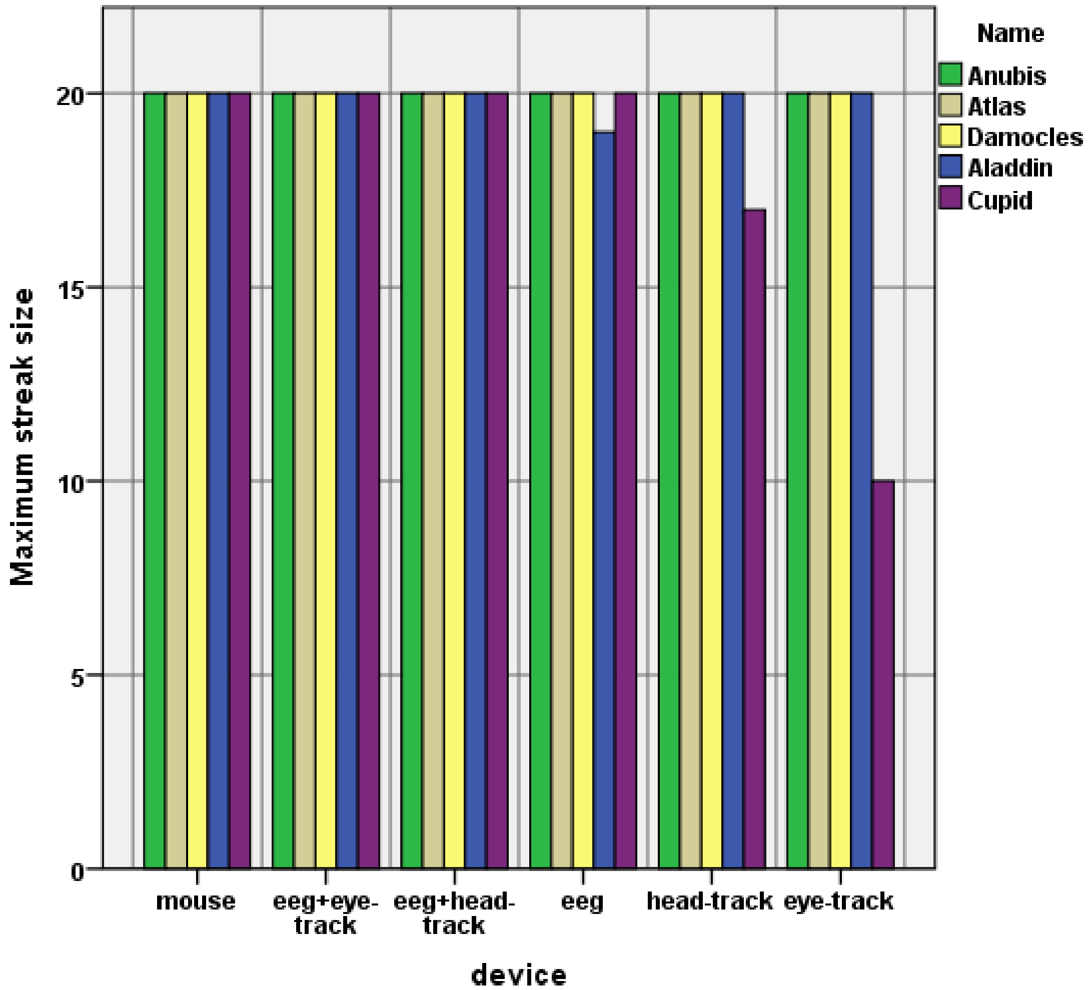
Lactose was successful with all the devices. She appeared to have an innate talent that allowed her to overcome the challenges the novel devices posed.

Mercury

Only the mouse was suitable for Mercury. Mercury developed a habit of resting his head against his hand which made the use of devices other than the mouse problematic. When the eye-tracker and head-tracker was used, the arm was an obstruction to tracking and he jerked his head to engage the tracker which had a negative result. He was not used to having a headset on and was restless when the EEG headset was used which caused unwanted interference. He was reluctant to try the multi-modal devices because he felt they were too complex.

Sinbad

The mouse, head-tracker and EEG headset suited Sinbad. Sinbad had problems with the eye-tracker because he could not avoid looking at the switch which caused false activation during periods when no responses were required. When using the head-tracker, Sinbad was impatient with the cursor latency and shook his head too forcefully to try to engage the head-tracker with negative results. Sinbad had poorer results with multi-mode hybrids than the component single mode devices.



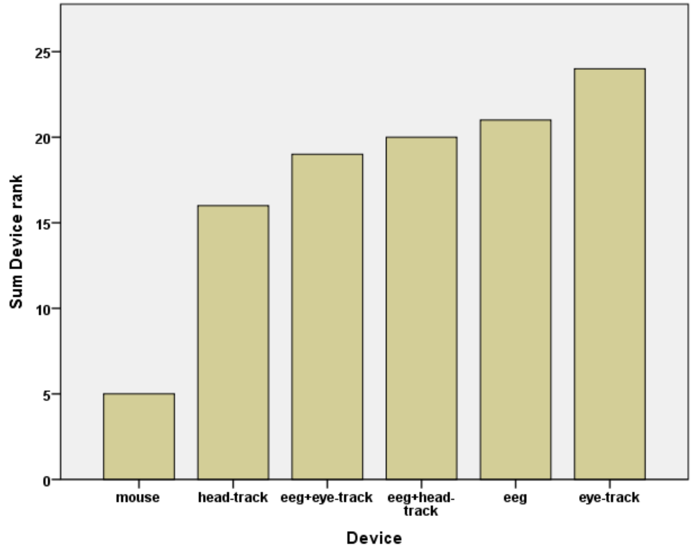


Figure 30 – (Left) Device competency of adults sorted by maximum streak size. (Right) Device competency sorted by preference. Devices were ranked by participants from 1 to 6 with 1 being the most preferred. Rankings were then summed for each device.

7.4.3. Results from the experiment with adults

Figure 30 (left) shows the results from COMPTEST for each device. The maximum streak measure is used to assess the competency of the user in using each device. The plot shows the results for a cluster of participants for each device. The results are sorted with the best results starting from the left to the right. The results show that for adults in general both test and device were accessible with little variation. In most cases, a maximum streak size of 20 was obtained within 32 trials which is the target size for success in the test. The only deviations to that was Aladdin who had a maximum streak of 19 when using the EEG headset, and Cupid who had a maximum streak of 17 when using the head-tracker and 10 when using the eye-tracker. Cupid was unusual in test behaviour for the small sample of adults in that he was exploring the capabilities of the devices as opposed to trying to achieve the test target. The results show a trend that adults have no significant problems with any device.

Figure 30 (right) is a summary of the preferences expressed by the participants for the use of each device. Each participant had to rank the devices (after the test session) in order of preference with 1 being the best device and 6 the worst. The rankings were then added up for each device and plotted as shown. The mouse (5) was unanimously the best device and by a long shot. What follows was the head-tracker (16), then the two hybrids (combination of EEG headset and eye-tracker (19) and EEG headset and eye-tracker (20)), followed by the EEG-headset (21) and the eye-tracker last (24). The results of the new devices tend to be closer to each other than the separation from the mouse. Variations of preferences reveal themselves through the deviation from the ideal sum for each rank. For example, the ideal sum for the second best device should be 12 (6 devices x rank 2) but instead, a sum of 16 was obtained showing that in some cases, the device was not second choice but worse.

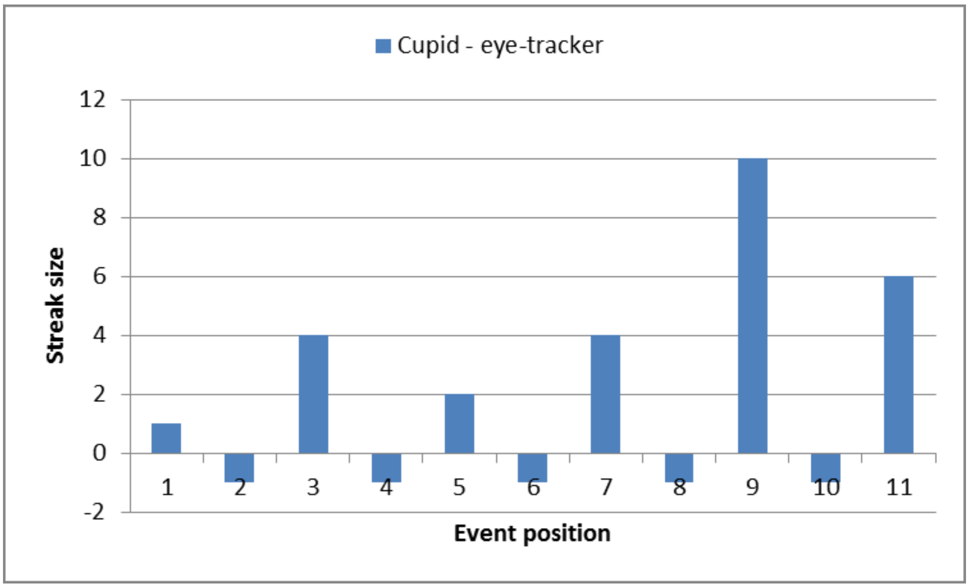


Figure 31 - Graphical interpretation of outcome event sequence collected from Cupid using eye-tracker in COMPTEST. Positive values indicate success streaks and negative values indicate failure streaks. The sum of all the streak sizes provides the number of trials (32) in this test run (or test block).

Figure 31 shows the results of a test run with Cupid using an eye-tracker on COMPTEST. The data capture produces a record of the individual trials as described in 7.3.4 which is graphically shown by the figure. The maximum streak is 10 which correspond to that shown in Figure 30 for the device competency for Cupid. Looking at Figure 31 however, we get a history of the events that took place leading up to the maximum streak and thereafter. We see that for Cupid, error streaks are the minimum size (1). Successful streak sizes build up from 1 to 4 back to 2 and up to 8 before decreasing to 6 (punctuated by single errors).

7.5. Discussion

7.5.1. Discussion of results from PNI children, non-PNI children and adults

The discussion that follows provides comparisons of the results of PNI children with the other control groups:

- Variation of maximum size distributions between PNI children and non-PNI children
- Variation of streak size distributions between PNI children and non-PNI children
- Variation of device ability between PNI children, non-PNI children and adults

- Comparison of device usage problems between PNI children, non-PNI children and adults
- Choosing a device fit for PNI children, non-PNI children and adults

7.5.1.1. Variation of maximum streak distributions between PNI children and non-PNI children

The maximum streak distributions of PNI children showed data clumped into three groups of achievement (low, medium and target attained) (Figure 28). The data for non-PNI children showed a similar trend. The difference is in the higher numbers of each group because the non-PNI children were able to engage in more devices. The test is a simple test that was intended to allow the target to be reached with an appropriate device. A clumping of results that show either a device is accessible or not is expected. The medium category of results that do not achieve the target streak size (20) are of interest because they represent a possible "ceiling" or limit reached in the results where subsequent re-tests may not improve because of impairment. The suite of three tests (section 6.1.4 Overview) allows opportunities for re-tests to compare this ceiling.

7.5.1.2. Variation of streak size distributions between PNI children and non-PNI children

The streak size distributions of PNI children show a steep decline which flattens out in the streak size count as the streak size increases due to the increasing difficulty of obtaining a large streak size (Figure 29). The confidence interval increases and show a similar flattening as a larger streak size is chosen to for the threshold that separates intention from noise. The streak size distributions of non-PNI children show a similar trend. The counts for non-PNI children tend to be higher as they were able to engage with more devices. The confidence interval chosen to determine the noise threshold of a streak size of 5 (implying integer values of 4 and below to represent noise) lies between 83 (PNI children) to (non-PNI children) 79%. The number of participants is small and comparisons between PNI and non-PNI confidence intervals need caution. It is possible that the higher number for PNI children is a result of relatively higher inability to use the devices. As the range of impairments is broad, it is possible that the confidence interval for PNI children is higher if the choice of participants is such that specific impairments result in even greater inability. The expectation would be that as ability is higher the confidence interval approaches that for non-PNI children.

The specific value of the noise threshold is a compromise between choosing a level that will reject too much of the data which may be valid and providing a high enough quality of results. For example choosing a noise level of 7 means that most of the results for a test run will be rejected as noise but of the few that are 7 or above would give great confidence that an intentional indication was given. The value of 5 was chosen for the noise threshold because the results from the test showed a reasonable confidence interval and also that it would provide compatibility with the target test (section 2.4.3) and similar psychological tests (Wisconsin Card Sorting Test (WCST)). Psychological tests such as WCST have psychological interpretations for ranges of consecutive successes. Using an appropriate value for the noise threshold in such a test allows a convenient fit for the interpretation of results for PNI children.

7.5.1.3. Variation of device ability between PNI children, non-PNI children and adults

The general results for PNI children, non-PNI children and adults show a trend of increasing ability from the PNI group to adults respectively (Figure 27, Figure 30). PNI children had mixed abilities with devices and tests, non-PNI children all showed ability with the test but mixed abilities with the device and adults all showed ability with both devices and test.

Some PNI children had to do something additional in a test that the other participant groups did not have to put up with. They had to mitigate impairments that prevent them from using a device. The range of impairments then determine how much effort they have to put in apart from the skill required to ordinarily use a device. Although the results show a poor ability to use a device, sometimes the children have an unimpaired function that has to compensate for the effects of other impaired functions. The results often have a ceiling. The performance will be limited to a certain peak. Detection of this is often accompanied by the detection that the child is cognitively more capable in a test than the result implies. This finer level of differentiation is important whenever a test that is crucial to the participant is applied to ensure that the results of the test are not contaminated by the effort required to handle the device.

Despite the general trend of least ability, PNI children have a wide range of impairments and demonstrate a broad range of abilities within their participant group as well. The control participant (Geronimo) was chosen to demonstrate that PNI children can develop the ability to use all the novel devices on offer. An added complication is that the physical age of a PNI child does not correspond to the mental age. It is thus possible that different developmental processes mature but others do not when compared with a non-PNI child contributing to differences observed between PNI and non-PNI children.

From the group of non-PNI children we find that we have more children demonstrating ability to be able to access more types of devices than PNI children. The trend however still show a mixture of abilities with one child (Achilles) requiring Signing to show that he did not have any cognitive problems with the test. The higher ability came from two girls with a higher mental age than the other non-PNI children (Lactose and Clementine) who were able to use the hybrid (mixed mode) devices in contrast to none of the PNI children apart from the control participant (Geronimo). The results suggest that at this age range, abilities develop to a point where the skills (D. A. Rosenbaum, Carlson, & Gilmore, 2001; Wade & Whiting, 1986) for all the novel devices start to emerge with some individuals having that ability. Demonstration of a spread of abilities was partly made by some non-PNI children who did not attempt the hybrid devices because of the perception that it was too complex and others who felt confident to try them. The results are consistent with findings that the motor abilities of children are still developing (Donker & Reitsma, 2007b; Kuhtz-Buschbeck, Boczek-Funcke, Illert, Joehnk, & Stolze, 1999).

Adults in contrast to children demonstrated almost perfect capability to handle all the devices. In general based on the participant feedback, we can conclude that user ability compensated for projected complexity of use by the theoretical model (section 2.3.6). An example is Damocles who complained of the Midas effect when using the eye-tracker and yet achieved the target criterion of a maximum-streak size of 20.

7.5.1.4. Comparison of device usage problems between PNI children, non-PNI children and adults

The trend with device usage problems were seen to be more problematic with PNI children and decrease with adults. The results for device usage problems (Table 12) for PNI children show most categories that are a result of their impairments. Non-PNI children exhibit additional device problems due to immaturity in operation because there is no impairment that masks the demonstration. Adults did not have significant device problems (Figure 30) as their ability overcame (7.5.1.3) the problems. However, common problems exist between PNI children and non-PNI children in terms of motor problems and common conditioned behaviour posed problems common to non-PNI children and adults.

PNI children have a range of device usage problems but their impairments tend to dominate, either masking the exhibition of other problems or the children could not engage in the devices to exhibit problems. Problems with device usage could be directly or indirectly attributable to impairment. Problems that are directly attributable to impairment include loss of motor control and attentional problems. Motor control problems such as involuntary muscular activity prohibit proper responses and triggers incorrect responses. Attentional (McCall, Kennedy, & Appelbaum, 1977; Wolff, 1965) problems tend to be one of the most problematic, excluding the participant from being tested. Signing (7.5.2) alleviates that to a certain extent with one-to-one engagement but will be difficult to duplicate in a machine. The progress with robotics (Dautenhahn et al., 2009) offer possibilities with the use of humanoid robots (androids) for engagement in limited test scenarios.

The eye-tracker software (Tobii Technology, 2013) uses the image of a teddy which shakes and rattles for infant calibration which is the sort of idea required for attentional deficits. The need for infant calibration in eye-tracking demonstrates one example of attention required for the training of some bio-modal devices besides the need to engage in the test. However, this did not work well for Lavender who had the attentional problem who required human interaction. Another problem is that the children may not understand the test or the device. Their limited communication vocabulary sometime limited to a handful of symbols prevents sophisticated communication (Waller, 2006). Further diagnostic tests can be subsequently run to track the source of the problem which is out of the scope of this study.

Device usage problems that are indirectly attributable to impairment are ones where the participant has to compensate for the side-effects of an impairment using a functional part of the body or the impairment prevents installation of the device. An example of a compensation that has to be made comes from Thor who suffers from a state of spontaneous torso movements that are irregularly timed, non-repetitive and randomly distributed. Although Thor was able to use an eye-tracker, he has to cope with the random movements of the torso causing errors in his responses. There are two problems with the effect. The first was that it was easy for the eye-tracker to lose tracking and the second was that there was an additional problem of mapping a number of relative movements; the body, the eyes, the cursor and the target.

Non-PNI children do not have impairments that prohibit access and tend to understand how to operate the devices more than PNI children. The device usage problems encountered are a stage further than PNI children in that having engaged the device their immaturity presents the next range of problems. Some children of the age range tested may not have spent much time in an environment where they have had to be still for some time on a focussed task and tend to be restless and distracted (Hanna, Risden, & Alexander, 1997). Children also develop habits that they generally grow out of as they mature such as shaking their legs when they are seated or a sense of play as a priority. Children tend to fatigue faster often resulting in changes of behaviour. Children may have a more physically demanding day with allocated times in the playground. The extraneous bodily activity (fidgeting) which include habitual brushing of hair, resting their heads on their arms, not being able to sit in any one position for a length of time children causes more interference that is detected by the input devices and provides erroneous inputs.

One comparison of interest between PNI and non-PNI children is the discipline shown in terms of attentional distractions. A previous paragraph mentioned that children of the age range tested not having the conditioning provided in a classroom. This applies for the non-PNI children but not the PNI children who normally have a physical age of 11. PNI children on the whole tend to be more disciplined and children with attentional problems like Lavender tend not to exhibit consistent unfocussed behaviour all the time.

Problems for PNI children that is common to non-PNI children rests in the domain of motor development and device deficiencies. A common motor problem is the problem of mapping a body movement with the movement of the cursor on the computer screen. The brain needs to establish a calibration between the abstract (egocentric) representation of the cursor position and exact (allocentric) physical movements of the body (eyes or head) needed to provide a precise response (Burgess, 2006; Ghahramani et al., 1995, 1996; Wolpert, Ghahramani, & Flanagan, 2001). This problem is followed with the more complex problem of breaking down the action into familiar chunks (Laird, Rosenbloom, & Newell, 1984; Miller, 1956; Rosenbloom & Newell, 1986) and applying it into another mapping required by the use of the hybrid devices (De Groot & de Groot, 1978; Fischer, 1980).

One common problem was observed with the adult and non-PNI children who could use the eye-tracker that was sometimes referred to as the Midas effect (R. J. Jacob, 1991; R. J. Jacob & Karn, 2003; Majaranta & Räihä, 2002). This effect is a result of normal use of the eyes which are constantly scanning for objects of interest. In this study, the switch image which was used to provide a YES response was accidently scanned. Special effort had to be provided by the participant to avoid looking at it when a NO response was required. Of the adult population Cupid had the most difficulty with this problem.

Throughout the study, the idea that cognitive effort is split between handling the device and dealing with the cognitive aspects of the test will be suggested. However, the amount of effort devoted to a task will be decreased if the participant is fatigued (including stress) or lack motivation (Xie & Salvendy, 2000). A weakness of this study is that these two factors were not investigated but can contribute to significant variation of the results.

7.5.1.5. Choosing a device fit for PNI children, non-PNI children and adults

As a summary we find a trend where PNI children with less developed ability and more challenges to cope with tend to use the device to satisfy basic needs and avoid impairment (7.5.1.3). As more ability develops and we progress to non-PNI children, confidence develops to use more sophisticated devices. At full maturity with adult abilities, the ability compensates for ease of use and more complex devices are sought which has better advantages.

In general device fit for PNI children revolve around fitting a device that gets round their disability. A PNI child could be fitted with a device on either an impaired or unimpaired function of the body. The differentiation intended here is that a person with disability/impairment could find a device fit using its features, either general features intended to match general abilities or personalised/adapted features that allow a match of a device to one's impairment.

A fit for an impaired function of the body is made when either the device is specially adapted to mitigate the effects of the impairment or in most cases where there are no better options available. An example of the concept for the former device is the use of streak measures that ignores responses made that may be due to involuntary movements. The only device that appears so far to be the perfect fit for impairment is Signing which seem to be able to get the target results for the tests when no other device could but the device is only conceptual. In most cases, a device fit is made for a device that gave the best results albeit not the required result. With a fit for impairments, a ceiling was normally imposed on the results because of the impairment; otherwise much better results would have been achieved. An example comes from Apollo who cannot use any device because of poor gross motor control (for example loss of basic anti-gravity postural control) but a best effort fit is a finger-on-palm mouse (maximum streak size = 7, Figure 27).

A fit for an unimpaired function of the body is fine if impairments do not have side-effects on device usage. An example of the former is in the case of a hemiplegic child where one side of the body is paralysed. In this case, very often the child is fitted with a device on the other side of the body. For example, Bacchus, a child with paralysis on the right side will be fitted with a left-handed mouse. In the case where there are side-effects from the impairment, a ceiling on results may also be imposed. This was the case for Thor who could use an eye-tracker (eye-tracker maximum streak size = 16, Figure 27) but had poor gross motor control like Apollo and involuntary torso movement. It is noted that this is a remarkable display of ability.

Non-PNI children are developing their abilities and the best device fit is the simplest non-intrusive device that fit the requirements of the test. The complexity of the device varies from single mode selection devices (mouse and EEG headset), single mode pointing devices (eye-tracker and head-tracker) to multi-mode pointing devices (hybrid devices). The EEG headset is intrusive in that the device presence is felt on the head. The result is that the mouse is still the best fit for non-PNI children and the only fit for most non-PNI children with other devices being good fits once the children develop the skills for them.

The high ability of adults resulted in a stronger preference for the more complex devices because they can use them to achieve advantages. Damocles rated the head-tracker second because he used some jerky movements so that the reference used by the head-tracker would be lost but realised that the reference could be regained by subsequent appropriate jerks ("shaken back into position"). In particular, the EEG and eye-tracking hybrid could be made to perform better than the single-mode eye-tracker as it provides a quick pointing facility and does not suffer from the Midas effect which is a pre-mature triggering of selection. For example, Cupid ranked the EEG headset and eye-tracker hybrid second because he noticed that he could look at the virtual switch and actuate it by flicking his right eyebrow. The trend in preferences suggests that preference for a device is very much individualized with the least fatigue playing the major role followed by other factors (Figure 30).

The mouse was by far the best choice for most of the participant groups as users have had a long time (years) to get used to the mouse. In this experiment, the mouse satisfied the criterion for test success for 3 out of 7 non-PNI children, 4 out of 5 non-PNI children and 5 out of 5 adults. There is a body of work that looks at the changes in the brain after repetition for a motor task. Recognition is made of a memory for motor skills (motor or procedural memory) such as learning to walk, swim or riding a bicycle (Longstaff, 2000; Mochizuki-Kawai, 2008; Shadmehr & Holcomb, 1997). Repetition is shown to be key in the success of many skills(Ericsson et al., 1993). New devices require a learning cycle dependent on the individual and are shown by the variation in individual performances. Skill acquisition of different levels depend on the specific type of training applied but ordinarily the initial level is where a stable phase of automated performance with minimal effort (Hill & Schneider, 2006) is reached (Ericsson, 2006).

The novel devices tested are only a part of the range of bio-modal devices available. The use of Signing (7.5.2) however, exposes the preferred modes of communication used by PNI and non-PNI children. The use of Signing suggests that non-hand held devices may be possibly better. For PNI children these could include head movements, hand movements, arm movements and each child have their own customized gestures due to impairments. For non-PNI children it includes pointing and voice. These devices have yet to be implemented and more investigation has to be done in this area.

7.5.2. Use of *Signing* – a versatile conceptual tool

Signing for the purpose of this study is any form of direct natural communication that could be used to interact with the children. The children would Sign to a carer or interpreter and their responses would be translated into mouse responses for the test. Signing was used in this test for children whose results were very poor or had no appropriate device to take the test. Signing was considered to require the least effort. Signing would be used from a very early age for communication and the participant would have had a lot of practice. The skill should have developed beyond the stage where the Signs would not require any cognitive association as indicated by Ericsson (Ericsson et al., 1993). Once that level of skill is reached, the brain is freed up to perform other cognitively demanding tasks (Smith & Chamberlin, 1992). Using this premise Signing was used to differentiate cognitive ability for the test from the cognitive demands of the device. An example of this was Apollo who had a poor result with a finger-on-palm mouse was probably due to poor hand control preventing proper use of the mouse. This assumption was also confirmed by the carer in her observation of Apollo's daily use of the mouse. However, when Apollo used Signing for COMPTEST, he achieved the test target (maximum streak size) of 20.

Signing for PNI children normally consists of hand, arm, head or body gestures. The symbol set used by PNI children would depend on the impairment of the child. As the impairments for each child vary, so does the method of Signing used. There is still room for an unfamiliar person to misinterpret as not all Signs are executed clearly. For example, Apollo will raise his right arm at right angles to the right to indicate YES and bring the arm to his left chest so that his hand touches his chest for a NO. There are occasions where the arm drops to his legs just as it reaches the chest when he means a NO. Signing could also be rich in vocabulary causing some confusion. For example Nimrod would nod his head, thumbs up or sign the animal shown in the test for a YES.

Signing for non-PNI children would mostly consist of speech but if tests have options, they may point to the correct answer. It was unexpected to use Signing for non-PNI children but sometimes test results would be poor and Signing was used to see if better results could be obtained. An example of this was Achilles achieved an outcome (maximum streak size) of 8 with a mouse and was subsequently tested using Signing. Achilles voiced YES or NO responses for the test. Achilles achieved the test target outcome of 20 with Signing validating that it was not a test capability that was the problem.

Signing is a personal form of communication. Within this communication, both verbal and non-verbal messages are exchanged. The non-verbal messages not only consist of the specific symbols that may be used for PNI children but also accompanying gestures that constitute body language. The use of a care-giver that is familiar to a PNI child and in the course of interaction between a non-PNI child and the interpreter of Signing, indications are both sent and received without conscious thought which contribute to the communication (Mehrabian, 1977). In the case of a child with attentional problems or less severely, moments of distraction or lack of motivation, the one-to-one communication acts as a compelled request-response protocol which demands the attention of the child. An example of this is Lavender who had attentional difficulties and could not understand her role in the test as she constantly kept hitting the physical switch during the test. Only a one to one interaction afforded by Signing was able to get her attention. In the case of a care-giver communicating with PNI children, Signing has greater significance due to impaired communication channels. The care-giver will be aware of subtle signals that are normally provided by the child that will not be obvious to an unfamiliar communicator. For instance, there may be a particular look that is given when the child means to say YES instead of NO. With computerised tests, these benefits are lost. The contributions of non-verbal communication in the form of gestures and other signs provided by body language such as pointing and expressions need further research. Pointing suggests that touch-sensitive screens may be replacing mice. The Kinect©[3] is an obvious candidate to replicate the detection of some non-verbal communication as it already uses gestures for gaming

[3] The Kinect is a product of the Microsoft Corporation.

(Bleiweiss et al., 2010) and Human Computer Interaction (HCI)(Ebert, Hatch, Ampanozi, Thali, & Ross, 2012). Work on expression analysis is also fairly common (Soleymani et al., 2012).

The combination of Signing being a conceptual tool that children require minimal cognitive effort with and the ability to override factors like attentional distractions make Signing an almost perfect tool. Tests in this experiment show Signing either outperforming (2 children, Apollo and Achilles) or matching (2 children, Nimrod and Lavender) other devices. These devices include the well-used mouse. Being a near perfect tool implies that it is possible to use Signing as a benchmark for comparing other devices. As a benchmark this suggests that a physical device that produces an outcome similar to Signing must also be a good device fit.

The implications that Signing is an ideal device go further. The mechanisms used by Signing in this study involve non-hand held device activations. The only gesture close to a hand-held activation was when children showed a preference for pointing to an image on the screen. A hand-held device that is close to this is the touch-screen devices that are currently widely available. However, pointing is a more difficult mode to detect because it is a gesture that can occur more than once at a distance from the screen. This implies that non-hand held devices have the potential to be better devices than hand held devices but the challenges to implement a good non-hand held device are still significant.

7.5.3. Use of measures

Figure 32 provides a comparison of a typical measure used in testing; the score (blue) with the maximum-streak measure (red). The score measures the number of successful trials in a block of tests and in the figure is expressed as a percentage of the number of trials executed in a block. The figure shows the results using the two measures and is clustered around children using various devices ordered by the child's ability using the maximum-streak measure. The children with the best outcomes start on the left followed by the device outcomes using the maximum-streak measure. Among the non-PNI children, Lactose has the best maximum streak outcomes followed by

Clementine, Sinbad, Mercury, Isis and Achilles. Among the PNI children (right), Geronimo has the best outcomes followed by Bacchus, Baldr, Apollo, Thor, Nimrod and Lavender. The results differ for scores and maximum-streak with scores being always higher general apart from when they both are 100%. The relative trends of both measures are mostly but not always aligned with the scores sometimes increasing when the maximum-streak shows an increase. The differences come from the design of the measures.

Scores are an averaging measure of the successes achieved spread among the total number of tries. Maximum-streak measures the best success achieved as a fraction of the maximum (20) that could be achieved. Scores do not identify events that are significant. For instance, if a child had 16 successive correct responses and 16 incorrect responses represented by an outcome event sequence of {16, 16x}, the score would still be 50%. The score is 50% as long as the total number of successes and errors are equal (e.g. {10, 6x, 6, 10x}). However, a maximum-streak measure would have indicated a significant achievement of 80% for {16, 16x} because a maximum-streak measure specifically distinguishes itself from average occurrences of which noise will be a strong component and acts more specifically as a measure of achievement. Scores are suited for cases where there are many re-attempts interspersed with occasional failures and results are from sources that are generally intentional with occasional errors. The averaging property is useful when it is required to have a measure of sustained capability or when comparisons with the actual noise levels are desired. Maximum-streaks show cases where singular strong attempts are made but there is possibly insufficient energy or opportunity to provide other significant intended outcomes. Scores require changes to provide information relevant to this test. Maximum-streaks are part of the steps required to provide this information.

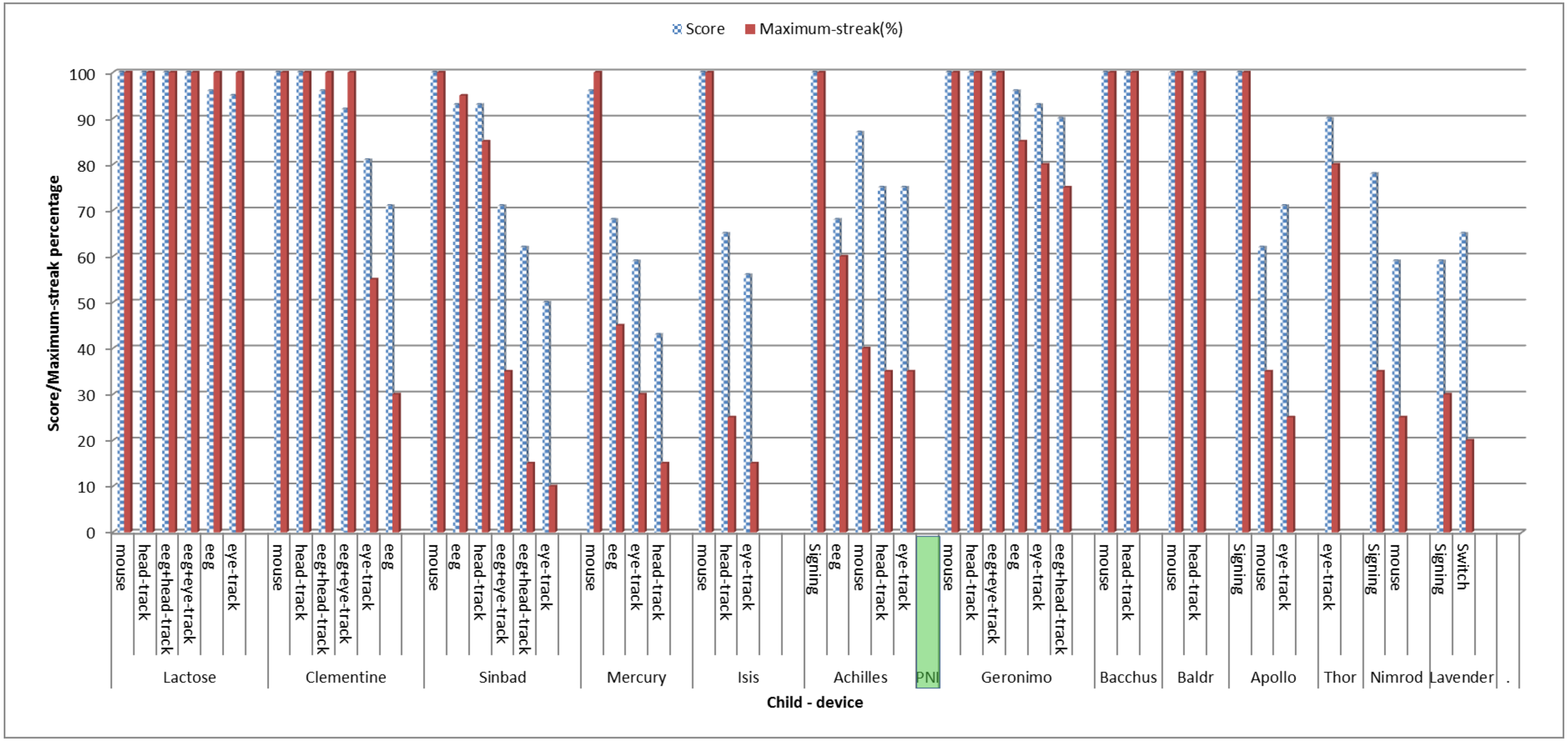
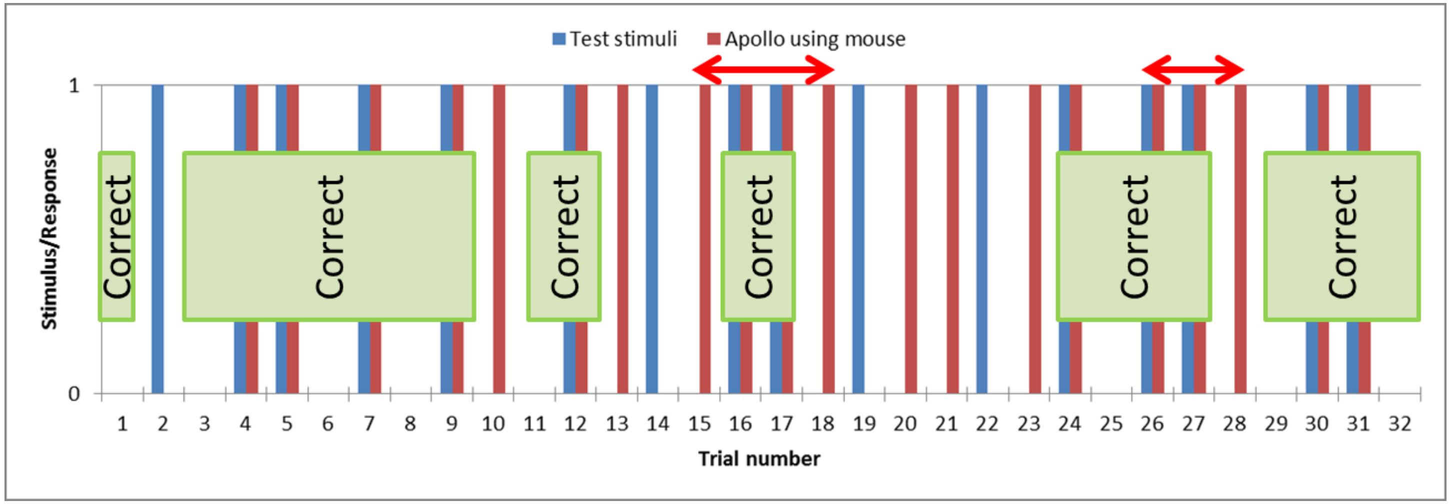


Figure 32- Comparison of maximum-streak measure with success score sorted by child and maximum-streak. Non-PNI children (left) are separated from PNI children (right) using the vertical divider (child) marked PNI.



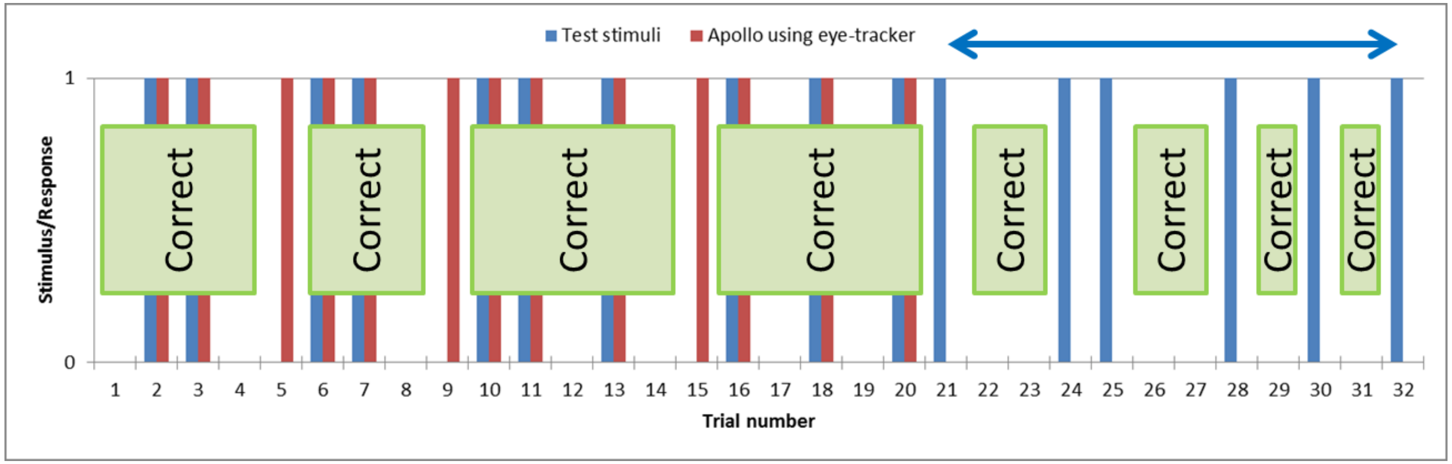


Figure 33 - Comparison of responses to stimulus. Apollo was using finger-on-palm mouse (top), eye-tracker (bottom). The stimuli and response are YES = 1/NO = 0 values. Single blue lines indicate inability to respond with a YES and single red lines indicate inability to respond with a NO. The red arrows indicate consecutive YES responses and the blue arrow indicates a period of consecutive NO responses. The pattern of responses provide a model for involuntary muscular actions. Different devices used look for areas of motor control that are less impaired.

With the use of streak measures it is hoped that possible noise that arises from involuntary muscular outcomes will be quantified and isolated. There are no models of noise generated by the involuntary effects. However, the tests done on Apollo provide some suggestions of the potential noise patterns. When tested using Signing, Apollo achieved the target streak size of 20 without errors. From this it was assumed that he had no cognitive problems with the test. What follows below are two different representations of Apollo's results with a mouse and an eye-tracker. The two different representations show the different focus one set of results can have and how effective a different representation can be on the same set of results. The results for the two devices were chosen because they show two different types of noise patterns that may not be uniquely attributed to a single cause but are convenient demonstrations. Apollo has poor hand control and suffers from quadriplegia CP. When the mouse was being used, involuntary muscular action stopped Apollo from clicking the mouse when he wanted to and caused him to click the mouse when he did not want to. A convenient hypothesis that can be drawn is that the involuntary muscular interventions present during hand control is absent from eye control. By looking at the results from these two devices we may get an idea of different noise models for impairments. However, the tests with the eye-tracker had to be done with a carer supporting Apollo's head as his neck was too weak to do so.

Figure 33 compares the stimuli presented with the responses made. The figure shows persistent patterns made in response to the stimuli. When the stimulus requires a YES response, the value 1 was given and for a NO, 0. The responses were also arranged such that a YES response provided 1 and a NO, 0. If the blue and red bars coincide in value, or the blanks coincide in value, the response made is correct, otherwise (single blue or red bar) an incorrect response was given. We see for the top figure (mouse) that initially there was good synchronism between stimuli and response (from trial number 3 to 9). There follows a burst of activity which had errors interspersed with successes (trial number 10 to 23). The series ends with another burst of good synchronism. Apart from the synchronism, there appears to be an intention to keep up with the stimuli requirements and YES and NO changes in responses are frequent. For the bottom figure with the eye-tracker, the synchronism viewed in general has occasional errors initially (from trial number 1 to 20) then follows a period of silence where there was no response for the rest of the test (from trial number 21 to 32).

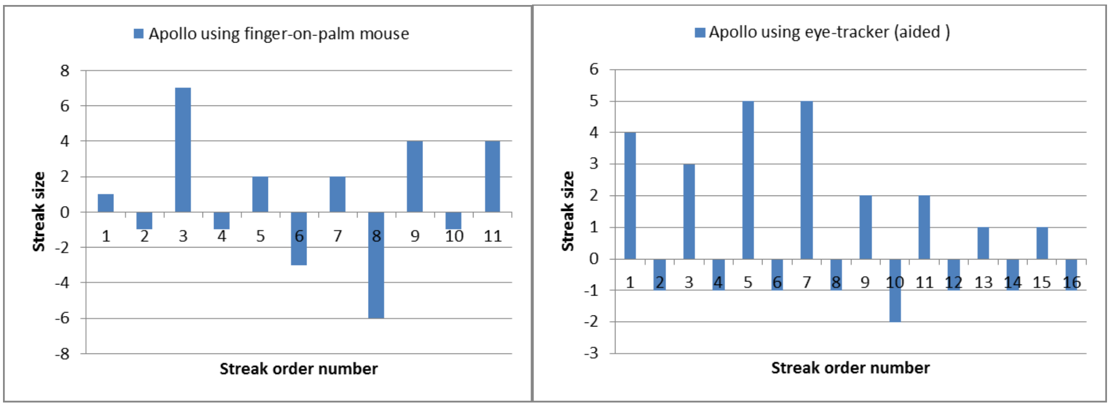


Figure 34 - Apollo using different devices generate different noise patterns (mouse left; eye-tracker right). Errors are depicted as negative values.

Figure 34 shows the same results as Figure 33 but displayed in a different way. The YES/NO results are converted into counts of consecutive successes and errors (streaks). The counts can be seen as synchronism counts for successes and out-of-synchronism counts for errors. Successes are positive and errors negative. The streak order number denotes the position of the streak in an ordered list of streaks (in the order that the streak occurs in the test). The mouse streaks show an increasing trend as the test progresses in both successes and errors. The eye-tracker streaks tend to start off around the same streak size of 4 and then drop off to 2 and 1. The error streak sizes also tend to be small (1) and constant.

From the two sets of representations of Apollo using the mouse and eye-tracker, we can see that Figure 33 is good for tracing test behaviour. The mouse behaviour appears to be one where Apollo is trying hard but possibly involuntary muscular action is interfering. The eye-tracker behaviour appears to be that there is fair control of the device but after a while, fatigue sets in and activity stops (no response). The figure is able to contrast patterns where there are periods of perseverations from periods where there is constant change in responses. Figure 34 shows level of synchronisations (in phase and out of phase) with the test using the same device.

The maximum streak measure is used as a first candidate measure as it is sensitive to users that suffer easily from fatigue due to the extra effort required to both mitigate impairment and drive the input device. It provides good traceability from results to events captured during testing. Other measures which provide a different emphasis to the results can also be devised. One such measure is the streak-sum(Gan, Frank, Amirabdollahian, Sharp, & Rainer, 2014c).

7.6.Conclusions

This experiment evaluated non-PNI adults and children as controls and PNI children for competency with a given range of devices. The objective was to provide a fit of devices to be used for a future test (ECDT). However, this study is also a test of the procedures and evaluation methods used to provide such an assessment of competency of devices.

The study concludes that of the 7 PNI children, 3 have the device to allow them to move to the next stage (Geronimo, Bacchus, Baldr), two need to find a better device but may not be able to do so unless a device is designed that is better suited to their needs (Thor, Apollo) and two indicate cognitive problems that need further investigation (Nimrod, Lavender).

In terms of new technology, the study concludes that their operation requires new skills that some children at that age of mental development (4 to 7) may already have, but others have to develop if possible. The skills involved depend on the nature of the device. These new skills are not a problem specific to PNI children but affect non-PNI children as well. Of the novel devices tested, only two of the PNI children (Bacchus, Baldr) managed to use one (the head-tracker). Results obtained from these devices demonstrate that a certain amount of cognitive ability is required which varies depending on the complexity of the device. The cognitive load imposed by the device adds on to the cognitive load needed to understand the test. The best demonstrations are from the head and eye-tracker where it is a common problem to first associate the head and eye-movements with cursor movement followed by the coordination required to move the cursor to the required position. The coordination problem may be the common thread that excludes some children from multi-modal devices even if they have the ability to drive each of the single mode devices separately. In general, the ability to handle a device in a test for a non-PNI child requires the cognitive ability to handle both the test and the device and for a PNI-child, they have to deal with the interference imposed by the impairment as well which may be direct or indirect. For instance impairment may directly impact performance by an unexpected spasm or indirectly by moving the torso so that the eyes which control the device have to compensate.

In terms of developing a model of the impairment to design devices against, this study has made a start with a hypothetical noise model. This study identifies two patterns of test behaviour using results from Apollo with a mouse and eye-tracker. One pattern consists of multiple attempts at responses in which high numbers of errors are escalated indicating confusion of some origin. The other pattern consists of perseveration in responses where long periods are spent with either positive or negative responses and is common with unfamiliarity (sticky response). The lack of impairment models suggests a need to obtain more participants with similar types of problems to gather results and to observe their clustering. The patterns themselves are not unique to impairment but provide supporting information.

This experiment also concludes that it is possible to have a measure to provide a ranking for device competency that is independent of the noise generated by impairments or the device in the form of maximum-streaks. This however, only provides a snapshot view of best attainment which is a first candidate measure. In order to get a view of sustained performance, it is possible to use an average measure of intentional responses in the form of streak-sum. The use of streaks as a basis for measures provide better measures for PNI children by ignoring responses that are possibly due to involuntary muscular actions. The evaluation is also compatible with the requirements of a future test (ECDT).

The use of Signing has proved good in helping to determine if a child has cognitive problems with the test rather than the device. Signing also provides better results in a case where a child has attention problems. Signing was used with Nimrod and Lavender to suggest that they had cognitive problems with the test. The one-to-one interaction with a familiar carer that a child normally gives an earnest answer to tends to provide an assured response. Signing with a carer that understands the child was seen to be the perfect device to use in a case where the physical devices appear not to be reliable. The problem of the presence of a carer in a test has problems of impartiality. Signing can be used to show the way for a machine that can understand the symbols used for communication as a future development try to get more of the advantages afforded by a carer.

A weakness of this study is transient behaviour that produces variable results for each trial. Even re-tests have a problem of the number of re-tests to carry out. The suggestion is to have at least 2 re-tests after the initial test to provide 3 results for comparison. Another related weakness is that the number of participants of similar groupings is too small to establish statistical trends.

Once the candidate devices have proven themselves to be usable for a simple competence test, a natural progression would be to see if it can perform under a greater cognitive load which leads us to the next test; CATTEST. We get to test the hypothesis mentioned above that the ability to handle a device in a test requires the cognitive ability to handle both the test, the device and deal with impairment.

8. MAIN PILOT STUDY: Light cognitive test (CATTEST)

8.1. Introduction

This is the second test in a series of three (section 6.1.4) tests. The first test (COMPTEST) ranks the devices tested in order of a child's best attainment using them. From the results of COMPTEST, a choice of at most two devices is made for this test (CATTEST). This test further validates the two devices. This test increases the cognitive complexity (Fischer, 1980) of COMPTEST and examines the effectiveness of the two devices under these conditions. In this study, an increase in cognitive complexity is seen as an increase in cognitive load (Hart & Staveland, 1988; Rubio, Díaz, Martín, & Puente, 2004). The third test (ECDT) is the target test and has the highest cognitive complexity. By increasing the cognitive complexity of the test and re-testing two devices of similar performance, ideally the two devices should produce similar results albeit different to the previous test (COMPTEST). If the results are vastly different under the increased cognitive load imposed by CATTEST then the hypothesis is that effects are not solely due to the increase in load but possibly an interaction occurs with the device as well. CATTEST continues the investigation of suitable devices for PNI children that began with COMPTEST that continues to be pursued by the third test (ECDT).

CATTEST requires the children to perform a simple categorisation which was designed to be a mechanical task (Fisher, 2011; Sloutsky & Fisher, 2011) for children of the age range tested (ages 4 to 7). According to Piaget and his 4 stages of cognitive development, a limited form of classification is a feature that occurs consistently in the Pre-operational stage (ages 2 to 7) (Sroufe et al., 1992). The categorisation test was seen as a more complex cognitive test than the motor-skills test (COMPTEST) as a taxonomic match(Blaye & Bonthoux, 2001) has to be made in CATTEST in addition to just recognising the images (in COMPTEST). Since PNI children were shown to have cognitive problems with even COMPTEST, this test allows for further verification of cognitive problems (Davidoff & Roberson, 2004) before the target test (ECDT). This test can be used to find out if a simple categorisation test can detect basic cognitive problems for use as an effective filter for PNI children who will fail a future learning test. Non-PNI children used as a control group for this test will verify if it is accessible (Osborne & Calhoun, 1998) to children of that mental age group.

Three sets of control groups (adults, non-PNI children, capable PNI child) were added to this study to provide a comparison for the devices and the tests. Together the control groups provide a variation of different developmental milestones for comparison. The control groups verify if they have any problems with a test that a 3 year old child(Dunham & Dunham, 1995) is expected to do.

An overview of the structure of this chapter follows. This section gave an introduction to a test that is intended to verify that the best device identified for a given participant by a previous test (COMPTEST) is still able to maintain its fit for use under cognitive load. Section 8.2 describes the aims and objectives of the test. Section 8.3 describes the experimental method used to achieve the aims. Section 8.4 describes the results obtained by the test. Section 8.5 is a discussion of the results of the test. Section 8.6 concludes this chapter.

8.2.Aim

The primary aim is to verify that the best device identified for a given participant by a previous test (COMPTEST) is still able to maintain its fit for use under cognitive load. The test uses the two best devices (if possible) identified by COMPTEST. The results for the two devices are compared. Significant deviation of results where one device outperforms another suggests the possibility of device-test interaction effects.

Research question
RQ 1.2. Does a PNI child have sufficient motor control to use the device and undertake a light cognitive test at the same time?

8.3. Method

8.3.1. Participants

Participants used in this study are common to all main study tests and are documented in section 6.1.6.

8.3.2. Procedure

The participants are each tested in a room equipped with a laptop, a separate screen monitor and hand-held and non-hand held devices for interaction with the software tests running on the laptop. Each participant was tested on 2 of their best devices; the usual device that the participant is most familiar with (typically a mouse) and one other device which testing with COMPTEST showed as appropriate, especially if the device provided better results than the usual device. The other device may not be the second highest in rank if it was judged that the child would be significantly more comfortable with another device that has similar results. Some participants could only use one device because of impairments which limits the assessment of device variation.

The screen monitor is arranged side to side with the laptop so that the participants with a view to the screen monitor are seated beside the researcher who has a view to the laptop. Although a seating with the researcher in front of the participant is preferred (to prevent distractions from the laptop), it is avoided because the headset Bluetooth connection is unreliable using this arrangement. The eye-tracker is mounted below the screen of the monitor using magnetic mounts. The face-tracker uses a remote web camera enabling tracking from the monitor. The order of devices tested was arranged to minimize anticipated boredom and fatigue so that the easiest would be done first.

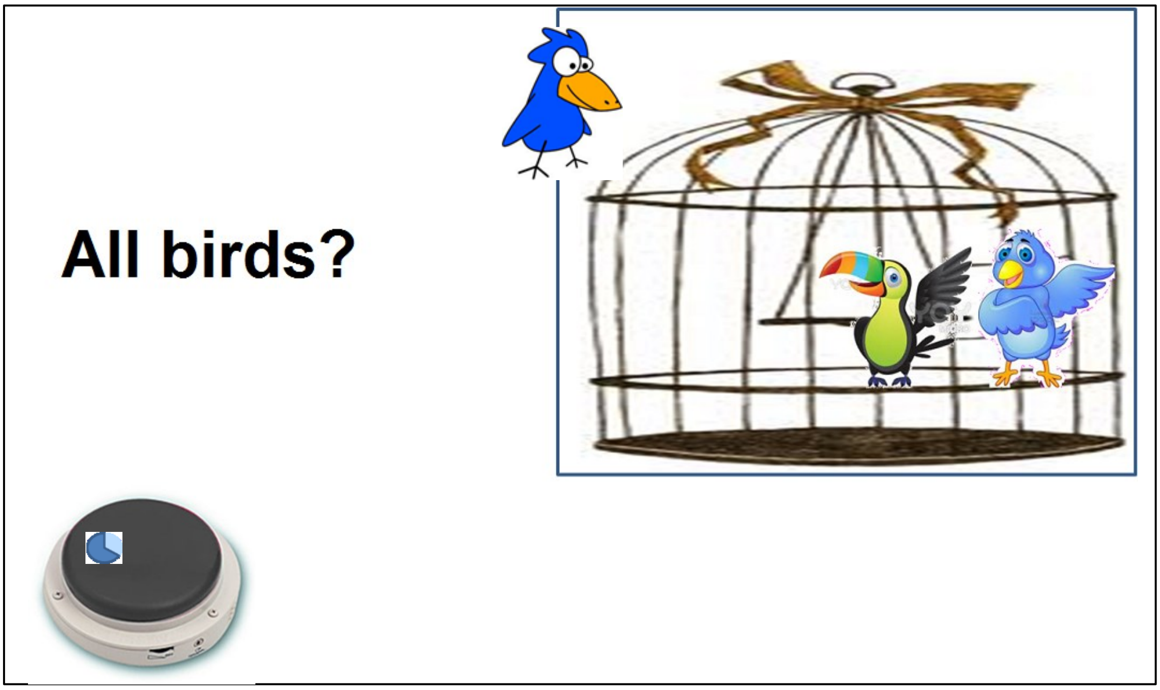


Figure 35 - Stimulus presentation for CATTEST. The blue bird (left) would appear in the left hand side of the display and float into the cage. The participant is required to respond with a YES if the display shows all birds and NO otherwise. In this case, the participant is responding with a YES. The virtual switch substitutes for a physical switch when a non-hand held device is used. When the EEG headset is used in single-mode, the blue pie indicates the strength of the selection. The light blue region increases clockwise with strength and at full strength, the dark blue region disappear leaving a twelve o'clock line (dark blue border). The strength of the selection is dependent upon the bite duration as well as clarity of the signal. Detectable bites sustained for a configurable duration is required for selection.

A set of stimulus (Figure 35) is produced for which the participant must provide a positive or negative response. The positive response is an active response which involves the actuation of a physical or virtual device. The negative response is a passive response requiring no action. 20 consecutive correct responses constitute a completely successful attempt by the child, which terminates the test. Otherwise, the trials (Figure 36 section 7.3.4) terminate after a block of 32 trials. A single session of two blocks of trials involving two different devices and tests are conducted within an hour. During the tests observations were made by the researcher regarding test response behaviour. Also notes were compiled after testing. Devices were run in order of increasing complexity. The stimulus for the CATTEST consists of images of birds and fruits (Quinn et al., 2004). Participants demonstrate ability to recognise a bird or fruit category by providing a positive response when they see either a bird go into a bird cage consisting of two different birds, or a fruit placed onto a fruit bowl holding a few different fruits. Conversely, a negative response is required when a bird is placed in the fruit bowl or a fruit is placed in the bird cage. No feedback is provided to indicate if the response is correct or incorrect but actuation of the virtual switch produces a click. Pre-test familiarization is provided to get the participants comfortable with the test. Participants are familiarized with a different set of birds and fruits which does provide feedback if the response is correct or incorrect. Once the participants indicate they are comfortable with the familiarization test, they are put on the real test.

When a mouse is being used, the participant would use a single mouse click for a positive response and do nothing for a negative response. When an eye-tracker is used, participants have to move a mouse cursor with their eyes to an image of a switch and "dwell" the cursor over the virtual switch for a period identified as the "dwell-time". When a head-tracker is used, participants have to move a mouse cursor using movement of their head to the virtual switch for the dwell-time.

This study also considers Signing as a method of input, and Signing provides a contrast to bio-modal inputs and the typical physical inputs. Signing in this case refers to a child who communicates using gestures to an interpreter. The gesture acts as the child's response to a test, and this response is then entered as a mouse input via the interpreter. Signing was used to confirm that a child was able to understand the test without the impact of dealing with a physical device in cases where there was uncertainty with the results. Signing was deemed to impose the least cognitive load for manipulating a device as the child has had years of using and developing it to use the parts of the body which were functionally capable for purpose and a carer who also was able to pick up on subtleties of the communication. The replacement of Signing by using a physical device loses some of those advantages.

8.3.3. Design

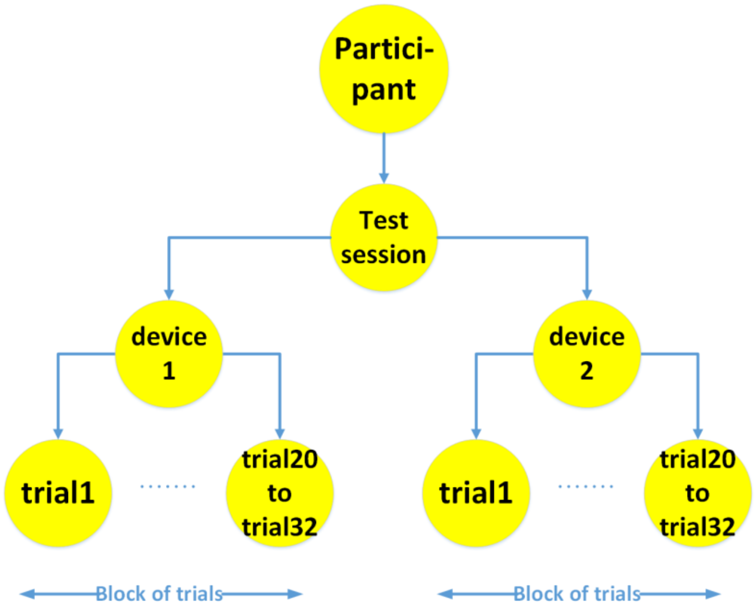


Figure 36 - Test design

Figure 36 show the test design. The experiment is a 2 x 32 within-subjects design for the maximum number of trials. For the participants who can only use one device due to impairments the design is 1 x 32. For each participant the following are components of the test:

- Device {(Signing), mouse/switch, eye-track, head-track, EEG, EEG and eye-track, EEG and head-track}. Signing is only used when no other devices can be used and does not increase the maximum count of devices for trials.
- Block {1 to a range between 20 and 32 trials}
- Test {CATTEST}

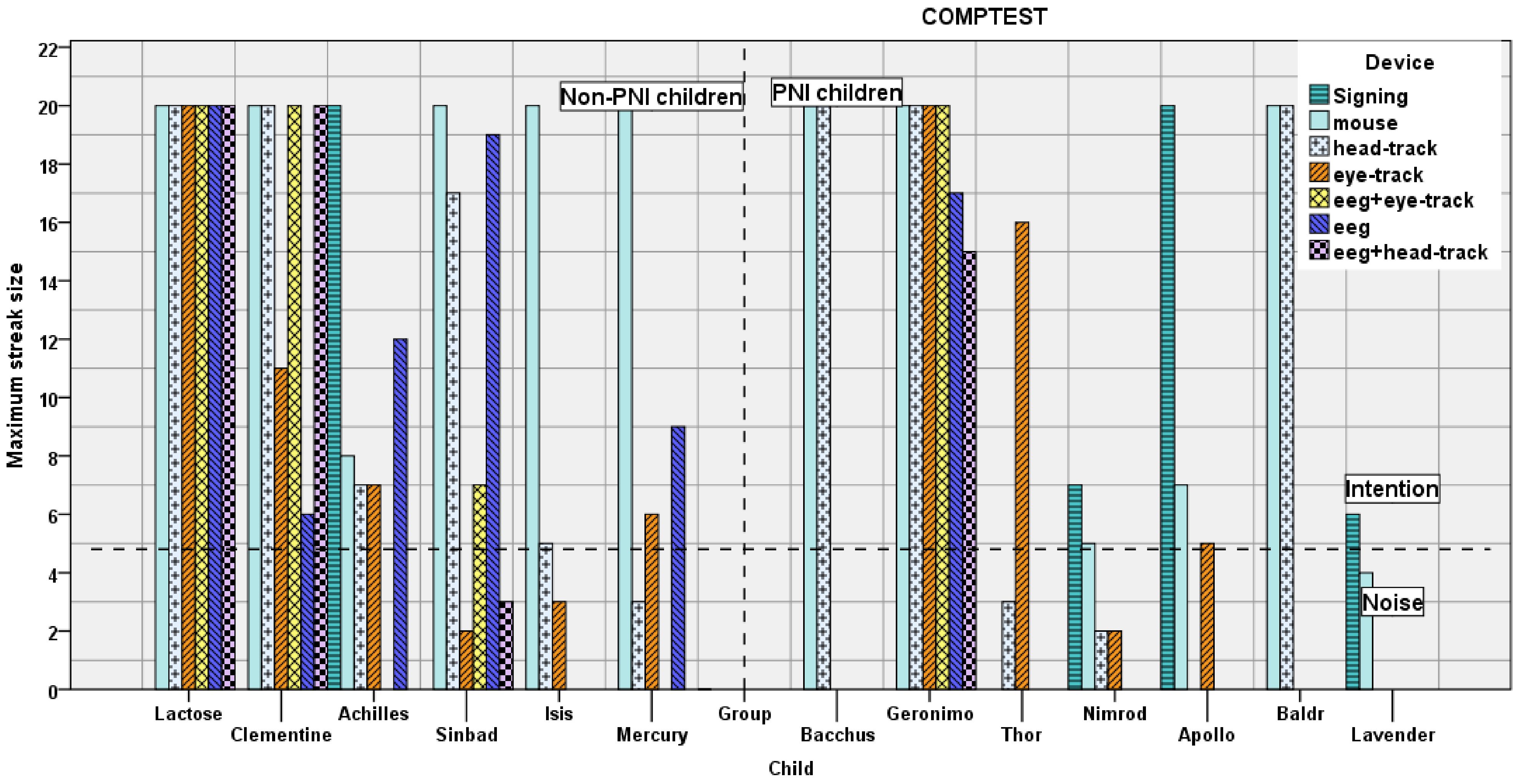
There are 18 participants giving a maximum total number of 18 x (2 x 32) trials. This is only the maximum total as it would depend on whether there was more than one appropriate device (1 or 2) or more than 20 trials (20 to 32).

8.3.4. Data capture

CATTEST results are represented as a 32-bit field, each bit representing an OK/NOK (not OK) outcome for a particular trial. The field is translated into a list of success and failure streaks. Failure streaks are suffixed with x. For example, for a list of 17 successes followed by 3 failures, 10 successes and 2 failures, the list is represented as {17, 3x, 10, 2x}. The consecutive successes and failure are referred to as success and failure streaks respectively. The entire list which captures an entire block of trials is referred to as an outcome event sequence. The maximum number of consecutive successes in the example list is 17.

8.4.Results

All results will be discussed further in the next section (8.5).


COMPTEST
Maximum streak size
0
2
4
6
8
10
12
14
16
18
20
22
Non-PNI children
PNI children
Device
Signing
mouse
head-track
eye-track
eeg+eye-track
eeg
eeg+head-track
Intention
Noise
Lactose
Clementine
Achilles
Sinbad
Isis
Mercury
Group
Bacchus
Geronimo
Thor
Nimrod
Apollo
Baldr
Lavender
Child

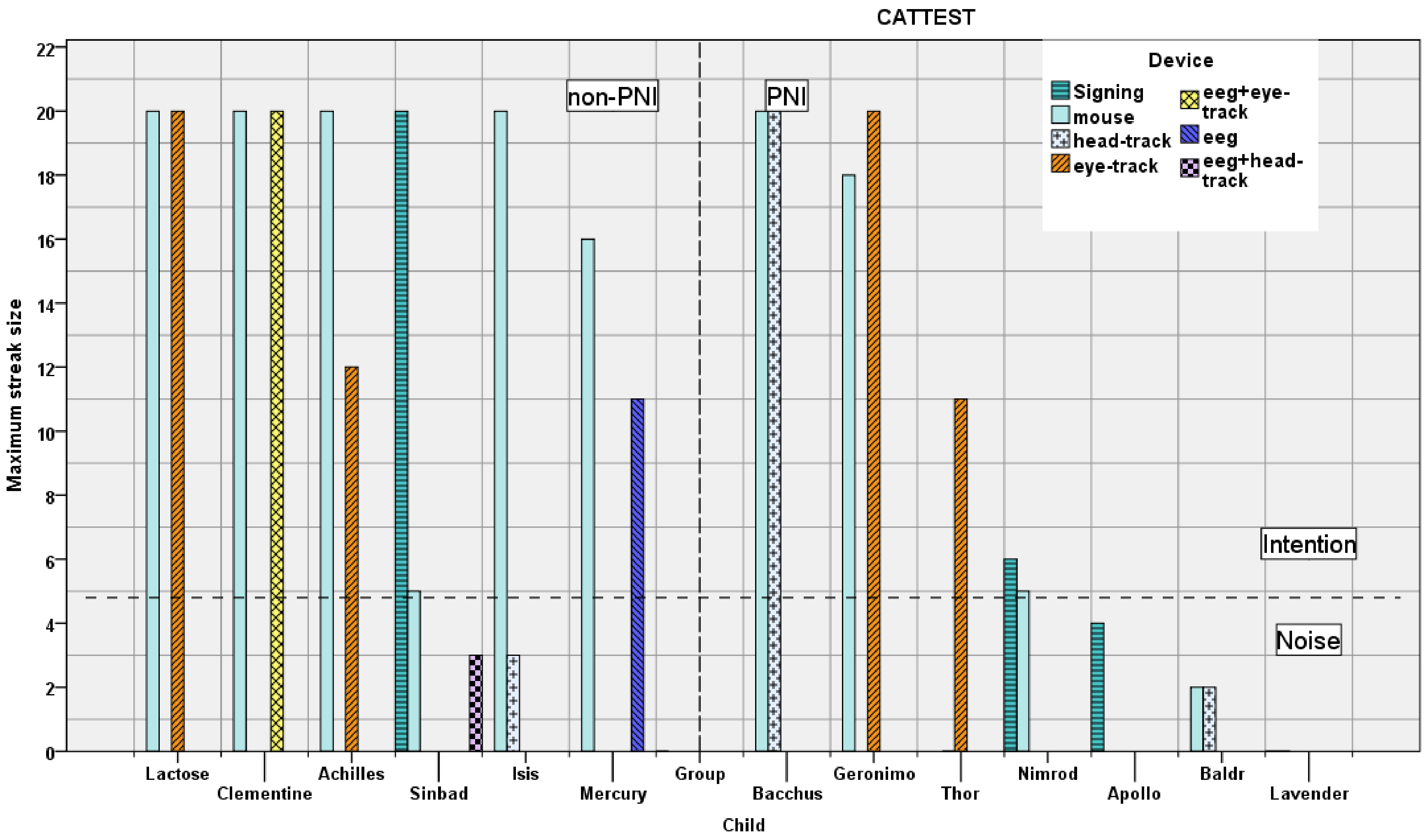


Figure 37 - Device competency of children evaluated using maximum streak as a measure in COMPTEST (top Pg. 163), CATTEST (bottom Pg. 164). The vertical dotted line separates non-PNI children (left) from PNI children (right). The dotted horizontal line represents a threshold that separates intention outcomes from possibly random outcomes.

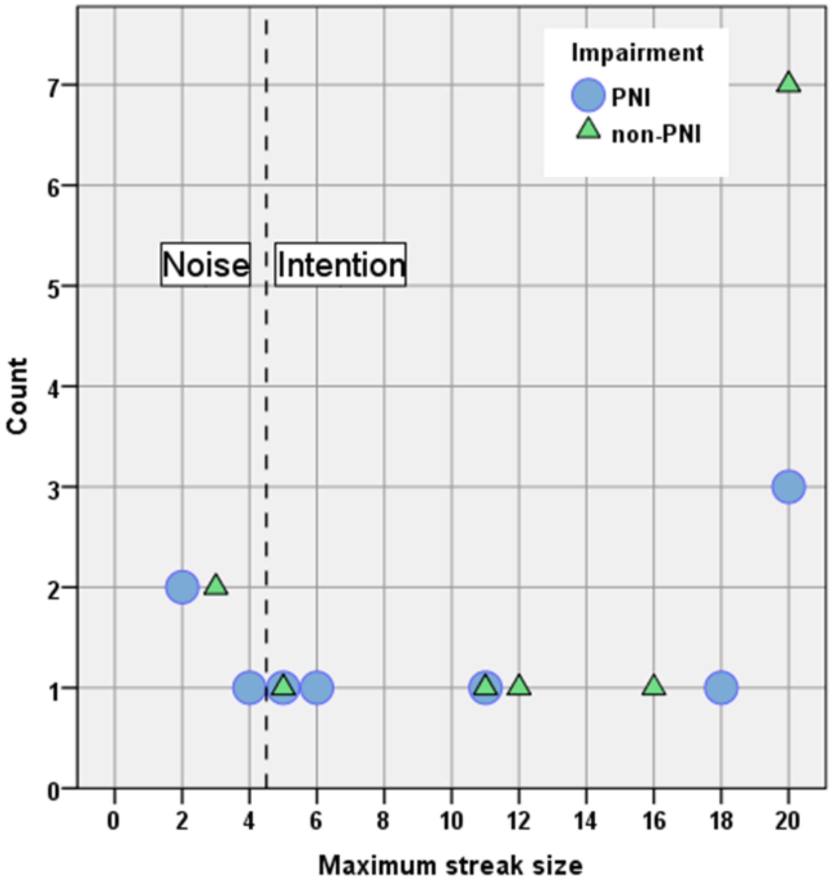


Figure 38 - Maximum streak size distribution for CATTEST. Results for PNI children are shown with circles and non-PNI children with triangles. Each count is a test run with a different device. Streak sizes of 0 are ignored as the participants did not engage with the test. The dotted vertical line represents a threshold that separates intention outcomes from possibly random outcomes.

8.4.1. Results from the experiment with PNI Children

Figure 38 provides an overview of the maximum streak size data for success streaks for and CATTEST. Each maximum streak size represents a block of trials that a child has had with a specific device. The count therefore shows the number of blocks that achieved a specific maximum. Streak sizes of 0 were ignored as they provided no inputs. 7 PNI participants were tested with a maximum of 2 physical devices but some were unable to engage with either the device or test leaving the total count (10) less than 14 (7 x 2). The distribution trend of the CATTEST figure has three ranges; a large high end, a thin middle range and a cluster at the low end (maximum-streak of 20, 11 to 18, 2 to 6). This pattern is due to the design of the tests. The high count of the maximum streak size of 20 is a tail effect due to the success cut-off for the test at a streak size of 20. If the cut-off was some larger number the count would be lower, spreading among the other possible values that were above 20.

Figure 39 shows the number of occurrences of a particular streak size (taken for all the tests) for the streak sizes ranging between 1 and 19. The results of children who succeeded in achieving the target (streak size of 20) and the children who could not engage in the test were not used in the plots as they do not represent a full test run of 32 trials. None of the children had 12 failures and 20 successes. The scatter plots gives the count of different streak categories (a specific streak size formed a category). PNI children were represented by blue diamonds.

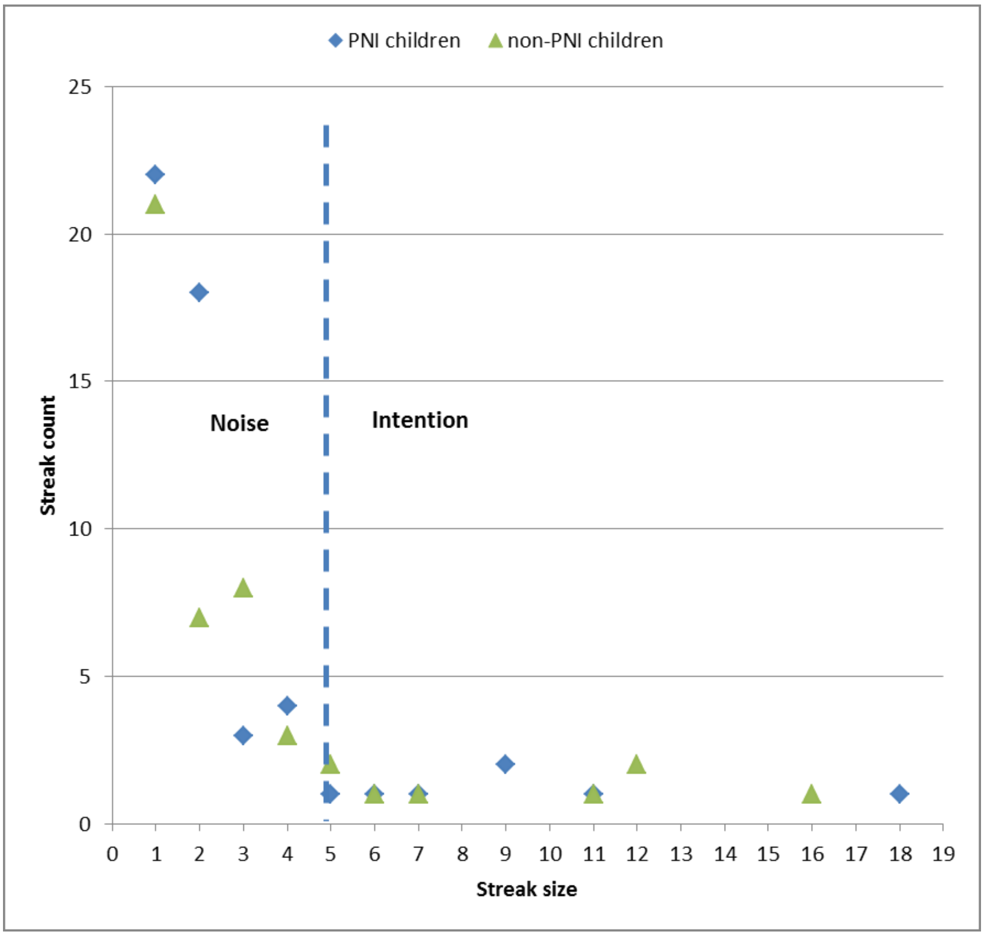


Figure 39 - Success streak size distribution of CATTEST (right) (0 < streak size < 20). Both plots are scatter plots of two series (PNI children, non-PNI children)

Streak sizes of 1 and 2 are very common because the test program generates that noise for random responses. Streak sizes between 5 and 19 are much less likely. A reasonable point to choose the noise threshold starts from the “knee” of the data as the data flattens out. In order to be compatible with a popular cognitive test (Wisconsin Card Sorting Test (WCST) (D. A. B. Grant, Esta A., 1981, 2003)) the noise threshold was chosen at 5 so that streak sizes of below 5 would be treated as noise and for sizes of 5 and above as demonstrating intention. The term intention is used to mean that the action is not the result of a reflex action but arising from a conscious decision (Libet et al., 1993).

Table 13 - COMPTEST and CATTEST success-streak size cumulative distribution

Streak size	COMPTEST		CATTEST	
	PNI (Cumulative Percent)	Non-PNI (Cumulative Percent)	PNI (Cumulative Percent)	Non-PNI (Cumulative Percent)
3	75.0	70.8	79.6	76.6
4	83.3	78.5	87.0	83.0
5	89.3	83.1	88.9	87.2
6	91.7	89.2	90.7	89.4

Table 13 shows the variation of the cumulative distribution of the success streak sizes that range between 3 to 6 which is around the size chosen to represent the noise threshold (5). For CATTEST the probability of getting a streak size less than 5 is around 89% for PNI children. This confidence level seems reasonable.

Figure 37 shows the maximum streak size achieved for each run of CATTEST (bottom figure) using a maximum of two different physical devices. The devices were among the best devices (determined from COMPTEST (top)) available for the child. The right plot shows results for PNI children. The children are sorted to show a general trend of ability starting from the left in the CATTEST figure. As the comparison is done for ability with physical devices, we can examine the trend discounting Signing (blue with horizontal lines). A child is considered to have the potential to carry out a target cognitive test (ECDT) if the child manages to achieve the test target streak size of 20 at least once. Streak sizes below 20 suggest that it is possible for the child to fail ECDT, due to the inability to sustain performance with the device used. This result does not conclude definite cognitive inability. 2 children (Bacchus, Geronimo) have passed the test and the others have failed. 1 child (Thor) has only one device that affords accessibility as other devices could not be used due to impairments. Thor showed cognitive ability in a more complex cognitive test using Signing and was not tested in this test with Signing.

2 children (Bacchus, Geronimo) used non-hand held devices and passed the test. None of the devices were hybrid devices.

The cognitive ability for PNI children can be divided into three groups:

- Children who cannot engage in the test
- Children who can engage but not pass the test
- Children who pass the test

Children who cannot engage in the test are distinguished by results that are noisy, falling below the noise threshold (Apollo, Baldr, Lavender). Children who can engage but not pass the test are distinguished by results that are above the noise threshold but not achieved the target criterion of a streak size of 20 using Signing (Nimrod). Children who pass the test have achieved the target criterion of a streak size of 20 (Geronimo, Bacchus, Thor). Thor is an exceptional case where he passed a more complex cognitive test (ECDT) using Signing.

Figure 40 shows a comparison of the maximum streak sizes achieved by the PNI children (right) while taking the COMPTEST and CATTEST. The devices that best suited the respective child were used and in most cases was the mouse. Signing was used when the results using physical devices were poor. The results are arranged in order of children achieving the best performances in COMPTEST. Their CATTEST results were equal (1 child) or lower (the rest of the children) in streak size than COMPTEST results showing that CATTEST has an effect on the children.

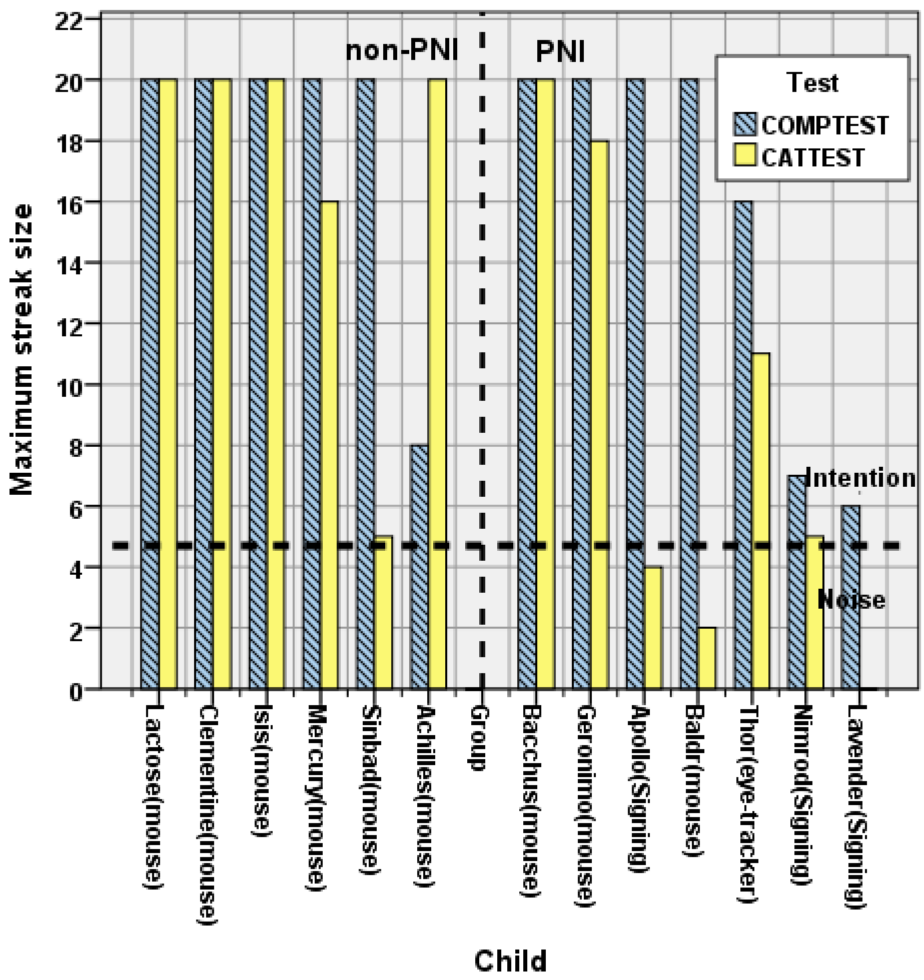


Figure 40 - Comparison of motor skills test (COMPTEST) and light cognitive test (CATTEST) using the best device. The vertical dotted line separates non-PNI children (left) from PNI children (right). The dotted horizontal line represents a threshold that separates intention outcomes from possibly random outcomes.

The mean maximum-streak size in CATTEST (M=8.57, SD=7.913) was lower than the mean maximum-streak size in COMPTEST (M=15.57, SD=6.373) resulting in a mean decrease (M=7, SD=7.141) in the ceiling (maximum-streak size). This decrease was statistically significant, t (6) =2.593, p<0.041, two-tailed. However, the sample size is small.

Figure 41 shows for each child, how motor control changes in the face of added cognitive load. The cognitive load change is represented by COMPTEST and CATTEST. The motor control changes are represented by the different devices used. Only devices where the child has proven stable ability are compared. The devices that are deemed to be stable are pairs of devices that achieved roughly equal streak sizes (differing within a streak size of 2) in COMPTEST above the noise threshold. A good benchmark was chosen as a basis for comparison. For one of the devices used, they must have achieved the target criterion in COMPTEST of a streak size of 20 or if they have not, then Signing was used.

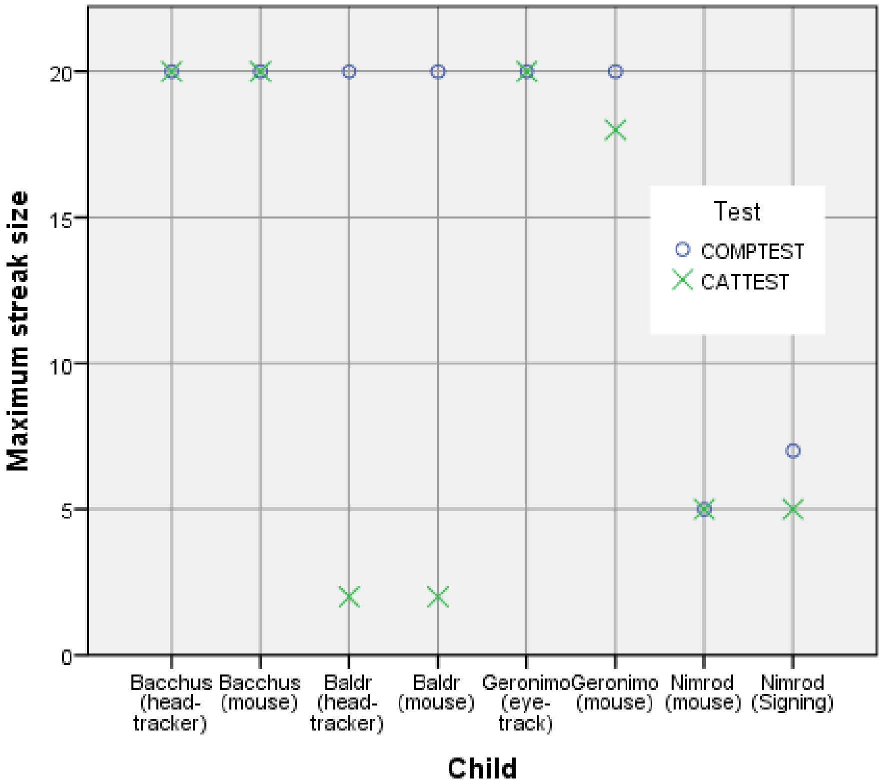


Figure 41 - Comparison of motor control performance with cognitive load change for PNI children. Each child has a pair of devices that is represented adjacently (device variation). The circle and cross represent different cognitive load conditions (test condition variation).

A line joining the two circles or cross for the pair of device each child use is almost horizontal showing small changes in motor control. The difference in streak size between cross and circle show different cognitive loads. This shows that the devices chosen by COMPTEST were valid for use. This result shows that the 4 children were probably not affected by the devices. One child (Baldr) was significantly affected by the change in cognitive load.

Apollo

Apollo was unable to do the test. The use of Signing showed that cognitively he was unable to do the test, achieving a maximum streak (4) which was in the "noise" region.

Bacchus

Bacchus had no problems with the test and managed to use the devices he was successful with in the COMPTEST (the mouse and the head-tracker) to do the test. The use of the devices did not hamper Bacchus from carrying out the test.

Baldr

Baldr was cognitively unable to do the test. COMPTEST established that he could use both the mouse and a head-tracker competently. His CATTEST results for both devices were identical and indicate results that could be due to random responses.

Geronimo

Geronimo had no problems with the test. His results using the eye-tracker was perfect but his results using the mouse was slightly less. It is probable that he had a momentary distraction when he was taking the test with a mouse.

Lavender

Lavender could not focus on the test. Her short attention spell prevented her from engaging in the test.

Nimrod

Nimrod was not able to attain the requirements of the test (streak size of 20). The results using Signing was poor (6). This result was similar to his result for COMPTEST (7). This may be a similar situation for the COMPTEST in which he gets confused after a short while.

Thor

Although Thor did not attain the requirements of the test (streak size of 20) the results (11) were substantially above the “noise” region (4). As he was able to do a more complex test (ECDT in chapter 9), it would suggest the possibility that he was hampered from attaining the requriements by the device he was using and not by cognitive abilities.

8.4.2. Results from the experiment with non-PNI Children

The maximum streak distribution of non-PNI children (Figure 38, triangles) show a relatively flat initial distribution from a streak size of 3 to 16 followed by a sharp rise at a streak size of 20. The high count of the streak size (20) is due to a “tail-end” effect where the test is terminated once 20 are reached. The count is a group count which represents streak sizes above 20. From this we see that further analysis of the data for maximum streak sizes of 3 to 16 is representative.

The streak distribution of non-PNI children (Figure 39, triangles) show high values for streaks of 1 to 4, and a relatively consistent low values for streaks thereafter from 5 to 16. This suggests that the probability of high value streaks decreases to a stable value after 5. We see the corresponding cumulative percentages in Table 13 show that for non-PNI children, the percentage is around 87%.The confidence level established is reasonable to have a noise threshold with a streak value of 5.

The results of CATTEST ability for non-PNI children is shown in Figure 37 (bottom) on the right side of the vertical dotted separator. 4 out of 6 non-PNI children achieved success in this test. One child who failed the test (Mercury) had a high maximum streak size with a mouse and was not re-tested using Signing. The other child (Sinbad) passed the test using Signing. 2 children (Lactose, Clementine) used non-hand held devices successfully in this test. 1 child (Clementine) of the 2 used a hybrid device.

Figure 40 (left) compares the maximum streak results of non-PNI children in COMPTEST and CATTEST using their best device. 3 children had the same results, 2 had poorer and 1 had better results in CATTEST. The results do not show that the children were significantly affected by CATTEST.

The mean maximum-streak size in CATTEST (M=16.83, SD=6.014) was lower than the mean maximum-streak size in COMPTEST (M=18, SD=4.899) resulting in a mean decrease (M=1.167, SD=8.681) in the ceiling (maximum-streak size). This decrease was not statistically significant, $t(5) = 0.329$, $p < 0.755$, two-tailed. However, the sample size is small.

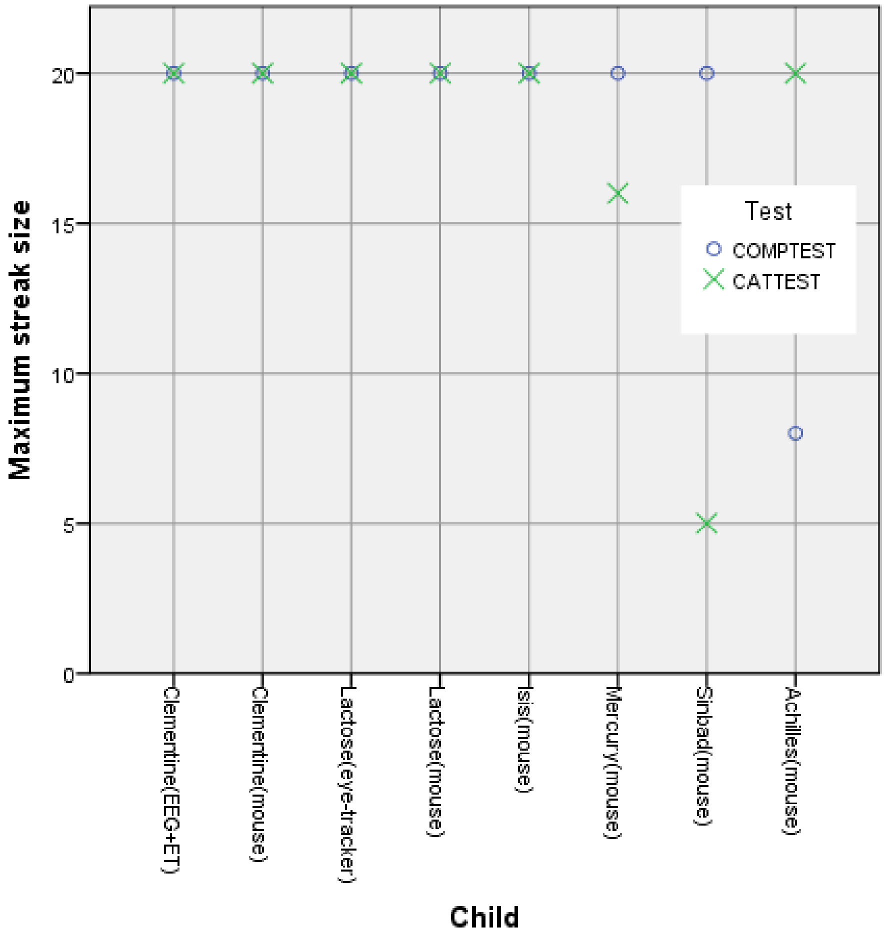


Figure 42 - Comparison of motor control performance with cognitive load change for non-PNI children. Each child has one or two devices being tested. The circle and cross represent different cognitive load conditions (test condition variation).

Figure 42 shows for each child, how motor control changes in the face of added cognitive load. Only devices where the child has proven ability (achieved target criterion in test with streak size of 20) are compared. 3 children (Clementine, Lactose, and Isis) prove stable ability with their devices, shown by results with cross and circle overlapping each other. The others only prove ability in one test.

Achilles

Achilles had no problems with the test using a mouse. He achieved better results than the simpler COMPTEST which suggested that the possibility of lack of focus for the COMPTEST. He was not able to master the eye-tracker well enough to succeed with it in the test.

Clementine

Clementine had no problems with the test using either a mouse or the multi-mode EEG and eye-track hybrid. This is consistent with her demonstrated abilities in previous testing with COMPTEST.

Isis

Isis had no problems with the test using a mouse. She had poor results with the head-tracker (streak size of 3) consistent with the results in COMPTEST (5).

Lactose

Clementine had no problems with the test using either a mouse or the eye-tracker. This is consistent with her demonstrated abilities in previous testing with COMPTEST.

Mercury

Mercury did not achieve the requirements of the test (20) but had a high result (16) using a mouse. Previous testing with COMPTEST showed he had no problems using a mouse (20). The possibility exists that both the cognitive requirements of the test and the use of the device is demanding enough to produce a degradation of the result. He had poor results with the EEG headset (11) consistent with the results in COMPTEST (9).

Sinbad

Sinbad had very poor results (5) with the test using a mouse. He took the test subsequently using Signing and showed that he had no problems with the cognitive elements of the test. The very poor result suggests the possibility of lack of motivation on this occasion of taking the test.

8.4.3. Results from the experiment with adults

Table 14 - CATTEST results for non-PNI adults. The adults obtained a streak-size of 20 without any errors for all devices.

Name	Device
Aladdin	Mouse
	Eye-tracker
Anubis	Mouse
	Eye-tracker
Atlas	Mouse
	EEG headset + eye-tracker
Cupid	Mouse
	Eye-tracker
Damocles	Mouse
	EEG headset + eye-tracker

Table 14 shows the results of CATTEST for adults. The results are uniform with all adults being able to do the test perfectly with the two devices allocated from the COMPTEST range. All adults had the mouse as their best device. Where there was more than one best device, the eye-tracker was used. In some cases, participants voiced a preference for the multi-mode combination of EEG headset and eye-tracker and the device was used instead of the eye-tracker.

8.5. Discussion

8.5.1. Variation of maximum streak distributions between PNI children and non-PNI children

The maximum streak distribution show higher values for the extreme streak value (streak value 20) (in Figure 38) because of the tail-end cutoff of the test design, the explanantion of which follows. Termination of the test at a streak value of 20 causes a discontinuity. If the test were not terminated at 20, a spread of values of above 20 would be obtained. From that point of view, the count of children having a streak score of 20 is a group count of children with a 20 plus maximum streak. The group count result in the higher count for the streak value of 20. The higher ability of the non-PNI children population hence results in a higher number who achieve the target of 20 than the PNI children. Apart from the extreme streak values, the distribution is fairly uniform. As the streak value of 20 is not representative of the distribution, only the streak values below 20 should be used for further analysis.

8.5.2. Variation of streak size distributions between PNI children and non-PNI children

The trend for the streaks distribution is for a much higher count of streaks for low streak values (from 1 to 4) and thereafter (5 and above), the count drops to a low almost constant value Figure 39. The streak size distributions contain some devices that participants did not control very well. These devices provide the data that forms the noise and help to establish where a sensible noise threshold should be. The combined effect of user, devices and test appear to stabilise the streak value counts to a low value when the streak value is around 5. An examination of the range of values of the cumulative distribution formed by PNI and non-PNI children for a streak size of 5 for CATTEST show a range (83% to 87%) comparable to that found for COMPTEST (83% to 89%). The noise threshold placed at a streak value of 5 is therefore reasonable.

8.5.3. Effects of user, test and device between PNI children, non-PNI children and adults

The results suggest that for 4 (Bacchus, Baldr, Geronimo, Nimrod) out of 7 PNI children, the ability with the device was good (Figure 41) although cognitive ability could be poor. All non-PNI children have good motor ability with their devices but only 3 show stable performances (Figure 42). Adults have the best ability, being able to do the test with all their devices perfectly (Table 14).

Four PNI children demonstrated good motor ability for two tests (COMPTEST, CATTEST). Demonstration consists of consistent ability with a trusted benchmark in both tests which were established by achieving target criterion in COMPTEST or the use of Signing. Nimrod was the only child who had poor results due to cognitive problems but established good motor ability with a mouse using Signing for both tests. Baldr demonstrated good ability with the devices in COMPTEST but problems with CATTEST.

Two PNI children who demonstrated poor device ability were Thor and Apollo. Thor demonstrated good cognitive ability with COMPTEST but poor cognitive ability with CATTEST using Signing. Thor demonstrated encouraging results in COMPTEST which became poorer in CATTEST using an eye-tracker. However, Thor had involuntary torso movements due to his impairment which made the eye-tracker lose track. Thor was not tested using Signing on both COMPTEST and CATTEST but in the most complex cognitive test (ECDT) of the test suite he was tested using Signing and showed no cognitive problems. However, his poorer results with CATTEST compared to COMPTEST suggest that an increase in cognitive load coupled with the need to manage his disability when using the eye-tracker decreases performance. This child would have failed the original ECDT if he used the only device that provided access.

One PNI child (Lavender) could not demonstrate good motor ability as she had cognitive problems with all tests (Figure 37, top). Although she demonstrated results with a switch that were fairly consistent with Signing, her switch results fell below the noise threshold and had to be discounted.

For non-PNI children, all children show ability with at least one device (Figure 42). For this test, all the girls show consistent performance across both tests and the boys have inconsistent performances, failing one of the tests. The performances of the boys may be due to fatigue and or motivation. Bruckman warns of short attention span that researchers of Human-Computer Interaction (HCI) for children of ages 2 to 7 have to cater for (Bruckman, Bandlow, & Forte, 2002; Piaget, 1970; Xie & Salvendy, 2000) which suggest an interaction effect within the user, device, and test triad. Possible interaction effects that are a combination of user with test and or device is possible. Where results have been poor (Sinbad, Achilles), the children were re-tested using Signing and both succeed in passing the failed tests then. The suggestion is that possibly another device may be more suitable for these children. Also there appear to be more instability with results using physical devices as the cognitive load is increased.

With the use of novel non-hand held devices, some PNI children have to use them because their impairments exclude them from using anything else. Non-PNI children are still developing (Gerber, Wilks, & Erdie-Lalena, 2010; Thelen, 1995) the skills required to use them and need the practice required to remove the need for cognitive association to manage the device(Ericsson, 2006). Adults have the latent skills to manage the device and have a preference for some hybrid devices as they have the potential to be task-saving devices.

Overall, results show that the increase in cognitive load imposed by CATTEST have enough of an influence in decreasing motor performance for some children.

8.5.4. Variation of cognitive ability between PNI children, non-PNI children and adults

The results show that PNI children performing more poorly than non-PNI children. The adults have no problems with the test (Figure 37). The increase in cognitive complexity in CATTEST appears to have more impact on PNI children than non-PNI children (Figure 40). CATTEST was designed so that a 3 year old should be able to pass the test (Dunham & Dunham, 1995) and the non-PNI control group verify it.

PNI children have three groups of results:

- Unable to engage test
- Able to engage test but cannot sustain performance
- Able to achieve required performance for target test (ECDT)

The children who were not able to engage in the test had attentional difficulties (McCall et al., 1977; Wolff, 1965) or did not seem to understand the test. The children who could not sustain performance could possibly not have developed a large enough working memory (Atkinson & Shiffrin, 1971; Baddeley, 1983). The hypothesis would be that these children would also not be able to pass ECDT.

All non-PNI children were shown to be able to pass CATTEST including those who used Signing except Mercury whose results were high and Signing was not performed.

8.6. Conclusions

The performance results (8.5.3) with maximum streak sizes show that generally, best devices chosen in COMPTEST produce consistent results when tested under increased cognitive load. The differentiation between motor ability and cognitive ability was achieved using trusted benchmarks that were approved by COMPTEST. Children are able to either achieve the test criterion with a physical device or the two devices used for CATTEST produce very similar performances. Some children can be affected by the devices while undertaking the test.

The poorer performance of PNI children on CATTEST, as measured by lower streak sizes than that obtained in COMPTEST, corroborates the hypothesis that CATTEST lowers the maximum performance that can be attained by a PNI child. The noise threshold of 5 established by COMPTEST (a confidence level of between 83% and 89%) was re-evaluated in CATTEST and found to be reasonable (a confidence level of between 83% and 87%). The result strengthens the use of a streak level of 5 as the noise threshold. We can also speculate that a lowering of the test performance ceiling increases the effective confidence level of CATTEST results above the noise level because the mean of the results are lowered nearer to the noise threshold. Another way of viewing this is that the noise will now engulf more intentional responses that are now relatively weaker in performance.

The t-test comparisons of CATTEST and COMPTEST maximum streak results provide further evidence that CATTEST lowers the ceiling of performance for PNI children using their best device but sample size is small. CATTEST is simple enough for cognitively unimpaired children and the t-test results show that there is no lowering of ceiling for the non-PNI children using the mouse but sample size is again small. There was not enough data to check if the change in ceiling is solely due to the test or combination interaction effects in the user-device-test triad. An improvement to the test design would be to use device and cognitive load as two factors in a two-way, in-between subjects

ANOVA test. The case of Sinbad succeeding with Signing but obtaining a poor result in CATTEST with a mouse suggests that the test by itself was not a problem. Since Sinbad had no problems with a mouse in COMPTEST the poor result is not limited to just the device. This opens up the possibility that the poor results could be due to the interaction of the user, device, and test triad. While we can speculate that both test and device effect performance by causing fatigue, this study raises the question of fatigue discussed later in section 11.1. Although the single instance of Sinbad performing very poorly with the mouse but perfectly with Signing is not conclusive, this opens up the possibility that a device used in a test can have significant impact on the results. It was also speculated that the device impacts become visible in the results only for participants where the cognitive load was high enough to cause an effect. In other words, cognitive complexity is relative to the participant (Rubio et al., 2004).

The CATTEST results show that PNI children can fail in the test. Comparing those results with that of the control groups, it was found that those groups are relatively unaffected by the test. The control groups tended to all achieve the criterion performance (streak-size = 20) with at least one device. Whilst the test appears to not have sufficient cognitive loading to affect the control groups, there is an effect on PNI children. These results provide support for the hypothesis that CATTEST can be used as a failure filter for ECDT results of PNI children.

Signing again proved to be the better "device" in being able to get better results than a physical device. The hypothesis is that Signing incurs less fatigue than the other physical devices. Support for this hypothesis was obtained in the next test (ECDT) which is a more complex cognitive test. In ECDT, Signing can be used in the cases where children fail to overcome the cognitive and physical challenges posed by a physical device.

9. MAIN PILOT STUDY: Psychological test (ECDT)

9.1.Introduction

9.1.1. General

This is the final test of a series of three (section 6.1.4.3). In this test, the cognitive skills in of PNI children will be assessed using a learning test (ECDT) (Khan, 2009; Sharp & Evans, 1981). The background to this test is covered in section 2.4.3. The test is essentially a non-verbal psychological test for PNI children to assess fundamental learning. It is regarded as a potential stakeholder test which means that it is being trialled for use to determine if PNI children are candidates for non-special needs schools. This test was only previously accessible via a mouse or switch. The version of ECDT being used for this study extends the capability of the previous test to use other devices besides the mouse/switch with the hope that it may extend accessibility. In addition, this test captures more data than the previous test. This test not only looks at whether the test can be made more accessible through the use of new devices but also if the use of a device presents a negative impact on the test.

The details of how the test is setup and how the test is run will be provided in a subsequent section (9.3). One of the main problems with PNI children is the lack of communication as a result of their impairments (Pennington & McConachie, 2001). Even when a PNI child is able to engage an input device, the responses may not reflect true intention due to involuntary muscular effects section 2.2. In order to judge accessibility to a test becomes a more complicated procedure of assessing intention in addition to ability. Three sets of controls (adults, non-PNI children, capable PNI child) are therefore added to this study to provide an indication of the range of expected results and benchmarks for the testing. The controls validate the setup. The controls look at how PNI population sample differs from a non-PNI sample. Some changes to the test (ECDT) were made so that it is able to assess intention better. Previously, ECDT used a success criterion which was 10 consecutive successes to serve the two purposes of assessing intention and ability to learn. This effectively restricted the metrics of the test to just one of indicating success. The work done in previous tests (COMPTEST) provides the ability to discriminate different levels of consecutive successes (streaks) with possible random inputs (noise). By assimilating this work with ECDT, the consecutive successes can provide finer levels of information. Inspiration on the possibility of other metrics can be drawn from a cognitive test that is similar to ECDT called the Wisconsin Card Sorting Test (WCST) (D. A. B. Grant, Esta A., 1981, 2003).

9.1.2. Adding to ECDT metrics

In order to distinguish device impacts on the results of the test (ECDT), the results can be interpreted in two ways; a low-level interpretation which considers performances under cognitive load (and has a stronger focus on device impacts) and a high-level interpretation which considers the psychology of the response (with a stronger focus on the cognitive objectives of the test). We first consider the low-level view from the COMPTEST study and then the high-level view.

Table 15 - Thresholds of streak-sizes using high and low level interpretations. The high-level interpretation is taken from the Wisconsin Card Sorting Test (WCST). Where there are no specific interpretations of the measure, the ranges are marked as unspecified. Conceptual level responses: Size ≥ 3; Failure to maintain set: 5 ≤ Size < 10.

Streak-size	Low-level interpretation	High-level interpretation (WCST)	Modified ECDT
0	Physical disengagement	Unspecified	Physical disengagement
1	Noise		Noise
2			
3		Conceptual level responses	
4			
5	Intention	Failure to maintain set	Failure to maintain set
6			
7			
8			
9			
10	Success	Success	Success

9.1.2.1. *Low-level interpretation*

From the view-point of the low level interpretation, the test (ECDT) is a cognitive test that imposes a cognitive load on the participant. The participant responds to the load with an attainment level that can be measured by the use of streaks. The higher the streaks the greater the effort placed on the response. As the cognitive load increases the streaks are anticipated to get lower (section 8.4.1). This follows from the assumption that a test outcome depend on the cognitive capability of the user, part of which is dedicated to manipulating the device and the rest to the cognitive demands of the test (Karni & Bertini, 1997; Karni et al., 1998; VanLehn, 1996) (a device-test cognitive load split assumption). If the effort required working the device and dealing with the test exceeds the capability of the user, the user will fail the test. The capability of the user for those tasks may change at different times depending on fatigue levels (Kahol et al., 2008). A second assumption is used in this study to demonstrate the split in cognitive effort between device and test. It is assumed that Signing (Iverson & Braddock, 2011) would outperform the physical devices used and demands the optimal amount of cognitive effort to execute compared to a physical device. For participants that fail the test with a physical device, Signing can be used to demonstrate that the effort imposed by both device and test exceeded that with Signing as a proxy for the physical device. The device-test cognitive load split assumption was made on the assumption that significant cognitive effort was not expended elsewhere due to distractions (Lamble, Kauranen, Laakso, & Summala, 1999). In the case of ECDT, when a participant fixates on an incorrect response, that cognitive effort can be viewed as a distraction. The distraction causes the cognitive attention to be split between dimensions resulting in measurement changes (section 6.1.5)(Irons, Folk, & Remington, 2012). However the low-level interpretation is not concerned with the changes between dimensions but just the quality of the measurements within each dimension. Changes between dimensions would be in the domain of the high-level interpretation. The additional consideration of distraction in the device-test cognitive load split hypothesis extends the hypothesis to generally be a consideration of simultaneous attentional control splits (Folk et al., 1994; Pashler, Johnston, & Ruthruff, 2001).

Table 15 maps the low-level interpretation with the streak-size in the second column. The low-level interpretation is generally an exclusive one in that meanings are assigned to specific ranges. 0 indicates that the participant was not physically engaged in the test, 1 to 4 indicates noise which is a random outcome from which no reliable information can be obtained, 5 to 10 indicates intention, and 10 indicates when the test is terminated due to meeting the success criterion.

9.1.2.2. High-level interpretation

The high-level interpretation can be equally applied on the maximum-streak measure. Although ECDT does not have thresholds other than the success criterion, similar tests (WCST) to ECDT, establish a set of psychological thresholds which we can compare with. The main one in WCST is the success criterion which indicates when confidence is obtained that the participant has learnt a solution. Two others are when the participant starts to demonstrate the ability to recognize the correct dimension (conceptual level responses) and when the participant starts to establish that the solution is specific to the dimension (set maintenance). The order in which the different thresholds would occur would be that conceptual awareness takes place first, followed by set maintenance and finally the success criterion is reached. In WCST, the threshold levels are set at the following streak sizes; for conceptual awareness 3, for set maintenance 5 and for the success criterion a size of 10(Nadeau, Routhier, & Tessier, 2008). Table 15 maps the high-level interpretation (WCST) to the streak size in the third column. The high-level interpretation is generally inclusive. 0 to 2 have no special significance in WCST and in the case of ECDT, 0 can be used to indicate disengagement and 1 to 10 are all the number of correct responses made. 3 to 10 demonstrates that the participant is demonstrating that they are choosing the right solution; 5 to 9 indicates that they have made an error before realizing the objective and have not maintained the required performance; 10 indicates the required criterion for success.

As indicated by the section on low-level interpretation (9.1.2.1), the interpretation is not limited to a single dimension. Observation of attentional changes between dimensions helps us determine how a PNI population sample differs from a non-PNI sample.

9.1.2.3. Rationalizing high and low-level interpretations

In order to reconcile the low-level interpretation with the high-level one, we first project the case where the low-level impacts are absent and only the high-level consequences on the outcome persist and for the latter part consider the impacts of both.

In the first instance, if the device was perfect (the low-level impacts would be absent), the participant fails to achieve the success criterion due to cognitive problems solely associated with the test objectives which is the high-level interpretation. The success criterion, the failure to maintain set range, the conceptual level response range, would apply.

In a practical situation, some level of cognitive effort is spent on manipulating the input device. It is hence possible that a child fails to achieve the test criterion (target streak size) because of the cognitive load for manipulating the device in conjunction with the cognitive load needed to resolve the test (ECDT) solution. Whereas the high-level interpretation would have focused on attributing outcomes solely to cognitive inability to resolve the problem posed by the test (ECDT), the low-level interpretation suggests that the device plays an equal part. In this case, there is a difference in interpretation between streak sizes of 1 to 5 which is considered as noise in the low-level interpretation and the sizes of 3 and above which is considered as conceptual level responses in the high-level interpretation. A rationalized metric would mean that the conceptual level responses of streak sizes of 3 and 4 cannot be used as there is an uncertainty to the quality of information shows the result of rationalizing the high and low interpretations in the fourth column to provide a modified interpretation for ECDT results.

9.2.Aim

The primary aim of this study is to find out if the outcomes of PNI participants taking a specific cognitive test (ECDT) are affected by the device that is used by comparing the results of two devices that was chosen from previous evaluations (COMPTEST and CATTEST). This test mainly addresses research question RQ1.3 highlighted below.

Research question
RQ1.3. Does a PNI child have sufficient motor control to use the device and undertake a learning test at the same time?

9.3. Method

9.3.1. Participants

Participants used in this study are common to all main study tests and are documented in section 6.1.6.

9.3.2. Procedure

The participants were tested in a room (located in a school) equipped with a laptop, separate screen monitor and hand-held and non-hand held devices for interaction with the software tests running on the laptop. The screen monitor is arranged side to side with the laptop so that the participants with a view to the screen monitor were seated beside the researcher who had a view to the laptop. The eye-tracker was mounted below the screen of the monitor using magnetic mounts. The head-tracker used a remote web camera enabling tracking from the monitor.

A presentation set of two stimuli (cards) were produced on the screen (Figure 43) and the participant must choose one by providing a positive or negative response to each card. The positive response was an active response which involved the actuation of a device. The negative response was passive requiring no action. Two responses were thus required for each stimuli presentation. The two stimulus cards had geometric figures that varied according to two perceptual dimensions (colour and shape). Each card had a time window within which the participant must provide a response. The time window was marked by a thick black frame which outlined the left card first, and then proceeded to outline the right card. Participants had to provide a positive response when the frame outlined the correct card and a negative response when the frame outlined the incorrect card. Feedback was provided in the form of the sound of a bell tinkle when the virtual switch was actuated and the response was correct. When the response was incorrect, a boing sound was produced.

When responses were negative, the frame would proceed from the left card to the right card and then to the next presentation. Success terminated the trials determined by two sequences (initial learning, transfer of learning) of 10 consecutive correct responses. Otherwise, the trials terminated after a block of 48 trials. Participants were familiarized with a few runs of the test to ensure they understood the test and were able to engage the devices. During the tests observations were made by the researcher regarding test response behaviour and notes compiled after testing. A maximum of four blocks of trials would be carried out in a session with each block dedicated to a specific sequence. Devices were run in order of the best physical device from a previous evaluation first. In general, this would involve running tests with the mouse/switch first, followed by the other physical device chosen. Signing was used to verify the results of the physical device when the child did not have adequate representation with the physical device.

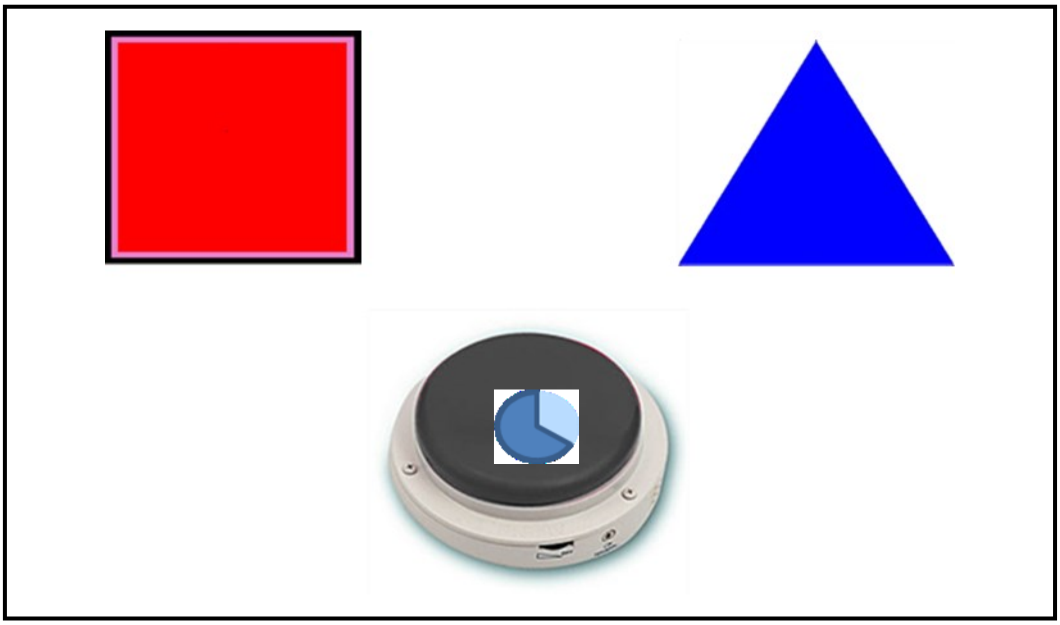


Figure 43 - Stimuli display for ECDT trial using a multi-modal point and select device. A red and blue card is displayed. The red card is highlighted by a black frame to indicate that a response for that card is required. The virtual switch substitutes for a physical switch when a non-hand held is used. The blue pie appears when the switch is pointed at and indicates the strength of the selection. The light blue region increases clockwise and at full strength, the dark blue region disappear leaving a twelve o'clock line (dark blue border). The strength of the selection is dependent upon the bite duration as well as clarity of the signal. Participants were told that they will be shown two pictures, one of which is right and the other, wrong. They have to work out which is the right picture. When they see the right picture in a black box, they are to "press" the switch. Participants also undertake a familiarisation test run which consists of a simple exercise to get them used to the environment.

When an eye-tracker was used, participants have to move a mouse cursor with their eyes to an image of a switch and “dwell” the cursor over the virtual switch for a period identified as the “dwell-time”. When a head-tracker was used, participants have to move a mouse cursor using movement of their head to the virtual switch for the dwell-time. When Signing was used, the participant was asked “Is this the correct answer?” If they were PNI children, they would use gestures that were specific to them to respond with YES or NO to a carer who will then interpret their responses. For non-PNI children, they used any form of direct communication they liked (e.g. touching the card they think is correct) and if it was not clear, they would follow the same procedure as for the PNI child.

The experiment is a maximum of 2 x 2 x 48 within-subjects factorial design for the maximum number of trials (Figure 44). For each participant the following are components of the test:

- Device {Pair combinations of Signing, mouse/switch, eye-track, head-track}
- Sequence {1 to 2}
- Block {1 to a range between 10 and 48}
- Test {ECDT}

Participants use a maximum of 2 devices. Using a device, participants have to complete 2 sequences of up to 48 trials. The second sequence cannot commence unless the participant succeeds in 10 consecutive successful responses. There were 18 participants giving a maximum total number of 18 x (2 x 2 x 48) trials.

9.3.3. Design

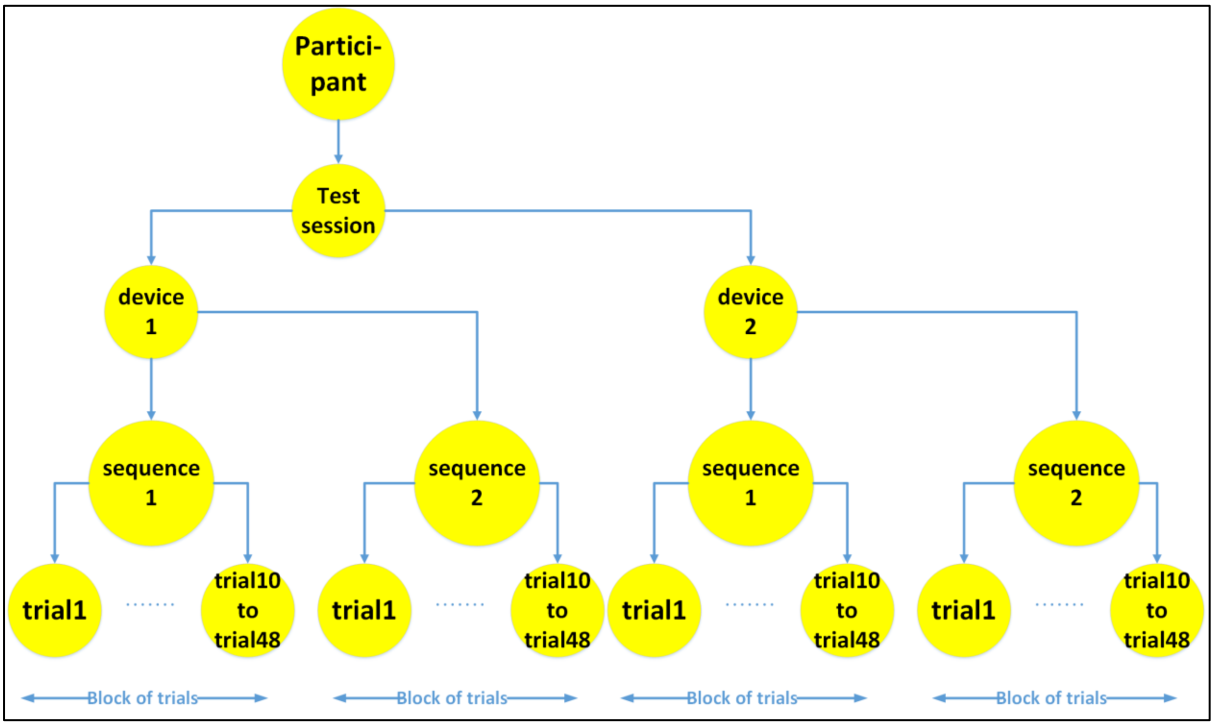


Figure 44 - Test design. A maximum of 2 devices were used for each participant.

9.3.4. Data capture

ECDT results were represented as a maximum of two 48-bit fields to represent 2 sequences; each bit represented an OK/NOK (not OK) outcome for a particular trial. The field was represented by a list of success and failure streaks. Failure streaks were suffixed with x. For example, for two sequences which terminated after 16 trials of 3 successes followed by 3 failures and 10 successes, were represented as {3, 3x, 10}, {3, 3x 10}. The consecutive successes and failure were referred to as success and failure streaks respectively. The entire list which captured an entire block of trials was referred to as an outcome event sequence. The maximum number of consecutive successes in the example list is 10. Multiple maxima may exist in general.

9.4.Results

This section describes the results obtained for an assessment of cognitive competency.

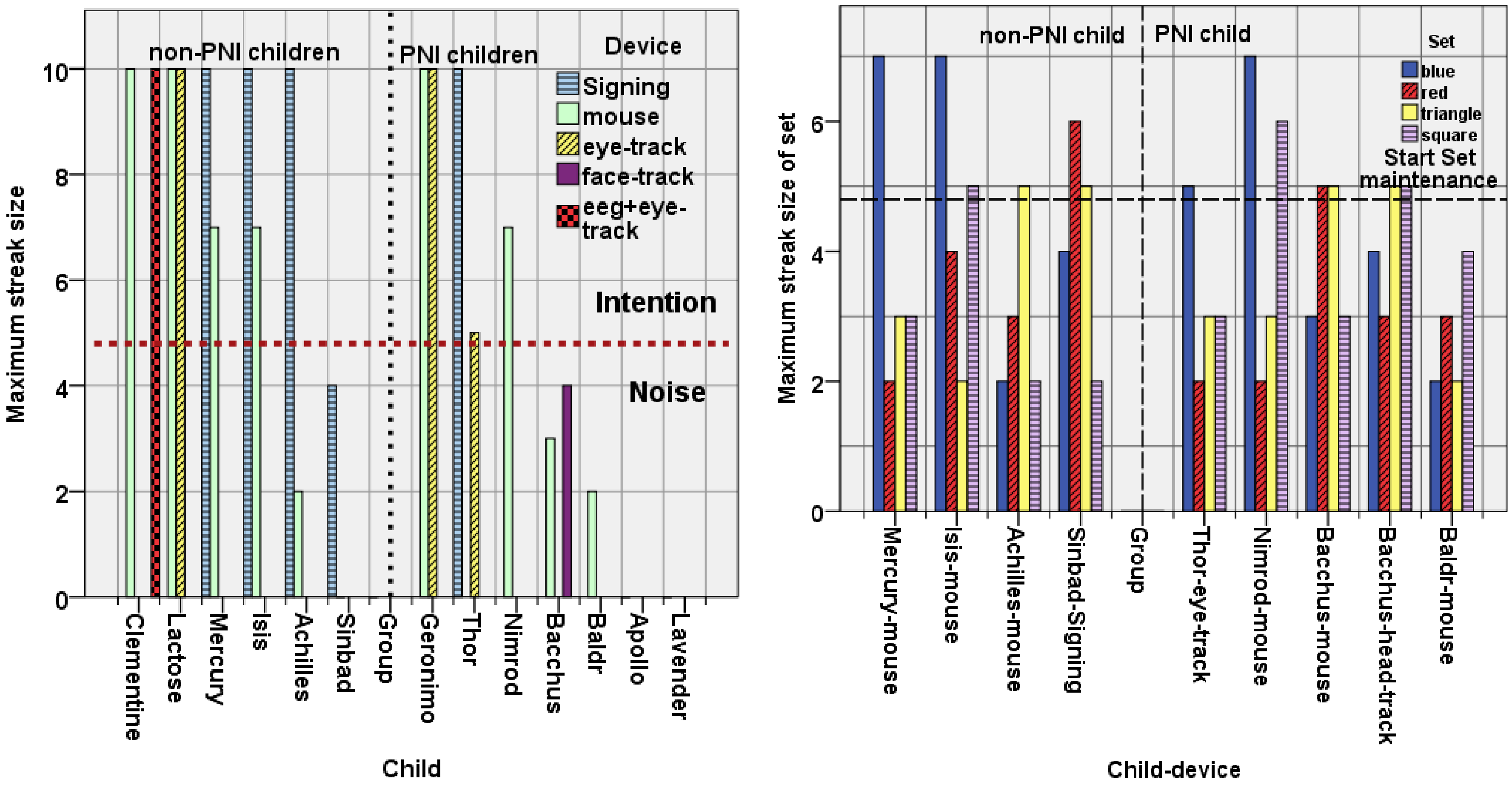


Figure 45 - (Left) Test (ECDT) outcomes of PNI and non-PNI children (separated by dotted vertical line) using a maximum of two devices. Children who had a streak size of 10 passed the test. Tests using Signing were performed after the physical device established poor performance. (Right) Fixations of children who failed ECDT are indicated by the maximum streak size obtained for responding to an incorrect (not blue) dimension. The level used for failure to start-maintain-set (shown by the dotted horizontal line) was used as a benchmark for maintaining the wrong set. However, this level is possibly more stringent than required. A noise-intention analysis of the dimensional results indicate that a 95% confidence for the intention zone for incorrect dimensions rests at a maximum streak size above 3.

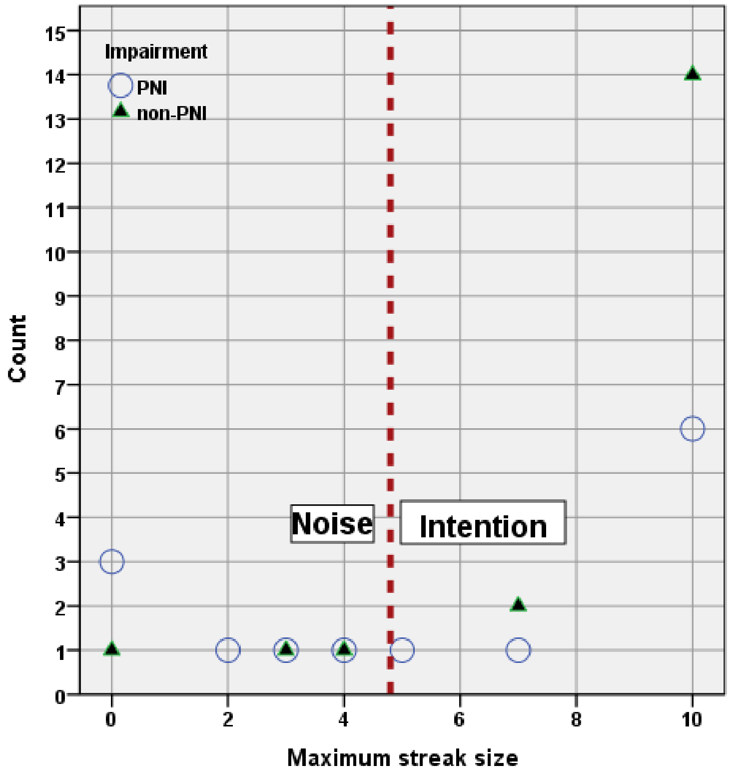


Figure 46 - Maximum streak distribution of PNI and non-PNI children taking test (ECDT). Results for PNI children are shown with circles and non-PNI children with triangles. Each count is a test sequence within a test run with a specific device. The dotted vertical line represents a threshold that separates intention outcomes from possibly random outcomes.

9.4.1. Results from the experiment with PNI Children

Figure 46 shows the number of sequences (count) that attained a specific maximum streak size. The counts for PNI children are shown by the circles. The figure shows three areas of interest; non-engagement with the test where the size is 0, engagement but no success where the size is between 2 to 7, and success where the size is 10. Non-engagement refers to instances where the test was actually not undertaken because the participants failed to understand the test or the test was aborted at the request of the participant. The PNI participants were deemed to have not understood the test when the carers indicated that they were not successful in explaining the test due to the limited symbols used for communication or there were problems with attention. The non-PNI child aborted the test as he got bored with not achieving the success streak size. Success sees an early termination of the test when the success criterion was achieved. The region of the graph that is of most interest is where the size is between 2 to 7 because it means that the participant has failed to meet the objectives of the test but that does not necessarily mean that the participant cannot meet the objectives under a different set of circumstances. In order to investigate further, we need to separate the outcomes which we can draw no conclusion from (the noise) and the outcomes that appear to be intentional.

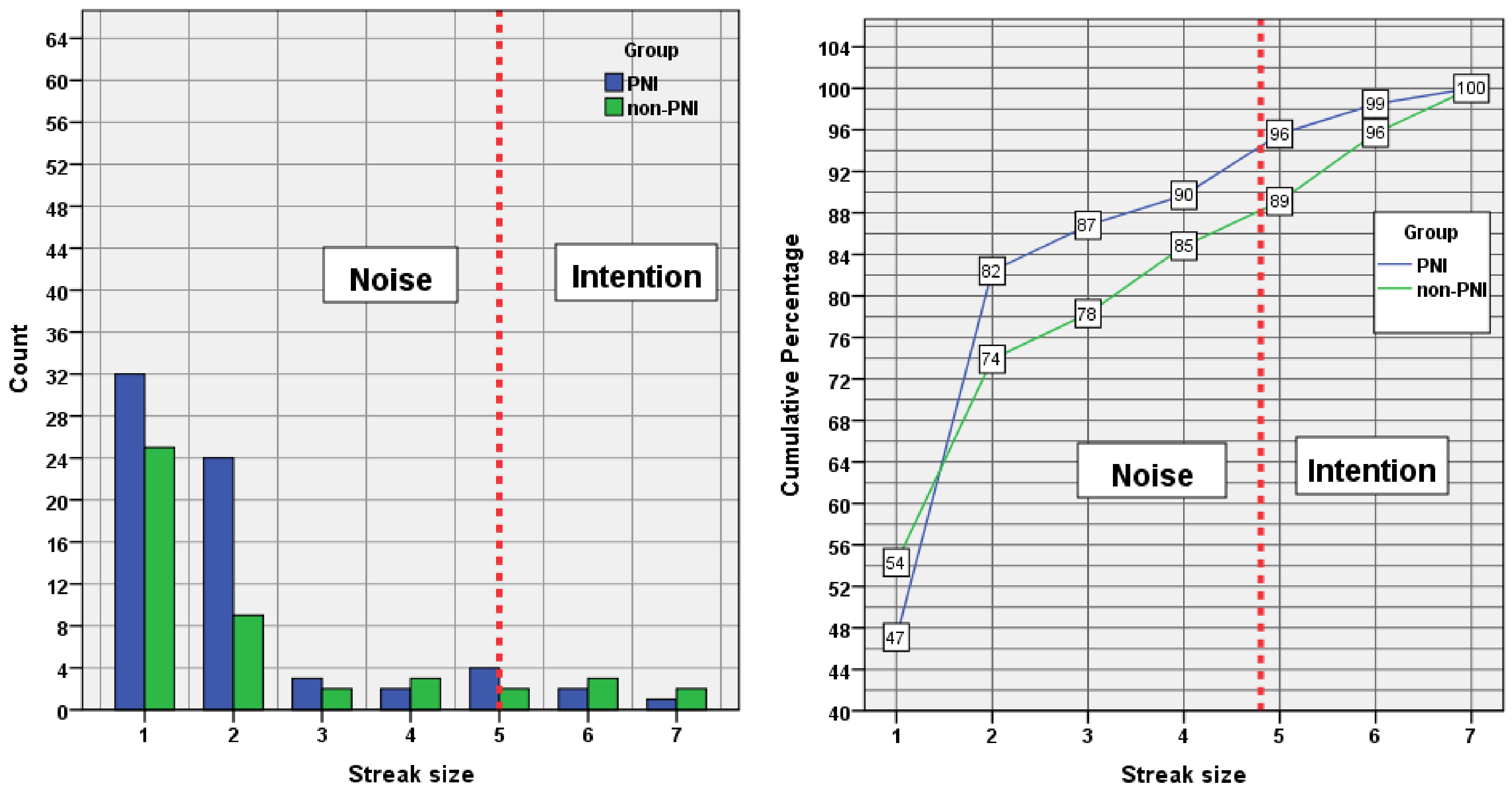


Figure 47 - Streaks size distribution of test runs for maximum-streak outcomes of between 2 to 7. PNI children are shown in blue and non-PNI children in green. The cumulative distribution (left) shown as a cluster plot and the frequency distribution (right) of the number of streaks of depicted size is shown. The plots indicate that the probability of responses decreases rapidly to small values with increasing streak sizes. The high initial values correspond to errors caused by involuntary motor actions, lack of attention or cognitive inability. A threshold can be established which distinguishes streak sizes that represent intention from the noise of involuntary motor actions and other poor results. Choosing a streak size threshold of ≥5 provides a confidence level of between 85% to 90% for an intentional response as indicated by the dotted vertical line.

Figure 47 (left) shows the distribution of streaks for the cases where the test runs have resulted in outcomes that are not at the extremes of non-engagement of the test or reaching the target criterion. These form cases where the number of trials (48) for each test run has not been shortened by achieving the target criterion. Each test run would have resulted in a number of streaks of varying sizes between 1 to 7 and having a maximum of between 2 to 7. The distributions for the PNI children are shown in blue. The figure shows most of the streaks have sizes of between 1 and 2 and a dramatic drop in numbers occur from 3 onwards.

Figure 47 (right) shows the corresponding cumulative percentages of the counts of various streak sizes. The significance of choosing the noise threshold of a streak value of 5 and above to represent intention and 4 and below to represent noise can be seen from the figure. The confidence levels range from between 90 to 96% for PNI children which are reasonable figures for acceptance.

Figure 45 (left) shows the maximum-streak outcome of the test (ECDT) for all the children. Each child would attempt the test with at most 2 devices. PNI children are shown grouped on the right hand side. The children are grouped in order of results with the best starting on the left. ECDT has two sequences (initial learning and transfer of learning) where target streak sizes of 10 have to be achieved for each sequence to pass the test. In all cases where the maximum-streak size was 10, the children have succeeded in both sequences and passed the test (there are two maximum-streaks of 10). Where the children have not managed to achieve the maximum-streak size, all of them had been stuck in the first sequence (initial learning). Two children passed the test (Geronimo, Thor). One child (Thor) had not passed the test using a physical device and was evaluated again using direct communication (Signing) and succeeded. The others failed the test. When we compare the performances of pairs of devices, we find that there was a discrepancy only in the case of Thor. For the other children, the results for both devices were comparable or only a single device was available.

Figure 45 (right) uses all the outcomes that have failed ECDT to compare the relative responses for the dimensions (of which the sets belong to) presented in the stimuli. The correct response is the colour blue. The plot shows the children's (and device they were using) greatest response (maximum streak) to the four main sets that each card image presented as stimuli may possess for each test run. The maximum-streak measure can be regarded as a measure of greatest perseveration for the different sets of dimensional responses. PNI children are shown grouped to the right. A horizontal dotted line separates the level at which set maintenance starts to take place. WCST uses this level to produce metrics for failure to maintain set where consecutive successes of 5 or more but less than the target 10 are counted. In this case, it is used to give a hint as to which set the child thinks is the solution to the problem. The results for each child-device pairing is clustered so that the correct response (blue) is shown first followed by its complement (red) followed by the sets of the incorrect dimension (of shape) which are a triangle and a square. The results are ordered according to the order in the left plot. 2 children (Thor using eye-tracker, Nimrod using the mouse) had significant fixation on the solution (blue). One child (Bacchus using the mouse) had significant fixation on the wrong colour (red). 2 children (Nimrod using the mouse, Bacchus using the mouse and head-tracker) had significant fixation on the wrong dimension (shape).

Table 16 - CATTEST and ECDT results for PNI children using their best devices. A Pass is given to those that achieved the target criterion for the test and a Fail for those who did not. Children were only re-tested on Signing if they had poor results. Failures in CATTEST were compared with ECDT. CATTEST is cognitively less complex than ECDT.

Child	Device	CATTEST	ECDT
		Pass/Fail	Pass/Fail
PNI			
Apollo	Signing	Fail	Fail
Bacchus	Mouse	Pass	Fail
Baldr	Mouse	Pass	Fail
Geronimo	Eye-track	Pass	Pass
Lavender	Signing	Fail	Fail
Nimrod	Mouse	Fail	Fail
Thor	Eye-track	Fail	Fail
	Signing	Untried	Pass

Table 16 indicates that PNI children who failed in the light cognitive test (CATTEST) were also unable to pass the more complex ECDT. The results were compared on a device basis. Only Thor did not have a CATTEST comparison for Signing. Children who managed to pass CATTEST can either succeed or fail in ECDT. The results are discussed in section 9.5.5.

Apollo

Apollo could not cognitively do the test. Both the researcher and his carer tried to explain the test but the vocabulary used in his Signing repertoire was too small and it was difficult to convey what had to be done for the test. In a case like this, sometimes the participants would intuitively understand what had to be done after a few tries but this did not happen with Apollo. In addition, Apollo was not able to do the previous simpler categorization test (CATTEST) and the result was consistent.

Bacchus

Bacchus was not successful in this test. However, during the test he verbalized his thoughts on the solution. It would appear that he was fixated with the wrong dimension and did not explore other dimensions enough. This suggests that if the test had the other dimension as a solution, he might have been successful. The learning process that was the aim of the test was demonstrated but the specific requirement was not. As a side-effect, the test demanded some lateral creativity.

Baldr

Baldr was unable to do the test. Baldr was more interested in the negative feedback which was a boing sound. However, his responses were mostly to hit the switch for every stimulus presented.

Geronimo

Geronimo had no problems with the test. Geronimo succeeded with both a mouse and an eye-tracker.

Lavender

Lavender could not do the test. Her short attention span prevented her from engaging with the test.

Nimrod

Nimrod could not satisfy the requirements of the test but progressed as far (streak size of 7) as all previous tests. This suggests the possibility that he got confused after a short time.

Thor

Thor could do the test but not with a physical device. Thor showed that he had no problems with the cognitive element of the test shen Signing was used. However, with the use of the only physical devices he was able to use (eye-tracker) the results were poor (streak size of 5).

9.4.2. Results from the experiment with non-PNI Children

Figure 46 shows the maximum streak size distribution for non-PNI children as triangles. The graph shows values that are low except when the maximum streak size is 10. Streak sizes of 0 mean the participant is disengaged from the test and 10 mean that the test target was reached. Values that are not 0 or 10 represent values that are not subjected to early termination.

Figure 47 (left) shows the distribution of streaks for non-PNI children in green, omitting streaks that are 0 or 10. The figure shows most of the streaks have sizes of between 1 and 2 and a dramatic drop in numbers occur from 3 onwards. The “knee” of the data starts from 3.

Figure 47 (right) shows the cumulative percentages for non-PNI children in a green line. The significance of choosing the noise threshold of a streak value of 5 and above to represent intention and 4 and below to represent noise can be seen from the figure. The confidence levels range from between range from between 85 to 89% for non-PNI children which are reasonable values for acceptance.

Figure 45 (left) shows the maximum-streak outcome of the test (ECDT) for all the non-PNI children on the left hand side. 2 children passed the test with their first device (Lactose, Clementine). The rest of the children except one (Sinbad) passed the test the second time round using Signing. When we compare pairs of devices used, we find that the children who passed the test using Signing had huge discrepancies but the others had comparable results with both devices or the results could only be collected for a single device (for the case of Sinbad).

Figure 45 (right) shows the children's attentional fixation on each dimensional set of the stimuli presented. Only children who failed the test are presented in order to observe their item of fixation. The non-PNI children are shown grouped to the left. Two of the children (Mercury, Isis, both using the mouse) had significant fixation on the solution (blue). One child (Sinbad using Signing) had significant fixation on the wrong colour (red). Two children (Achilles using the mouse, Sinbad using Signing) had significant fixation on the wrong dimension (shape).

Table 17 – CATTEST and ECDT results for non-PNI children using their best devices. A Pass is given to those that achieved the target criterion for the test and a Fail for those who did not. Children were only re-tested on Signing if they had poor results. Failures in CATTEST were compared with ECDT. CATTEST is cognitively less complex than ECDT.

Child	Device	CATTEST	ECDT
		Pass/Fail	Pass/Fail
Non-PNI			
Achilles	Mouse	Pass	Fail
	Signing	Untried	Pass
Clementine	Mouse	Pass	Pass
Isis	Mouse	Pass	Fail
	Signing	Untried	Pass
Lactose	Mouse	Pass	Pass
Mercury	Mouse	Fail	Fail
	Signing	Untried	Pass
Sinbad	Signing	Pass	Fail

Table 17 was only able to identify one child (Mercury) non-PNI who failed CATTEST. He was not evaluated using Signing for CATTEST. He failed ECDT with a mouse as well but passed using Signing. The number is small and this result is discussed in section 9.5.5.

Achilles

Achilles can do the test but on this occasion not with a mouse. Achilles had no problems with the test when Signing was used subsequent to having done the test using a mouse indicating that Achilles had no problems with the cognitive requirements of the test. Achilles appeared to concentrate mostly on a set (triangle) in the incorrect dimension (shape) so the possibility was that Achilles decided to use an alternative strategy when Signing was used.

Clementine

Clementine had no problems with the test using both a mouse and an EEG and eye-tracker hybrid. She represents the higher performing sample of population with both good motor and cognitive skills.

Isis

Isis can do the test when Signing was used subsequent to her doing the test using a mouse. She had focused on the right solution (the colour blue) with a mouse but did not maintain the set long enough to satisfy the test requirements. She had a couple of errors and was maintaining set before the test terminated. The possibility was that she had the solution but did not have the time to express it with a mouse.

Lactose

Lactose had no problems with the test using both a mouse and eye-tracker. She represents the higher performing sample of population with both good motor and cognitive skills.

Mercury

Mercury was another child who was successful in the test using Signing but not using a mouse. He had focused on the right solution (the colour blue) with a mouse but could not maintain the set subsequently. This suggests the possibility that the effort of both the use of a mouse and the cognitive requirements of the test was high enough to prevent Mercury from being successful.

Sinbad

Sinbad gave up on the test using a mouse. When Signing was used, the results were also negative. Sinbad fixated on the wrong set (the colour red) and gave up on the test.

9.4.3. Results from the experiment with adults

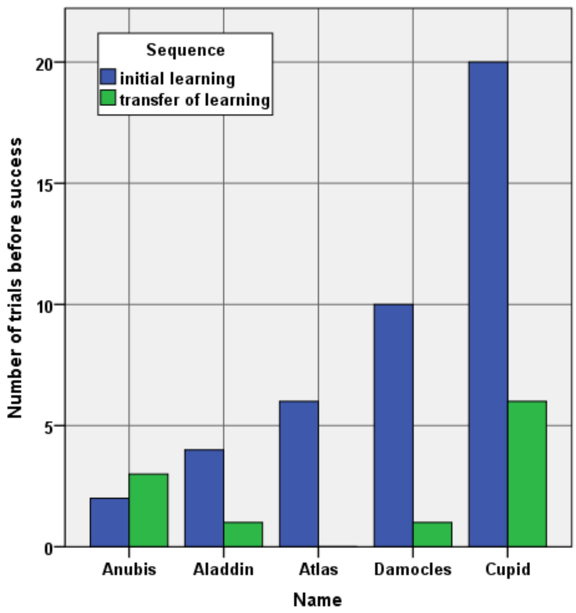


Figure 48 - Number of trials before learning for adult sample.

Figure 48 shows the number of trials that an adult will take before the target number of correct responses is reached during the two sequences (initial learning and transfer of learning) in ECDT. The graph is ordered in ascending order according to the number of trials before success in the first sequence of the test (blue). This is compared with the number of trials before success in the second sequence of the test (green). We see that the number of trials taken before success is always less in the second sequence than the first sequence except for Anubis where the difference is marginal (2 trials for initial learning and 3 trials for transfer of learning).

The mean number of trials before success for initial learning (M=8.40, SD=7.127) was higher than the mean number of trials before success for transfer of learning (M=2.20, SD=2.387) resulting in a mean decrease (M=6.200, SD=5.718) in the learning time (number of trials before success). This decrease was not statistically significant, t (4) =2.424, p<0.072, two-tailed. However, the sample size is small. The t-test results support the transfer of learning in the adult sample during ECDT.

9.5.Discussion

9.5.1. Variation of maximum streak distributions between PNI children and non-PNI children

The maximum streak distribution show higher values for the extreme streak values (streak values of 0 and 10) (in Figure 46) because of the test design. Apart from the extreme streak values, the distribution is fairly uniform. A streak value of 0 means that a child is not engaged in the test. The disengagement is a group count of children who had problems with the devices and understanding the test. The PNI participants were deemed to have not understood the test when the carers indicated that they were not successful in explaining the test due to the limited symbols used by their devices for communication or there were problems with attention. The non-PNI child aborted the test as he got bored with not achieving the success streak size using a mouse. Similarly a streak value of 10 will terminate a sequence in the test and begin another sequence. If the sequence were not terminated at 10, a spread of values of above 10 would be obtained. From that point of view, the count of children having a streak score of 10 is a group count of children with a 10 plus maximum streak. The higher ability of the non-PNI children hence results in a higher number who achieve the target of 10 than the PNI children.

9.5.2. Variation of streak size distributions between PNI children and non-PNI children

The trend for the streaks distribution is for a much higher count of streaks for low streak values (from 1 to 2) and thereafter (3 and above), the count drops to a low almost constant value (Figure 47 (left)). Although the streak distribution should be strongly driven by cognitive responses as the best devices are used, one child (Thor) had to use the only physical device that can afford representation which presents a limit to his responses. The combined effect of user, devices and test appear to stabilise the streak value counts to a low value when the streak value (3) is small. The noise threshold placed at a streak value of 5 is therefore quite adequate.

9.5.3. Effects of test and device between PNI children, non-PNI children and adults

The results of ECDT for children contain examples where some children would fail the test using their best device (usually the mouse) and pass the test using Signing (Figure 45). Others would have comparable results with both devices. The results of the adults do not carry any such discrepancies. In this section, we discuss whether a device, especially a proven device can have an impact on cognitive test results.

The demands of both device and test impose a cognitive load on the children. A device that a child has developed a good skill for (such as their best device) would mean that the cognitive effort used to work the device is minimal (Ericsson et al., 1993; Shadmehr & Holcomb, 1997). It is therefore an expectation that devices that have been found to fulfil the performance requirements set by previous motor skill and cognitive tests (COMPTEST, CATTEST) achieve the same result. All the children except one (Thor) used devices for ECDT which COMPTEST and CATTEST indicated would represent the child. The results from either COMPTEST and CATTEST would show that either the child had reached the target criterion for the test (streak size of 20) or the results for the pair of devices, one of which was Signing, was comparable. The discrepancy obtained between Signing and the best device suggests that the effort demanded by Signing has enabled a good cognitive response, outperforming the best device. Conditions existed for these children where the best device was not appropriate but Signing was. It remains for further research to be done to understand those conditions (Xie & Salvendy, 2000). A simple view would be that the effort required by Signing is much lower than that required by any existing physical device.

In the case of Thor, the device was the only one he could use and provided fair results for previous tests because of his impairments, so it was not a device that could give the required results in the first place. Signing confirmed that this was the case.

The use of Signing suggests the possibility that certain combinations of user and device affect the results. In order to investigate user, device and test interactions, this study has to be extended. More participants are required and the test targets need to be evaluated for extension.

9.5.4. Variation of cognitive ability between PNI children, non-PNI children and adults

The results show that PNI children performing more poorly than non-PNI children. The adults have no problems with the test (Figure 45, left, Figure 48).

Of the PNI children, 4 failed due to true cognitive deficit (Nimrod, Baldr, Apollo and Lavender), 1 failed possibly due to test design (Bacchus) (explained shortly) and 2 passed. Of those that failed due to cognitive deficit, 2 failed to engage with the test (Apollo, Lavender) and the other 2 who engaged (Baldr, Nimrod) provided data that showed their attentional focus (Figure 45, right). One (Baldr) had equal attentional focus on the test stimuli and possibly did not know how to proceed with the test. Nimrod's results were interesting as he showed a focus on the correct solution but was not able to maintain the focus. This is a metric used in a similar cognitive test (Nyhus & Barceló, 2009) to detect perseveration failure with the solution referred to as failure to maintain set. Huizinga (Huizinga & van der Molen, 2007)reported studies on different groups to establish different levels of developmental maturity for set maintenance. The results exhibited by Nimrod suggest a possibility that his type of failure is a developmental condition. The one who possibly failed due to test design was regarded so as he verbalised the expectation for an incorrect dimension. The test design could be changed to allow for options where different dimensions can be used for the initial learning. Since the objective of ECDT was to establish initial learning and show transfer of learning, Bacchus did not get a chance to demonstrate initial learning. 1 of the children (Thor) who passed could only do so with Signing since the device he was using imposed limits due to his impairments.

All the non-PNI children except one passed the test. 3 out of 6 children only passed the test using Signing after failing with a physical device. The results suggest that the cognitive load imposed by both device and test was sufficient to cause failure and the substitution of a device with lower cognitive load would enable success. The child who failed the test used Signing as he initially used a mouse but gave up mid-test as he could not arrive at the solution. His attentional focus was shown (Figure 45, right) to be high on the wrong colour (red). Of interest is attentional focus of the 3 children who passed the test using Signing but failed using a physical device (Mercury, Isis, Achilles). Two (Mercury, Isis) had a strong focus on the solution (blue) but failed to maintain set to pass the test using the device. This supports their passing the test using Signing as they appeared to be on the right track already. The other was focussed on the wrong dimension but perhaps could not provide the effort to provide the correct solution using the mouse.

9.5.5. Pre-test predictors of ECDT

The results show that children who fail the light cognitive test (CATTEST) also fail in ECDT (Table 16, Table 17). The results are taken on a device basis as the failure may be due to the combined load of both device and test.

This suggests that CATTEST is suitable as a pre-test filter of PNI children for ECDT. CATTEST does not predict all failures of ECDT. Those who pass CATTEST is not assured to pass ECDT. Since ECDT is intended to be only run once, CATTEST is invaluable that either initial problems can be detected before-hand and more detailed investigations carried out or if developmental progress is expected, appropriate re-tests can be carried out before ECDT is finally run.

9.6. Conclusion

Results show that devices do have a range of impacts on the test (ECDT). There were two groups of results; one where the device was adequate and the child was capable and the child was successful regardless of the device the child used. The second group of results would show the child having vastly better results in the test using a procedure (Signing) but fail in the test using a physical device. The primary differentiation that was made in most cases was between the mouse and Signing. The particular form of Signing would vary between participants depending on impairments and preferences. For non-PNI children, the cause for the difference between Signing and a physical device was possibly the effort required between the two with Signing being an agreeable choice. With the PNI child, the only device (eye-tracker) that he (Thor) could use was not able to afford good control. So for the non-PNI children, we could in general attribute the cause to fatigue as a speculation and for the PNI child it was due to both fatigue and impairment. The success of Signing implies that the potential exists for a range of devices could be designed that matches the preferences and abilities (or impairments) of individuals.

From the small sample, there was poorer performance from the PNI children. When compared with the non-PNI children, none of the PNI children except the control (Geronimo) and one other (Thor) out of the 7 could do the test (1 out of 6). In the case of non-PNI children, all except one were able to do the test (5 out of 6). Another major characteristic for PNI children that is different from non-PNI children is failure due to non-engagement in the test as opposed to failure when engaged in the test. Due to the limited communication, it was difficult to work out how much of the test was understood by the PNI children as a cause for not being able to engage. Out of the 5 PNI children that failed the test, 3 could not engage but all of the non-PNI children appeared to engage in the test. One PNI child exhibited an interesting characteristic of not being able to maintain set for all the 3 tests (COMPTEST, CATTEST, ECDT) but could engage in the tests. It would thus appear that Nimrod has the potential to pass all the tests if he could attain a higher level of performance.

The implications for the results obtained in this study suggest necessary refinements for ECDT. As a test that is going to determine the future of a PNI child (stakeholder test), it is insufficient because the PNI child can fail the test not because of an inability to learn. PNI children can fail because of a variety of other reasons that do not befall that of a non-PNI child. A series of tests of increasing difficulty is suggested similar to COMPTEST and CATTEST. If the child fails the test, it is suggested that a re-test be done using Signing. The final suggestion for refinement concerns an observation made regarding children that start the test by fixating on the wrong dimension. The child fails because one dimension seems to be the obvious answer. As test to detect learning, this prevents the child from getting past the initial learning sequence even though the child can learn a property of the stimuli. As a lower rank of ability, it is possible to allow the child to use the wrong dimension in a re-test to allow demonstration of transfer of learning.

10. MAIN PILOT STUDY: Integration and evaluation of results

10.1. Introduction

This chapter provides an evaluation of results obtain in the three tests (COMPTEST, CATTEST, ECDT) of the main study test-suite (section 6.1.4) and provides an integration of the results to the initial theoretical model discussed in the literature review (section 2.3.6).

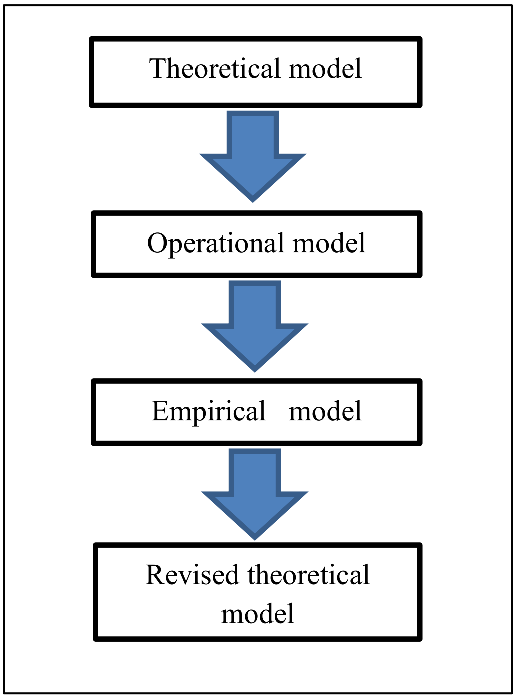


Figure 49- Development of theoretical and operational model for device ability

Figure 49 shows the theoretical model as the starting point of a development process. An operational model is developed to support the theoretical model. The operational model consists of the implemented environment to support the study using the theoretical model as a hypothesis. The results and data gathered from the running the operational model provides the empirical model. This is a model where the actual observations either support, rejects or modifies the theoretical model of understanding. The observations provided by the empirical model may suggest new factors to consider resulting in a revised operational and theoretical model for future work.

Previous discussions of the study have been limited to specific tests although there have been a small amount of association from one test to another. The tests have in a sense been kept constant and the user ability and device variation have been discussed. In this chapter we will keep the user constant and have a look at variation across devices and tests.

10.2. Research questions

The research question that this chapter answers is the high level question concerning the choice of the device:

Research question
RQ1. What device helps a PNI child successfully attempt a cognitive test?

10.3. Development of a theoretical model

The development of the theoretical mode was discussed in section 2.3.6. the result of which was a ranked table of devices as shown in Table 18.

Table 18- Devices ranked in order of complexity of use starting with the least complex devices at the top of the table and the more more complex at the bottom. Single-mode devices are seen to be less complex than multi-modal devices. In addition a device that requires a selection operation is less complex than a device with both pointing and selection.

Device	Mode	Operations
Signing	Single-mode	Selection
Mouse[4] or switch		
EEG headset		
Head-tracker		Pointing and selection
Eye-tracker		
EEG headset + head-tracker	Multi-mode	
EEG headset + eye-tracker		

10.4. Development of an operational model

In developing the operational model to verify our theoretical model, we identify three basic components of the system used in the study that can influence the results as being the user, the

[4] Although the mouse can be used for pointing, this is not used in the study.

device and the test. Having considered complexity as a major factor in the theoretical model, we choose device complexity and test complexity to be two independent variables for the empirical model. Since latent ability was also identified in the theoretical model, an attribute associated to latent ability was used as the third independent variable. We chose mental age as the attribute since we undergo a developmental cycle that reaches its peak in adulthood. This follows a body of work as charted by Piaget (Piaget, 1970; Sroufe et al., 1992). The expectation is that latent ability for device increases with use and matures in adulthood. As shown in Figure 50, user variation was set up by having groups of PNI children, non-PNI children and adults, device variation was provided by the assumed device complexity discussed in the theoretical model set up and test variation was set up by having three tests of increasing cognitive complexity. Although it is also possible to split the user ability into sub-attributes of impairment (with the PNI children) and non-impairment (with non-PNI children), a simplification has been made in considering impairment as a degeneration of general ability.

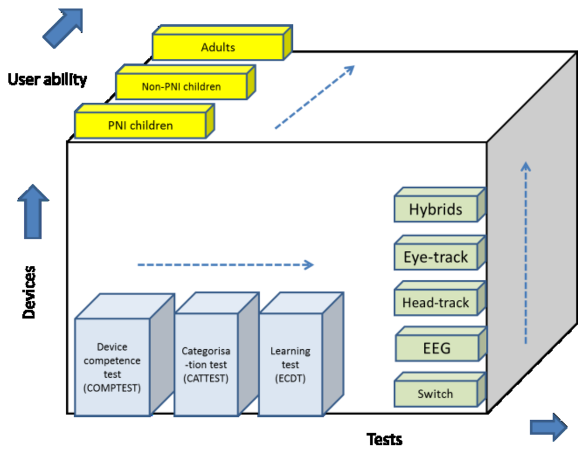


Figure 50- Framework for operational model consisting of the user- device-test triad as independent variables

In order to implement the empirical model, we need to have a system of devices, tests and measures. We find that in general, devices, tests and measures are not designed for PNI children. We therefore have to adapt the devices, simplify the tests and develop new measures. The devices were adapted so that the PNI children can understand how to use them (section 6.1.3). The devices were operated using functions that the children already understand, in particular for the EEG headset, mind control was not used but they were told to use bites instead. The tests were simplified to have the minimum number of simultaneous variation of attention points (section 2.4). The three tests in the test suite provide the participant with only one stimulus for COMPTEST (either cat or dog) and two stimuli for CATTEST (fruits and birds) and ECDT (shape and colour). Procedures were introduced in the tests to isolate the effect of devices from the cognitive targets of the tests (section 6.3.3). Measures using streaks as a basis (section 0) were developed to ignore errors that are randomly generated due to impairments and focus on intentional responses instead.

Using the implementation of the empirical model we establish a set of results that enable us to determine device fits for PNI children and when the best fit has been made.

10.5. The empirical model

In order to put the problems faced by the PNI children into perspective, Table 19 gives a summary of the a child's impairment and the effect on device handling. Although a child may have severe impairments, some are still able to handle a device competently. Table 21 provide the results for the empirical model. The central issue observed was that PNI children with good cognitive ability are able to overcome a lot of other problems. At this stage we see that the well-practised mouse is dominant. The device is commonly used to interface to the computer and children would have had long periods of exposure using the device. The device has also been in popular use for a long time and been highly optimised in design as a result of feedback from many users. Any second best device is generally managed by children who have well developed all round skills and they manage to have good results for any device. We see this from Geronimo, Lactose and Clementine in their COMPTEST results (Figure 27) and their results in the table. The implication is that a required modification of the initial theoretical model based on latent ability will be discussed further in the next section.

Table 19 - PNI child's impairment and extent of impact on testing. Although the impairment is severe, the severity may not be high enough to affect the child's performance with a device.

Child	Impairment or problem	Impact device handling?	Other issues
		PNI	
Geronimo	Quadriplegic	No	
Bacchus	Hemiplegic	No	
Thor	Quadriplegia. Involuntary torso movement	Yes	Limited communication
Baldr	Paraplegic + autism	No	Limited communication
Apollo	Quadriplegic	Yes	Limited communication
Nimrod	Clumsy with mouse	Yes	Communication is confused - inconsistent
Lavender	Quadriplegic	No	Attention difficulties stops testing

Table 20 - Legend for test grading to be applied

Legend	Streak value (COMP/CAT TEST)	Streak value (ECDT)	Grading
	20	10 + 10	Test pass
	20 > value >5		COMPTEST, CATTEST: Possible short-term deviation or ceiling reached
		10 > value > 5	ECDT: Failure to maintain set
	≤ 4	≤ 4	No information (noise)

Table 21 - Children's device capability. The results show the limits that the children reach with increasing test and device complexity. Some children marked "not tried"are untested due to availability in the simpler tests and achieve success in the more complex tests. These children were expected to also succeeded in the simpler tests. Children (Thor, Isis, Mercury, Achilles) were observed doing poorly in ECDT with a good device (e.g. mouse) but achieving good results with Signing. This suggest that cognitive test results can be distorted even with a good device when the cognitive effort required to multi-task between test and device becomes appreciable for the children.

Child (ranked)	Mental age	Device	COMP TEST	CAT TEST	ECD TEST
PNI					
Geronimo	11	Mouse	20	18	10 + 10
		Eye-tracker	20	20	10 + 10
Bacchus	4	Mouse	20	20	3
		Head-tracker	20	20	4
Thor	7	Eye-tracker	16	11	5
		Signing	Not tried	Not tried	10 + 10
Baldr	< 4	Mouse	20	2	2
		Head-tracker	20	2	Not tried
Apollo	4	Mouse	7	Not tried	Could not understand
		Signing	20	4	Could not understand
Nimrod	4	Mouse	5	6	7
		Signing	7	5	Not tried
Lavender	< 4	Switch	4	Could not understand	Could not understand
		Signing	6	Could not understand	Could not understand
Non-PNI					
Lactose	5	Mouse	20	20	10 + 10
		Eye-tracker	20	20	10 + 10
Clementine	5	Mouse	20	20	10 + 10
		EEG + ET	20	20	10 + 10
Isis	4	Mouse	20	20	7
		Signing	Not tried	Not tried	10 + 10
Mercury	4	Mouse	20	16	7
		Signing	Not tried	Not tried	10 + 10
Achilles	4	Mouse	8	20	2
		Signing	20	Not tried	10 + 10
Sinbad	4	Mouse	20	5	Aborted
		Signing	Not tried	20	4

10.6. Revision of the theoretical model from empirical model

The results from the empirical model suggest that the theoretical model is strongly affected by the development of latent ability as seen from previous section (10.5). Children in the mental age group of between 4 to 6 years old appear to be on the borderline of those who are still developing the ability relevant to handle the novel devices or have already developed those abilities. Those that have already developed the ability find few problems using any of the new devices. Those that are still developing those abilities find difficulty using the novel devices. The complexity that was envisaged in the theoretical model ceases to exist once the latent ability for handling the devices is developed.

The understanding for the results comes from a number of areas. The first is that the brain is organized into specialized functions that are generic for any skill. That is to say that although the brain has specialized areas, combinations of these areas tend to be required for different skills we employ rather than a specific area of the brain that is specialized for a specific skill. For example Broca's area is used when hierarchical abstractions are made of a group of functions or chunks (Clerget, Poncin, Fadiga, & Olivier, 2012; Fadiga, Craighero, & D'Ausilio, 2009). The organisation of a set of motor actions for tool use and the organisation of words into a grammar use the same area of the brain (Higuchi, Chaminade, Imamizu, & Kawato, 2009). Albouy (Albouy et al., 2008)demonstrates the involvement of the hippocampus for the consolidation of motor memory between practice sessions. However, the hippocampus is involved in rapid encoding of associations in general (Henke, 2010). Previous work has indicated that the hippocampus is involved in declarative memory which is used in explicit learning such as memorizing phone numbers (Squire, 1992; Tulving & Markowitsch, 1998). When the combination of areas required for handling the novel devices mature in development, proficiency in handling the devices become much better (Kuhtz-Buschbeck et al., 1999). The transference of skills by the brain is exemplified by the same areas of the brain that process music and language for instance (Fadiga et al., 2009; Patel, 2003). The general skill that is required in the case of device operation is one of sensorimotor integration (Ghahramani et al., 1995). For example the eye makes use of sensorimotor integration for visual motion processing often described as smooth pursuit eye movements (Lisberger, Morris, & Tychsen, 1987). It is a reasonable assumption that the skills acquired for smooth pursuit is transferable to the skill required for the operation of an eye-tracker because it is a similar skill. The eye must now move an object by moving the eyes instead of the eyes moving to accommodate a moving object.

The second is when a skill is being developed performance starts low and reaches a plateau after a period of practice. This skill can be compared to that which is required to ride a bicycle. Once mastered, it is never forgotten, so the plateau lasts forever. The plateau is marked by an automative stage of performance (as shown in Figure 51) where conscious thought regarding each set of movement is not required. This is the best stage to compare ability with different devices as a stable state is reached. In addition, in order to improve the skill to an expert stage, it is required to enter the associative cognitive state again to attain a higher level of automative state in performance. New tricks have to be learnt and old habits discarded (Ericsson, 2006). From a neuro-science perspective, procedural memory is established. A kind of hard-wiring of essential procedures required for the task takes place and the brain is able to engage the set of procedures more efficiently (Longstaff, 2005).

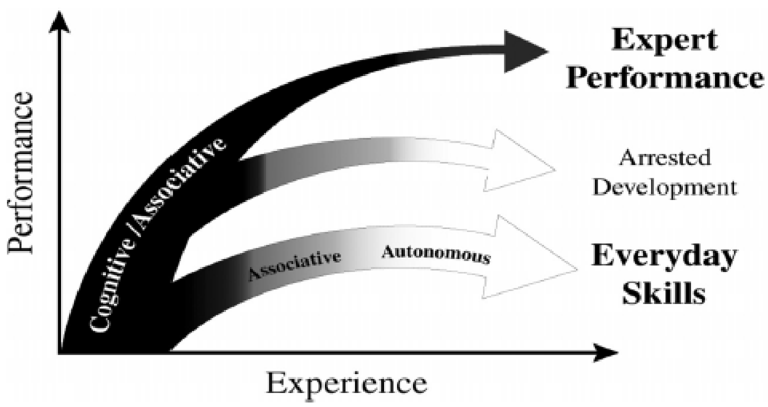


Figure 51 - Automative and associative motor control (Ericsson, 2006)

Results obtained from testing suggest that for some children, the theoretical model where hybrid mode devices are more difficult to manage is true but for other children who have more developed latent device skills, hybrids may achieve a better performance. The hybrids were designed to compensate for weakness present in the single mode devices. However, the difficulty is that the ability to handle both devices has to be present. In addition, the ability to coordinate the handling of both devices also has to be present. The case for latent ability overcoming the projected complexity of the theoretical model comes from two sources, Clementine and Geronimo (Figure 27, page 121) who performed better with a multi-modal device than the individual modal devices and adults who had a higher preference for some multi-modal devices instead of the single mode ones (e.g. Cupid in Figure 30, page 133) .

The results of the study suggest that the theoretical model needs to account for stages of development of children under test who appear to be undergoing a transitional phase. Results are more predictable once the children have gone past the transitional phase otherwise the results tend to be fairly negative and unstable and thus individual models are required. Studies indicate that motor skills of children undergo developmental stages (Kuhtz-Buschbeck et al., 1999) as also noted for their cognitive development (Piaget, 1970). Children who have gone past the transitional phase have developed more capabilities into the automative stage and will learn new skills faster than children still in the transitional phase. These results are consistent with studies made of the abilities of young children with a computer mouse (Donker & Reitsma, 2007a). Abilities with the mouse were partly attributed to fine motor skills acquired when they learn to write.

10.7. Re-revision of theoretical model from consideration of other factors

This section discusses the results that suggest that even after the revision of the theoretical model to one which is dependent upon the latent ability of the child, other factors that were beyond the scope of the study require consideration. These factors are ones which were better handled once the framework for the observation of the more basic factors were set up such as general ability and impairment. These more advanced factors include motivation, fatigue, stress (Xie & Salvendy, 2000), experience and memory consolidation. Motivation includes the desire for engagement into the activity and is held back by boredom. Fatigue refers to the lack of energy to provide the effort required for a particular task. Stress refers to some extenuating circumstance that the participant is experiencing and therefore distracting. Experience refers to the amount practice that the participant has already had with a device. Memory consolidation refers to the availability of an interval between tests with a device for the brain to adjust the information content to have more efficient access.

The study results show that the cognitive requirements of a specific device and test combination are sometimes too high for some children to handle which can be viewed from a relative perspective. The possibility exists that the requirements may be sufficient for the child to handle but not enough effort by the child has been put into handling the demands of both test and device. This is especially true in cases where non-PNI children have been provided with a mouse and put on a relatively simple cognitive test (CATTEST) in which they demonstrate that they are able to achieve using Signing instead. There was even a demonstration where a child (Achilles) fails in the simpler motor-skills test (COMPTEST) but succeed in the more difficult cognitive test (CATTEST). The possibility was that factors such as fatigue, motivation and stress have decreased the effort applied on the failed test. Human Computer Interface (testing) of pre-schoolers have often warned that children of that age group often have short attention span (Hanna et al., 1997). The children in the study are also exposed to a fair amount of activity throughout the day both in and out of the playground which contributes to fatigue. Motivation can either be positive, resulting in engagement and better results, or negative, sometimes resulting in boredom and worse results. Stress usually results in worse results but it may be possible that depending on the type of stress an opposite reaction occurs. The consideration of emotions opens up a large area of investigation because emotions have complex interactions and mixtures of emotions form new emotions (Ekman, 1992; Kort, Reilly, & Picard, 2001; Plutchik, 2001).

The previous paragraph briefly discusses emotional factors. Other factors that are not emotion based are experience and memory consolidation. Experience is the result of practise. In this study, the child has in general three attempts at using a couple of novel devices and one attempt for the rest of the novel devices. The previous section (10.5) indicates that with experience, performance should reach a particular ceiling which is automative. Although it is possible to take a test when experience with a device has not reached this stage, it would be best as it would mean that device effects are at a minimum. Memory consolidation suggests that the brain re-organizes the information learnt through a period of time after practice and efficiencies are gained through the consolidation resulting in better performance.

In order to perform these studies the same empirical model can be used except that the attribute for the user as an independent variable becomes one of these other factors discussed in this section as opposed to mental age that was used in the current study.

10.8. Case evaluations

10.8.1. PNI children

Apollo

The impairments of Apollo are severe. The closest fit of device was a finger-on-palm mouse. Other devices may possibly be used but requires customisation to take into account both muscular weakness, involuntary muscular actions and lack of muscular control. The success in use of Signing suggests that gesture recognition may be a promising pathway to a solution for a suitable device for Apollo. Apollo has problems understanding possibly due to the lack of a rich enough reliable communication channel. ECDT appear currently to be beyond the reach of Apollo. The categorization test is a possible starting point to understand Apollo's cognitive challenges.

Bacchus

Bacchus works well with a mouse. Some novel devices require coordination skills that Bacchus needs to develop. Results suggest that practice in more lateral thinking may help Bacchus be more successful in learning tests like ECDT.

Baldr

Baldr works well with a mouse. Baldr does not often communicate. Baldr seem to suddenly get the idea but requires the opportunity to demonstrate it. Some novel devices require coordination skills that Baldr needs to develop. ECDT appear currently to be beyond the reach of Baldr. The categorization test is a possible starting point to understand Baldr's cognitive challenges.

Geronimo

Geronimo works well with a mouse. Geronimo has both the skills to engage novel devices successfully and ECDT.

Lavender

Lavender is prevented from acquiring both motor and cognitive skills due to her short span of attention. Lavender can use a push button switch. However, she had problems with the cognitive elements of the motor-skills test.

Nimrod

Nimrod is clumsy with a mouse but can use it. Nimrod results are interesting in that he has consistent results for the 3 tests of varying complexity. Although the results are poor, they do not fall into the noise region. Nimrod has cognitive problems but it would appear that he seems to get confused after a short period of focus. This prevents his progress in any test and learning new skills. In addition because of his need to remain in his comfort zone, his progress tends to be slow in order to get used to new situations. Novel devices are precluded until Nimrod understands their usage or are willing to try the ones that require body contact.

Thor

Thor's only device of use is the eye-tracker. Unfortunately, involuntary muscular activity around the trunk makes the device unreliable. However, Thor was able to indicate by Signing that he had no problems with ECDT. Thor is cognitively able to do the 3 tests but is unable to communicate the proper responses except with the use of Signing.

10.8.2. Non-PNI children

Achilles

Achilles was a typical 5 year old active child. His results in the motor-skills test (COMPTEST) were poor (with all devices) compared with the categorisation test (a more complicated test) in which he succeeded well using a mouse. This suggests that children may have other priorities. Achilles only succeeded the most complicated test using Signing. This suggests the possibility that the use of both the mouse and the test may impose a sufficient load to cause or be an associative cause for failure.

Clementine

Clementine was a 6 year old child. Clementine had relatively well developed overall skills. She did well engaging all the new devices and choosing to use a hybrid device as an alternative best device. She had no problems with all the tests.

Isis

Isis was a typical 5 year old active child. She could use a mouse well. She succeeded in both the COMPTEST and CATTEST using a mouse but had problems with the other devices. She did not succeed in ECDT using a mouse but subsequently succeeded using Signing. This suggests the possibility that the use of both the mouse and the test may impose a sufficient load to cause or be an associative cause for failure.

Lactose

Lactose was a 5 year old child. Lactose had relatively well developed overall skills. She did extremely well engaging all the new devices and had no problems with all the tests.

Mercury

Mercury was a 6 year old child. The mouse was his best device but his skill-set with other devices were poor. This may partly be due to a habit that he had of always needing to support his head with his arm. Although he was requested to not use his arm for support it appeared to be an unbreakable habit. He did not succeed in ECDT using a mouse but subsequently succeeded using Signing. This suggests the possibility that the use of both the mouse and the test may impose a sufficient load to cause or be an associative cause for failure.

Sinbad

Sinbad was a typical 4 year old active child. He could use a mouse well and did fairly well with some of the other devices. He had very poor results in the categorization test (CATTEST) and succeeded well when Signing was used subsequently. This suggests that Sinbad had other priorities and the possibility that the use of both the mouse and the test may impose a sufficient load to cause or be an associative cause for failure. Sinbad could not do ECDT and the results suggested that he focused on the wrong set.

10.8.3. Adults

All the adults provided almost homogenous results. They could use all the devices without much problems and do all the tests.

10.9. Conclusion

Although we do not see a specific physical device being identified as the best device, the models provide a process of identifying devices appropriate for each child. A child in the developmental phase tends to have more variance in the results and it is more appropriate to have a series of tests point to particular solutions with confidence.

11. Conclusions

11.1. Research questions

In order to answer the research question “RQ1. What device helps a child successfully attempt a cognitive test?” we divide the question into the user-device-test triad using the model presented in section 1.3.

User

In order to make a good assessment, we needed to develop strategies that can effectively show how well a child has *successfully attempted* a cognitive test.

We found that common measures that make use of averages are not appropriate for PNI children as they take into account the high number of random errors due to impairments. It was more appropriate to use a measure that uses streaks (consecutive successes) as a basis to separate results that provide no information from intentional responses that show ability. A streak threshold that matches the target application as a criterion for success demanded a performance from the child that matched the final use of the device and proved to be an effective measure of successful attempts. The use of trusted benchmarks allow for a differentiation of device effects from cognitive ability. One trusted benchmark was found to be the use of *Signing* which uses a system of communication via a carer. Other trusted benchmarks are established by the pairing of devices that perform to criterion in a motor skills test and exhibit similar results in other tests.

We find a threshold where participants have developed their general skills to a degree that they tend to be good with all devices and the term latent skill was used to describe this. This is in line with work with the brain that suggests the brain uses a selection of specialized functions to perform a particular skill and there is not an area of the brain for a specific skill (Higuchi et al., 2009). We also find that skill levels can be divided into an automative level where the cognitive effort required to handle the device is a minimum and a skill level that requires cognitive association which demands more cognitive effort as indicated by Ericsson (Ericsson, 2006). We find that in a test, attention is divided between the device and the test. Within a test, it is possible to further sub-divide this attention between solution elements and incorrect solution elements.

Device

We needed a process to determine *which of a range of devices* are appropriate for a child and be able to distinguish the effects of the device from cognitive ability.

We used a range of devices that demanded motor control from different regions of the body so that there was a chance that there was no impairment in that region.

A motor skills test was devised that would demand the same sustained device skill performance required by the final application (ECDT). The motor skills test used a streak threshold that would be demanded as a minimum criterion by the target application. Care had to be taken that the motor skills test did not provide practise that might bias the target test application. Strategies described in a previous paragraph using a streak threshold and tested benchmarks then separated device effects from cognitive ability.

Devices were evaluated by non-PNI children and adults to ensure that the device could be operated in general circumstances.

Communication and understanding proved to be a hurdle to the use of novel devices for PNI children. A case found that the child (Baldr) had not initially understood what was required to operate the head-tracker. However, it was observed that he was using head movements to nudge the eye-cursor when using an eye-tracker and a subsequent switch to the head-tracker again proved to be successful. In a case like this, the simpler a device is to operate, the higher the chances of success.

Hybrid devices were successful for participants that were comfortable with multi-tasking. Hybrid devices were built with compensations for the weaknesses of single-mode devices. However, the multi-tasking required to use both modes together was a hurdle that a participant has to be willing to do. A case however, was found that a non-PNI child (Clementine) would do better at using a hybrid than using the individual modes that composed the hybrid. In general latent ability was found to be a key factor in being able to use a device. A certain threshold of general cognitive development provided enough nurturing of a spread of abilities that enabled mastering the skill for a novel device.

Results suggested that devices may play a significant role in influencing the success of a test for children of a developing age (4 to 5 years old). Learning tests (ECDT) carried out on non-PNI children provided negative results when using a well-used device (mouse) but when Signing was employed subsequently showed that they had the correct solutions all along.

Tests

There was a need to be able to separate the issues of the test from the issues of the device. In order to do that *appropriate tests* needed to be conceived to show device effects and cognitive effects.

It was recognised that all tests carry a cognitive load. The cognitive load may require sets of different skills to handle. Even so elements of the same types of cognitive loading may be present in different tests. For example, although a motor skills test requires device handling ability, some cognitive effort had to be devoted to recognising images so that correct responses can be provided. That cognitive effort is required in presumably all tests. The use of a set of tests which checks fundamental cognitive ability before reaching a level of complexity of the target application test (ECDT) was found to be a useful strategy in the case of PNI children for checking test effects from device effects.

A motor skills test (COMPTEST) designed to be compatible with the target application test was required to separate concerns of cognitive ability from device handling ability. A light cognitive test (CATTEST) was a useful addition in the case of PNI children as it could show basic cognitive problems and allow for a schedule of re-training and re-testing without compromising the target application test that had a higher cognitive complexity.

Learning tests like ECDT need to have a success criterion based on consecutive successes (a success streak threshold) to provide indication that the participants have learnt. There is finer granularity of results to indicate some form of understanding based on an interpretation of the success streak size (maximum streak size). The interpretation however has to be reconciled with random errors especially from PNI children who are liable to produce a higher rate of errors due to their impairments. This study suggests a specific threshold for the streak size where interpretation is not fruitful based on the empirical study. The integration of interpretation of the maximum-streak measure from motor skills test to a high level psychological test such as ECDT allows a perspective on performance of PNI children across different tests. For instance, a case of Nimrod who took all the tests in the test suite with increasing cognitive complexity showed consistent results with similar maximum streak size (5, 6, and 7) in the intention range in all three tests. The numbers suggests a ceiling to performance that is independent of cognitive complexity.

11.2. Relation to work done by others and contribution to knowledge

Current studies done using a learning test like ECDT (Khan, 2009; Sharp & Evans, 1981) have problems as they only make evaluations of children who can succeed in the test and can be considered as a PASS/FAIL test. The problems occur for the children who fail the test. If the test is a stake holder test that has significant consequences, disagreements form with parents who feel that the evaluation of the cognitive ability of their child is inaccurate. Other tests of similar nature like the Wisconsin Card Sorting Test (WCST) (D. A. B. Grant, Esta A., 1981, 2003) are primarily designed for healthy participants and developed metrics do not seem to take into account the high rate of random errors that are made by PNI children using input devices.

This study makes a contribution by suggesting that results can have a finer granularity and more discrimination of outcomes can be made for tests such as ECDT by integrating the interpretations of WCST with a maximum-streak measure (Gan, Frank, Amirabdollahian, Sharp, & Rainer, 2014a; Gan et al., 2014b) used for differentiating random errors from intentional responses. Using this measure, a more meaningful conversation can be obtained with PNI parents because there is more resolution within the user-device-test triad. The parents can more readily agree to results because it is not a PASS/FAIL conclusion. The extension of ECDT to include streak indications of failure to maintain set and the monitor of wrong dimensional outcomes show how the children are responding.

This study extends ECDT as a single test to a test battery of three tests (including COMPTEST and CATTEST). The suite of 3 tests of varying cognitive complexity enables a finer resolution as to the level of cognitive ability and allow for a system of re-training and re-testing. Testing can progress from the test with the simplest cognitive complexity to the most complex one. Failures in the first two tests allow for a period of re-training or developmental progress before the test is re-applied to check for improvements.

Current studies tend to be singular in nature and concentrate on assessments of device skills (Man & Wong, 2007) or cognitive testing of PNI children (Nadeau et al., 2008). This study approached the problem of testing PNI children in a more holistic approach where the user-device-test triad is recognised instead of isolated tests which is important when children of a developing age are concerned. This operational model allows a structured comparison of interactions which can be extended. The results with non-PNI children as a control group allow a comparison to be made to indicate the extent that outcomes can be different between the two groups. We can also show the additional behaviour that non-PNI children would have in the test outcomes. Although there are lots of studies with non-PNI children in using a mouse (Donker & Reitsma, 2007b; Lane & Ziviani, 2010), studies of non-PNI children using non-hand held devices need searching. Comparison with adult participants as another control group shows a trend of increasing novel device handling ability as a result of developmental maturity.

This study considered an initial theoretical model of device skill required based on existing work on cognitive complexity (De Groot & de Groot, 1978; Fischer, 1980) and found that the model had to be modified to take into account latent skills as the result of developmental maturity.

11.3. Ethical issues

A concern in the testing of children is the level of fatigue that will be imposed by the studies. Exploratory pilot testing is useful to determine whether the time allotted and the nature of the tests prove to be demanding for the children. In general it was found that testing with children should not exceed an hour. The better indication is that the body language of the child needs to be observed. In addition, it is helpful to establish a level of familiarity in order that the researcher gets an idea of patterns of behaviour of individual children especially for a PNI child. Before testing it is useful to engage in small talk with the child to set the child at ease as well as establish a relationship that does not place the adult in the usual authoritarian role. Children are then more comfortable with making requests and providing feedback if they are fatigued. Unfortunately, learning tests sometimes require a spark of inspiration and that arrives randomly, possibly requiring memory consolidation. The tests carried out terminated after a specific interval as children would complain of tests running too long but future tests may need to run differently.

Many children try hard to please adults and the researcher has to be careful that some children perceive failure in a test too personally and become emotionally distressed. It was required to explain beforehand that any tests conducted should be perceived as trying out toys that sometimes may not work and any failure is due to the toy rather than the individual.

PNI children can have certain unexpected sensitivities. For example, Nimrod was very uncomfortable with being out of his comfort zone. Any interaction with new devices that have to be worn makes Nimrod panic. It was therefore a good idea to arrange sessions observing PNI children in different classes before hand to get used to children's abilities, challenges and reactions. A very wide range of abilities and behaviour was encountered in the studies. The presence of a carer for a PNI child was extremely useful for the studies. From that perspective, testing of PNI children seem to involve two stages: (1) a view where a carer is present and the contributions of the carer (2) a result where the carer is not present and the reasons for any differences.

The presence of a carer is required as gestures used by the PNI child may be unfamiliar to the researcher. In addition, if the PNI child requires structural body support, the best people to handle it are the carers. For example, when Apollo used an eye-tracker, his neck and head had to be supported and that was best left to the carer. Some PNI children need to try hard to overcome their impairments and a carer would be able to distinguish if the child was just providing a serious attempt from levels of discomfort.

11.4. Limitations

The lack of required types of data is one limitation of this study. There is a difficulty in getting parents to commit their children for experimental studies. The number of specific types of participants tends to be small. The testing of children of developing age (4 to 6 years old) tends to be challenging as there can be large variation in behaviour. A child that is not used to a structured learning environment of sitting in a class for a length of time in a focussed activity is not usually an ideal participant. The yield of the study could not be foreseen in a lot of cases.

Arranging to have an overview of PNI children in a classroom environment of all children within a certain educational ability group was useful for selecting participants. There was a huge range of abilities due to the diversity of impairments. There was a tendency for the very severely affected children whose parents give consent for testing. This provided a strong bias towards poorer success rates for PNI children due to cognitive problems. This coupled with the tendency of children at that mental age to be highly excitable can render a participant unsuitable for some sessions.

For PNI children, the availability of carers was important for this study because of the use of Signing. The carer was required to interpret the meanings of large variations of signs depending on the child. At times the child would also provide gestures that are not part of the documentation accompanying the child which require the knowledge of a carer to interpret. For a large part of this study, carers were not available.

Carers not being available also meant that windows for available testing become more restrictive. The carers had to be there in case of problems with discerning levels of discomfort the PNI children often suffer because of their condition. Carers may also be required for providing unforeseen levels of structural body support to investigate possibilities with devices that could be modified to assist areas affected by weak motor control.

PNI children require regular medication which will not be publicised. The schedule may disrupt a test session and it is not always possible to avoid such occurrences which tend to decrease the yield of a study. PNI children often suffer from personal and medical conditions which makes them unavailable at scheduled times for several sessions.

Many PNI children communicated with a symbol set that is severely restricted (sometimes using only a few symbols) and some tests could not be carried out because there were challenges faced by both the carer and the researcher in trying to explain the procedures of the test.

PNI children require tests that have simple visual representations. In order not to bias ECDT, it was not possible to use the simplest representation in the main tests preceding ECDT. A child who is unable to progress to ECDT however, could be re-tested with tests with simpler visual represenations as diagnostic measures if required.

The use of equipment is fairly mobile and provides an advantage as the testing can be carried out on school premises. However, ideal conditions for the equipment placement will be rare and exact replication of placement for all tests are not possible. One problem is the EEG headset is sensitive to radio black spots and tend to fail in certain environments because it could not communicate with the Bluetooth receiver. The ideal case of having the research monitor and laptop placed opposite to the test screen monitor so that the participant cannot see the research monitor cannot be avoided because of the need to get the headset side by side with the laptop.

Ethics approval from the university required for children under the age of 5 proved to be more involved and required a longer time. As the periods of study required for the exploratory study and the main study required a span of time two approvals were required.

11.5. Future work

11.5.1. Bio-modal data

Multi-modal devices provide more and better levels of information which can be exploited. This study was able to collect EEG data as well as eye-tracking data for participants while they undertook ECDT.

For EEG data this includes the EEG time series that can be analysed using suitable processing software (Delorme & Makeig, 2004). Previous work indicates a possibility of establishing correlation with learning (Chabot, di Michele, Prichep, & John, 2001; Ciçek & Nalçaci, 2001). It was suggested that there was a correlation between higher EEG alpha power levels during rest and performance in cognitive tests similar (WCST) to ECDT. A negative correlation was suggested between alpha power levels and good performance in WCST because larger alpha suppression indicates higher cortical activity. The interest from the perspective of this thesis is whether the same correlations would be observed in PNI children and more importantly if it is possible to differentiate between PNI children who are capable of ECDT and those who are not so that the chidren do not actually require physical manipulation of input devices.

For eye-track data fixation of solutions compared with actual responses may reveal a greater efficiency for establishing learning (E. R. Grant & Spivey, 2003; Nakano, 1971; Thomas & Lleras, 2006).

11.5.2. Extension of the user-device-triad

The use of the user-device-test model in this study has established a framework to base further work.

In terms of the user, a large number of user attributes could be used for a study. These include fatigue, experience, motivation, stress and consolidation of memory. Fatigue could involve studying children at different times of the day as they progress through different physical and mental routines. Experience could be a study of the number of hours a child has at practicing with the devices. This study would be closely tied to the consolidation of memory which involves the observation of device performance for different periods after use of the device. Motivation could involve the stimulation of engagement. Some children are drawn towards particular sounds (such as a boing sound) or certain cartoon imagery which would promote engagement or distractions could be provided. Stress could involve specific time durations required to meet objectives. The role of emotional factors in the study such as motivation, stress, and fatigue would act to change the amount of cognitive effort that is placed into a particular task such as the operation of a device.

In terms of the device, a range of other devices could be used to expand the study. The use of Signing provides a good initial guide to the requirements of a user for a good device. The use of Signing in this study suggests that two further devices that may prove useful. The first is a touch-screen device as children are more used to these devices with the explosion of smartphones and tablets which tend to use these devices. The reason a touch-screen device was not considered in the studies was due to the fact that touch-screen devices had not taken off in popularity when this PhD was started. Also, a touch-screen device was seen as one requiring motor movements and there was already a device in the studies which used motor movements. The other devices considered tried to use a different part of the body to interact with the computer. The second is the Kinect™ which is able to track limb movements. The Kinect™ could be used to recognise common Signing gestures that are used by the PNI children.

A detailed study of Signing itself should prove useful. We need to understand more about the different approaches that PNI children use for Signing as each impairment forces the use of different Signing mechanisms. Carers also look for other accompanying gestures during Signing to support the communication. If a machine like the Kinect™ is required to replace Signing, a detailed understanding of all gestures used is important.

In terms of tests, the three used in the current study could be substituted to provide the variation of complexities required. For example, replacing ECDT with the Wisconsin Card Sorting Test (WCST) is a fairly complementary substitution.

If greater numbers of results become possible, the study would be able to use the triad as three independent variables in a multi-factorial design which would open up the use of ANOVA for repeated measures.

11.5.3. iCASE work

Part of the iCASE requirement was to produce a version of Psyborg with capability to register users and store study data on a web server. The motivation would be that participant recruitment would be increased greatly due to the widespread penetration of the web.

The Psyborg 2 software is already a shrink wrapped prototype that can be delivered as a single executable. Only the registration of a Psyborg user is implemented as proof of concept for web capability. The storage capability requires implementation.

Bibliography

Albouy, G., Sterpenich, V., Balteau, E., Vandewalle, G., Desseilles, M., Dang-Vu, T., . . . Degueldre, C. (2008). Both the hippocampus and striatum are involved in consolidation of motor sequence memory. *Neuron, 58*(2), 261-272.

Allsop, D., Delic, E., Della Torre, M., Do, N., King, W., Le, T. T. T., . . . Thie, J. (2006). Method and System for Detecting and Classifying Mental States: Google Patents.

Altmann, E. M., & Burns, B. D. (2005). Streak biases in decision making: Data and a memory model. *Cognitive Systems Research, 6*(1), 5-16.

Amirabdollahian, F., Gomes, G. T., & Johnson, G. R. (2005). The peg-in-hole: a VR-based haptic assessment for quantifying upper limb performance and skills. *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, 422-425.

Amirabdollahian, F., Munih, M., Kouris, F., Laudanna, E., Stokes, E., & Johnson, G. (2005). The I-match project: A VR based system to allow matching of an optimum interface to a user of assistive technology. *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, 526-529.

Andersen, S. M., Glassman, N. S., Chen, S., & Cole, S. W. (1995). Transference in social perception: The role of chronic accessibility in significant-other representations. *Journal of Personality and Social Psychology, 69*(1), 41.

Ask-a-mathematician. (2014). Ask a Mathematician. http://www.askamathematician.com/2010/07/q-whats-the-chance-of-getting-a-run-of-k-successes-in-n-bernoulli-trials-why-use-approximations-when-the-exact-answer-is-known/. Retrieved June, 2014

ASL. (2014). Applied Science Laboratories (ASL) Mobile Eye-XG Eye Tracking Glasses; http://www.asleyetracking.com/Site/Products/MobileEyeXG/tabid/70/Default.aspx. Retrieved August, 2014

Atkinson, R. C., & Shiffrin, R. M. (1971). *The control processes of short-term memory*: Institute for Mathematical Studies in the Social Sciences, Stanford University.

Ayton, P., & Fischer, I. (2004). The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness? *Memory & cognition, 32*(8), 1369-1378.

Baddeley, A. (1983). Working memory. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 302*(1110), 311-324.

Barreto, A. B., Scargle, S. D., & Adjouadi, M. (2000). A practical EMG-based human-computer interface for users with motor disabilities. *Journal of Rehabilitation Research & Development, 37*(1), 53-64.

Bax, M., Tydeman, C., & Flodmark, O. (2006). Clinical and MRI Correlates of Cerebral Palsy. *JAMA: The Journal of the American Medical Association, 296*(13), 1602-1608. doi: 10.1001/jama.296.13.1602

Betke, M., Gips, J., & Fleming, P. (2002). The camera mouse: visual tracking of body features to provide computer access for people with severe disabilities. *Neural Systems and Rehabilitation Engineering, IEEE Transactions on, 10*(1), 1-10.

Birbaumer, N. (2006). Breaking the silence: Brain–computer interfaces (BCI) for communication and motor control. *Psychophysiology, 43*(6), 517-532.

Biswas, K., & Basu, S. K. (2011). *Gesture Recognition using Microsoft Kinect®.* Paper presented at the Automation, Robotics and Applications (ICARA), 2011 5th International Conference on.

Blaye, A., & Bonthoux, F. (2001). Thematic and taxonomic relations in preschoolers: The development of flexibility in categorization choices. *British Journal of Developmental Psychology, 19*(3), 395-412.

Bleiweiss, A., Eshar, D., Kutliroff, G., Lerner, A., Oshrat, Y., & Yanai, Y. (2010). *Enhanced interactive gaming by blending full-body tracking and gesture animation.* Paper presented at the ACM SIGGRAPH ASIA 2010 Sketches.

Boston College. (2013). CameraMouse. http://www.cameramouse.org/. Retrieved June, 2013

Bowler, M., Amirabdollahian, F., & Dautenhahn, K. (2011). *Using an embedded reality approach to improve test reliability for NHPT tasks.* Paper presented at the Rehabilitation Robotics (ICORR), 2011 IEEE International Conference on.

Broomfield, J., & Dodd, B. (2004). Children with speech and language disability: caseload characteristics. *International Journal of Language & Communication Disorders, 39*(3), 303-324.

Bruckman, A., Bandlow, A., & Forte, A. (2002). HCI for kids: Lawrence Erlbaum Associates.

Burgess, N. (2006). Spatial memory: how egocentric and allocentric combine. *Trends in cognitive sciences, 10*(12), 551-557.

Cans, C. (2000). Surveillance of cerebral palsy in Europe: a collaboration of cerebral palsy surveys and registers. *42*(12), 816-824.

Cans, C., Surman, G., McManus, V., Coghlan, D., Hensey, O., & Johnson, A. (2004). *Cerebral palsy registries.* Paper presented at the Seminars in pediatric neurology.

Ch'ien, L. T., Aur, R. J., Verzosa, M. S., Coburn, T. P., Goff, J. R., Hustu, H. O., . . . Simone, J. V. (1981). Progression of methotrexate-induced leukoencephalopathy in children with leukemia. *Medical and pediatric oncology, 9*(2), 133-141.

Chabot, R. J., di Michele, F., Prichep, L., & John, E. R. (2001). The clinical role of computerized EEG in the evaluation and treatment of learning and attention disorders in children and adolescents. *J Neuropsychiatry Clin Neurosci, 13*(2), 171-186.

Ciçek, M., & Nalçaci, E. (2001). Interhemispheric asymmetry of EEG alpha activity at rest and during the Wisconsin Card Sorting Test: relations with performance. *Biol Psychol, 58*(1), 75-88. doi: S0301-0511(01)00103-X [pii]

Clerget, E., Poncin, W., Fadiga, L., & Olivier, E. (2012). Role of Broca's area in implicit motor skill learning: evidence from continuous theta-burst magnetic stimulation. *Journal of cognitive neuroscience, 24*(1), 80-92.

Dautenhahn, K., Nehaniv, C. L., Walters, M. L., Robins, B., Kose-Bagci, H., Mirza, N. A., & Blow, M. (2009). KASPAR–a minimally expressive humanoid robot for human–robot interaction research. *Applied Bionics and Biomechanics, 6*(3-4), 369-397.

Davidoff, J., & Roberson, D. (2004). Preserved thematic and impaired taxonomic categorisation: A case study. *Language and Cognitive Processes, 19*(1), 137-174.

De Groot, A. D., & de Groot, A. D. (1978). *Thought and choice in chess* (Vol. 4): Walter de Gruyter.

Debaere, F., Wenderoth, N., Sunaert, S., Van Hecke, P., & Swinnen, S. (2004). Changes in brain activation during the acquisition of a new bimanual coordination task. *Neuropsychologia, 42*(7), 855-867.

Deco, G., & Rolls, E. T. (2005). Attention, short-term memory, and action selection: a unifying theory. *Progress in neurobiology, 76*(4), 236-256.

Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of neuroscience methods, 134*(1), 9-21.

Donker, A., & Reitsma, P. (2007a). Aiming and clicking in young children's use of the computer mouse. *Computers in Human Behavior, 23*(6), 2863-2874.

Donker, A., & Reitsma, P. (2007b). Young children's ability to use a computer mouse. *Computers & Education, 48*(4), 602-617.

Duchowski, A. T. (2007). *Eye Tracking Methodology: Theory and Practice*: Springer-Verlag New York, Inc.

Duncan, J., & Owen, A. M. (2000). Common regions of the human frontal lobe recruited by diverse cognitive demands. *Trends in neurosciences, 23*(10), 475-483.

Dunham, P., & Dunham, F. (1995). Developmental antecedents of taxonomic and thematic strategies at 3 years of age. *Developmental Psychology, 31*(3), 483.

Dunn, L. M. (2009). *The British picture vocabulary scale*: GL Assessment Limited.

Dworetzky, B., Herman, S. T., & Tatum IV, W. O. (2011). Artifacts of recording. *6*, 239-266.

Ebert, L. C., Hatch, G., Ampanozi, G., Thali, M. J., & Ross, S. (2012). You Can't Touch This Touch-free Navigation Through Radiological Images. *Surgical innovation, 19*(3), 301-307.
Ekman, P. (1992). Are there basic emotions? *Psychological review, 99*(3), 550-553.
Emotiv. (2013). Emotiv EEG Features. http://www.emotiv.com/eeg/features.php. Retrieved June, 2013
Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science, 11*(1), 19-23.
Erasmus, C. E., Van Hulst, K., Rotteveel, L. J., Jongerius, P. H., Van Den Hoogen, F. J., Roeleveld, N., & Rotteveel, J. J. (2009). Drooling in cerebral palsy: hypersalivation or dysfunctional oral motor control? *Developmental Medicine & Child Neurology, 51*(6), 454-459.
Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. *The Cambridge handbook of expertise and expert performance*, 683-703.
Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological review, 100*(3), 363.
Fadiga, L., Craighero, L., & D'Ausilio, A. (2009). Broca's area in language, action, and music. *Annals of the New York Academy of Sciences, 1169*(1), 448-458.
Fischer, K. W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *87*(6), 477.
Fisher, A. V. (2011). Automatic shifts of attention in the Dimensional Change Card Sort task: Subtle changes in task materials lead to flexible switching. *Journal of experimental child psychology, 108*(1), 211-219.
Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of experimental psychology, 47*(6), 381-391.
Folk, C. L., Remington, R. W., & Johnston, J. C. (1992). Involuntary covert orienting is contingent on attentional control settings. *Journal of Experimental Psychology: Human perception and performance, 18*(4), 1030.
Folk, C. L., Remington, R. W., & Wright, J. H. (1994). The structure of attentional control: contingent attentional capture by apparent motion, abrupt onset, and color. *Journal of Experimental Psychology: Human perception and performance, 20*(2), 317.
Gan, H., Frank, R., Amirabdollahian, F., Sharp, R., & Rainer, A. (2014a). Bio-digital device impact on a constant load cognitive test of children with physical and neurological impairments. *International Journal of Advances in Computer Science & Its Applications, 4*(4), 99 - 105.
Gan, H., Frank, R., Amirabdollahian, F., Sharp, R., & Rainer, A. (2014b). Development of the maximum-streak measure for evaluating the suitability of non-handheld devices in cognitive tests of Physically and Neurologically Impaired(PNI) children. *International Journal of Advances in Computer Science & Its Applications, 4*(4), 130 - 136.
Gan, H., Frank, R., Amirabdollahian, F., Sharp, R., & Rainer, A. (2014c). Use of re-attempts measure for evaluating device test results of children with neurological impairments. *Human System Interactions (HSI), 2014 7th International Conference on*, 206 - 211. doi: 10.1109/HSI.2014.6860476
Gastaut, M. (1952). *Etude electrocorticographique de la reactivite des rythmes rolandiques.* Paper presented at the Revue Neurologique.
Gerber, R. J., Wilks, T., & Erdie-Lalena, C. (2010). Developmental milestones: Motor development. *Pediatrics in Review, 31*(7), 267-277.
Ghahramani, Z., Wolpert, D. M., & Jordan, M. I. (1995). Computational Structure of coordinate transformations: A generalization study. *Advances in neural information processing systems*, 1125-1132.
Ghahramani, Z., Wolpert, D. M., & Jordan, M. I. (1996). Generalization to local remappings of the visuomotor coordinate transformation. *The Journal of neuroscience, 16*(21), 7085-7096.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of experimental psychology, 38*(4), 404.

Grant, D. A. B., Esta A. (1981, 2003). Wisconsin Card Sorting Test® (WCST). Retrieved 3/4/2012, 2012, from https://shop.acer.edu.au/acer-shop/group/QBQ/32

Grant, E. R., & Spivey, M. J. (2003). Eye Movements and Problem Solving. *Psychological Science, 14*(5), 462-466.

Grol, M. J., de Lange, F. P., Verstraten, F. A., Passingham, R. E., & Toni, I. (2006). Cerebral changes during performance of overlearned arbitrary visuomotor associations. *The Journal of Neuroscience, 26*(1), 117-125.

Hanna, L., Risden, K., & Alexander, K. (1997). Guidelines for usability testing with children. *4*(5), 9-14.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Advances in psychology, 52*, 139-183.

Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature Reviews Neuroscience, 11*(7), 523-532.

Higuchi, S., Chaminade, T., Imamizu, H., & Kawato, M. (2009). Shared neural correlates for language and tool use in Broca's area. *Neuroreport, 20*(15), 1376-1381.

Hill, N. M., & Schneider, W. (2006). Brain changes in the development of expertise: Neuroanatomical and neurophysiological evidence about skill-based adaptations. *The Cambridge handbook of expertise and expert performance*, 653-682.

Huizinga, M., & van der Molen, M. W. (2007). Age-group differences in set-switching and set-maintenance on the Wisconsin Card Sorting Task. *Developmental Neuropsychology, 31*(2), 193-215.

Irons, J. L., Folk, C. L., & Remington, R. W. (2012). All set! Evidence of simultaneous attentional control settings for multiple target colors. *Journal of Experimental Psychology: Human Perception and Performance, 38*(3), 758.

Iverson, J. M., & Braddock, B. A. (2011). Gesture and motor skill in relation to language in children with language impairment. *Journal of Speech, Language, and Hearing Research, 54*(1), 72-86.

Jacob, R., Leggett, J., Myers, B., & Pausch, R. (1993). An Agenda for Human-Computer Interaction Research: Interaction Styles and Input/Output Devices. *Behaviour and Information Technology*. doi: citeulike-article-id:421453

Jacob, R. J. (1991). The use of eye movements in human-computer interaction techniques: what you look at is what you get. *ACM Transactions on Information Systems (TOIS), 9*(2), 152-169.

Jacob, R. J., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *Mind, 2*(3), 4.

Jacobson, L., Flodmark, O., & Martin, L. (2006). Visual field defects in prematurely born patients with white matter damage of immaturity: a multiple-case study. *Acta Ophthalmologica Scandinavica, 84*(3), 357-362.

JASPER, H. H. (1958). The ten twenty electrode system of the international federation. *Electroencephalography and clinical neurophysiology, 10*, 371-375.

Johnson, S. C., Baxter, L. C., Wilder, L. S., Pipe, J. G., Heiserman, J. E., & Prigatano, G. P. (2002). Neural correlates of self-reflection. *Brain, 125*(8), 1808-1814.

Kahol, K., Leyba, M. J., Deka, M., Deka, V., Mayes, S., Smith, M., . . . Panchanathan, S. (2008). Effect of fatigue on psychomotor and cognitive skills. *The American Journal of Surgery, 195*(2), 195-204.

Karni, A., & Bertini, G. (1997). Learning perceptual skills: behavioral probes into adult cortical plasticity. *Current opinion in neurobiology, 7*(4), 530-535.

Karni, A., Meyer, G., Rey-Hipolito, C., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1998). The acquisition of skilled motor performance: fast and slow experience-driven changes in primary motor cortex. *Proceedings of the National Academy of Sciences, 95*(3), 861-868.

Khan, C. (2009). *Development of a tool to assess cognitive ability in school aged children with cerebral palsy.* (MSc), University of Hertfordshire.
Kieras, D. E., & Santoro, T. P. (2004). *Computational GOMS modeling of a complex team task: Lessons learned.* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems.
Kort, B., Reilly, R., & Picard, R. W. (2001, 2001). *"An affective model of interplay between emotions and learning: reengineering educational pedagogy-building a learning companion,".* Paper presented at the Advanced Learning Technologies, 2001. Proceedings. IEEE International Conference on.
Kuhtz-Buschbeck, J., Boczek-Funcke, A., Illert, M., Joehnk, K., & Stolze, H. (1999). Prehension movements and motor development in children. *Experimental Brain Research, 128*(1-2), 65-68.
Laird, J. E., Rosenbloom, P. S., & Newell, A. (1984). *Towards Chunking as a General Learning Mechanism.* Paper presented at the AAAI.
Lamble, D., Kauranen, T., Laakso, M., & Summala, H. (1999). Cognitive load and detection thresholds in car following situations: safety implications for using mobile (cellular) telephones while driving. *Accident Analysis & Prevention, 31*(6), 617-623.
Lane, A. E., & Ziviani, J. M. (2010). Factors influencing skilled use of the computer mouse by school-aged children. *Computers & Education, 55*(3), 1112-1122.
Lazzari, S., Vercher, J.-L., & Buizza, A. (1997). Manuo-ocular coordination in target tracking. I. A model simulating human performance. *Biological cybernetics, 77*(4), 257-266.
Libet, B., Wright Jr, E. W., Feinstein, B., & Pearl, D. K. (1993). Subjective referral of the timing for a conscious sensory experience *Neurophysiology of Consciousness* (pp. 164-195): Springer.
Lisberger, S. G., Morris, E., & Tychsen, L. (1987). Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Annual review of neuroscience, 10*(1), 97-129.
Longo, L., & Barrett, S. (2010). A computational analysis of cognitive effort *Intelligent Information and Database Systems* (pp. 65-74): Springer.
Longstaff, A. (2000). lnstant Notes in Neuroscience. *The United Kingdom: Bios Scientific Publishers Limited*.
Longstaff, A. (2005). Instant notes in neuroscience (BIOS instant notes): Abingdon: Taylor & Francis.
Mahncke, H. W., Connor, B. B., Appelman, J., Ahsanuddin, O. N., Hardy, J. L., Wood, R. A., . . . Merzenich, M. M. (2006). Memory enhancement in healthy older adults using a brain plasticity-based training program: a randomized, controlled study. *Proceedings of the National Academy of Sciences, 103*(33), 12523-12528.
Majaranta, P., & Räihä, K.-J. (2002). *Twenty years of eye typing: systems and design issues.* Paper presented at the Proceedings of the 2002 symposium on Eye tracking research & applications.
Makri, F. S., Philippou, A. N., & Psillakis, Z. M. (2007). Shortest and longest length of success runs in binary sequences. *Journal of Statistical Planning and Inference, 137*(7), 2226-2239.
Man, D. W., & Wong, M. S. (2007). Evaluation of computer-access solutions for students with quadriplegic athetoid cerebral palsy. *Am J Occup Ther, 61*(3), 355-364.
Marcus, G. (2008). *The birth of the mind: How a tiny number of genes creates the complexities of human thought*: Basic Books.
Martinez-Conde, S., Macknik, S. L., Troncoso, X. G., & Hubel, D. H. (2009). Microsaccades: a neurophysiological analysis. *Trends in neurosciences, 32*(9), 463-475.
McCall, R. B., Kennedy, C. B., & Appelbaum, M. I. (1977). Magnitude of discrepancy and the distribution of attention in infants. *Child Development*, 772-785.
Mehrabian, A. (1977). *Nonverbal communication*: Transaction Publishers.
Miller, G. A. (1956). The magical number seven, plus or minus two: some limits on our capacity for processing information. *Psychological review, 63*(2), 81.

Mitchell, D. P., & Netravali, A. N. (1988). *Reconstruction filters in computer-graphics.* Paper presented at the ACM Siggraph Computer Graphics.
Mitchell, H. B. (2007). Multi-Sensor Data Fusion - An Introduction: Springer - Verlag.
Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The Unity and Diversity of Executive Functions and Their Contributions to Complex “Frontal Lobe” Tasks: A Latent Variable Analysis. *Cognitive Psychology, 41*(1), 49-100. doi: 10.1006/cogp.1999.0734
Mochizuki-Kawai, H. (2008). [Neural basis of procedural memory]. *Brain and nerve= Shinkei kenkyu no shinpo, 60*(7), 825-832.
Muthukumaraswamy, S. D., Johnson, B. W., & McNair, N. A. (2004). Mu rhythm modulation during observation of an object-directed grasp. *Cognitive Brain Research, 19*(2), 195-201.
Nadeau, L., Routhier, M. E., & Tessier, R. (2008). The performance profile on the Wisconsin Card Sorting Test of a group of children with cerebral palsy aged between 9 and 12 *Dev Neurorehabil* (Vol. 11, pp. 134-140). England.
Nakano, A. (1971). Eye movements in relation to mental activity of problem-solving. *Psychologia, 14*(3-4), 200-207.
Nicolas-Alonso, L. F., & Gomez-Gil, J. (2012). Brain computer interfaces, a review. *Sensors (Basel), 12*(2), 1211-1279. doi: sensors-12-01211 [pii]

10.3390/s120201211

Nyhus, E., & Barceló, F. (2009). The Wisconsin Card Sorting Test and the cognitive assessment of prefrontal executive functions: a critical update. B*rain Cogn, 71(*3), 437-451. doi: S0278-2626(09)00045-1 [pii] 10.1016/j.bandc.2009.03.005
Organization, W. H. (1980). International classification of impairments, disabilities, and handicaps: a manual of classification relating to the consequences of disease, published in accordance with resolution WHA29. 35 of the Twenty-ninth World Health Assembly, May 1976.
Osborne, J. G., & Calhoun, D. O. (1998). Themes, taxons, and trial types in children's matching to sample: Methodological considerations. J*ournal of Experimental Child Psychology, 68(*1), 35-50.
Oskarsson, A. T., Van Boven, L., McClelland, G. H., & Hastie, R. (2009). What's next? Judging sequences of binary events. P*sychological bulletin, 135(*2), 262.
Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: the multiple processes by which past behavior predicts future behavior. P*sychological bulletin, 124(*1), 54.
Oviatt, S. (1997). Multimodal interactive maps: Designing for human performance. H*uman-Computer Interaction, 12(*1), 93-129.
Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. (2003). Cognitive load measurement as a means to advance cognitive load theory. E*ducational psychologist, 38(*1), 63-71.
Pashler, H., Johnston, J. C., & Ruthruff, E. (2001). Attention and performance. A*nnual review of psychology, 52(*1), 629-651.
Patel, A. D. (2003). Language, music, syntax and the brain. N*ature neuroscience, 6(*7), 674-681.
Penfield, W., & Boldrey, E. (1937). SOMATIC MOTOR AND SENSORY REPRESENTATION IN THE CEREBRAL CORTEX OF MAN AS STUDIED BY ELECTRICAL STIMULATION. B*rain, 60(*4), 389-443.
Pennington, L., & McConachie, H. (2001). Predicting patterns of interaction between children with cerebral palsy and their mothers. D*evelopmental Medicine & Child Neurology, 43(*2), 83-90.
Pfurtscheller, G., & Lopes da Silva, F. H. (1999). Event-related EEG/MEG synchronization and desynchronization: basic principles. C*lin Neurophysiol, 110(*11), 1842-1857.
Piaget, J. (1970). Science of education and the psychology of the child. Trans. D. Coltman.
Pizzagali, D. (2007). E*lectroencephalography and High-Density Electrophysiological Source Localization in Book: Handbook of Psychophysiology* (J. T. Cacioppo, L. G. Tassinary, & G. Berntson Eds.): Cambridge University Press.

Plutchik, R. (2001). The Nature of Emotions. A*merican Scientist, 89(*4), 344. doi: citeulike-article-id:7326293
Quinn, P. C., Oates, J., & Grayson, J. O. A. (2004). Early category representation and concepts C*ognitive and language development in children* (pp. 21-60). Maidenhead, BRK, England: Open University Press.
Ren, Z., Meng, J., Yuan, J., & Zhang, Z. (2011). R*obust hand gesture recognition with kinect sensor.* Paper presented at the Proceedings of the 19th ACM international conference on Multimedia.
Reyes, C. E., Rugayan, J. L. C., Jason, C., Rullan, G., Oppus, C. M., & Tangonan, G. L. (2012). A study on ocular and facial muscle artifacts in EEG signals for BCI applications. T*ENCON 2012-2012 IEEE Region 10 Conference,* 1-6.
Rhodes, M. G. (2004). Age-related differences in performance on the Wisconsin card sorting test: a meta-analytic review. P*sychology and Aging, 19(*3), 482.
Rosenbaum, D. A., Carlson, R. A., & Gilmore, R. O. (2001). Acquisition of intellectual and perceptual-motor skills. A*nnual review of psychology, 52(*1), 453-470.
Rosenbaum, P., Paneth, N., Leviton, A., Goldstein, M., Bax, M., Damiano, D., . . . Jacobsson, B. (2007). A report: the definition and classification of cerebral palsy April 2006. D*ev Med Child Neurol Suppl, 109,* 8-14.
Rosenbloom, P. S., & Newell, A. (1986). The chunking of goal hierarchies: A generalized model of practice. M*achine learning: An artificial intelligence approach, 2,* 247-288.
Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods. A*pplied Psychology, 53(*1), 61-86.
San Agustin, J., Skovsgaard, H., Mollenbach, E., Barret, M., Tall, M., Hansen, D. W., & Hansen, J. P. (2010). E*valuation of a low-cost open-source gaze tracker.* Paper presented at the Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications.
Schalk, G., & Mellinger, J. (2010). A *Practical Guide to Brain–Computer Interfacing with BCI2000:* Springer.
Shadmehr, R., & Holcomb, H. H. (1997). Neural correlates of motor memory consolidation. S*cience, 277(*5327), 821-825.
Shannon, C. E. (1948). A Mathematical Theory of Communication. *27(*3), 379-423.
Sharp, R. J., & Evans, P. L. (1981). Verbal regulation of behaviour and discrimination shift learning in severely retarded children. B*ehavior Research of Severe Developmental Disabilities, 2(*1), 37-49.
Shiffrin, R. M., & Nosofsky, R. M. (1994). Seven plus or minus two: a commentary on capacity limitations.
Sloutsky, V. M., & Fisher, A. V. (2011). The development of categorization. P*sychology of learning and motivation, 54,* 141-166.
Smith, M. D., & Chamberlin, C. J. (1992). Effect of adding cognitively demanding tasks on soccer skill performance. P*erceptual and Motor Skills, 75(*3), 955-961.
Soleymani, M., Lichtenauer, J., Pun, T., & Pantic, M. (2012). A Multimodal Database for Affect Recognition and Implicit Tagging. A*ffective Computing, IEEE Transactions on, 3(*1), 42-55.
Squire, L. R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. P*sychological review, 99(*2), 195.
Sroufe, L. A., Cooper, R. G., DeHart, G. B., Marshall, M. E., & Bronfenbrenner, U. E. (1992). C*hild development: Its nature and course:* Mcgraw-Hill Book Company.
Staudt, M. (2010). Reorganization after pre-and perinatal brain lesions*. J*ournal of anatomy, 217(*4), 469-474.
Thelen, E. (1995). Motor development: A new synthesis. A*merican psychologist, 50(*2), 79.
Thomas, L. E., & Lleras, A. (2006). Moving eyes and moving thought: The spatial compatibility between eye movements and cognition. J*ournal of Vision, 6(*6), 871-871.

Tobii Technology. (2013). Portable lab - Tobii X2-30 Eye Tracker. http://www.tobii.com/fr/eye-tracking-research/global/products/hardware/tobii-x2-30-eye-tracker/. Retrieved June, 2013

Tulving, E., & Markowitsch, H. J. (1998). Episodic and declarative memory: role of the hippocampus. H*ippocampus, 8(*3).

Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. N*ature, 377,* 155-158.

van Ee, R., van Boxtel, J. J., Parker, A. L., & Alais, D. (2009). Multisensory congruency as a mechanism for attentional control over perceptual selection. T*he Journal of Neuroscience, 29(*37), 11641-11649.

VanLehn, K. (1996). Cognitive skill acquisition. A*nnual review of psychology, 47(*1), 513-539.

Vidal, J. J. (1973). Toward Direct Brain-Computer Communication. A*nnual Review of Biophysics and Bioengineering, 2(*1), 157-180. doi: 10.1146/annurev.bb.02.060173.001105

Wade, M. G., & Whiting, H. T. A. (1986). M*otor development in children: Aspects of coordination and control:* Nijhoff.

Waller, A. (2006). Communication access to conversational narrative. T*opics in Language Disorders, 26(*3), 221-239.

Ware, C., & Mikaelian, H. H. (1986). An evaluation of an eye tracker as a device for computer input2. S*IGCHI Bull., 18(*4), 183-188. doi: 10.1145/1165387.275627

Webb, E. A., & Dattani, M. T. (2010). Septo-optic dysplasia. E*uropean Journal of Human Genetics, 18(*4), 393-397.

Wolff, P. H. (1965). The development of attention in young infants. A*nnals of the New York Academy of Sciences, 118(*21), 815-830.

Wolpaw, J., & Birbaumer, N. (2006). B*rain–computer interfaces for communication and control in Book: Textbook of Neural Repair and Rehabilitation: Volume 1, Neural Repair and Plasticity* (M. Selzer, S. Clarke, L. Cohen, P. Duncan, & F. Gage Eds.): Cambridge University Press.

Wolpaw, J. R., McFarland, D. J., & Vaughan, T. M. (2000). Brain-computer interface research at the Wadsworth Center. R*ehabilitation Engineering, IEEE Transactions on, 8(*2), 222-226. doi: 10.1109/86.847823

Wolpert, D. M., Ghahramani, Z., & Flanagan, J. R. (2001). Perspectives and problems in motor learning. T*rends in Cognitive Sciences, 5,* 487-494.

Xie, B., & Salvendy, G. (2000). Review and reappraisal of modelling and predicting mental workload in single-and multi-task environments. W*ork & stress, 14(*1), 74-99.

Zander, T. O., Kothe, C., Jatzev, S., & Gaertner, M. (2010). Enhancing human-computer interaction with input from active and passive brain-computer interfaces B*rain-computer interfaces* (pp. 181-199): Springer.

Appendix

12. Appendix – Ethics approval (1)

Revised (September 2006)

SCHOOL OF PSYCHOLOGY ETHICS COMMITTEE APPROVAL

Student Investigator: Hock Chye Gan

Title of project: Early Concept Development Test for Children with Complex Physical and Neurological Impairment.

Supervisor: Ray J. Frank

Registration Protocol Number: PSY/12/10/HCG

The approval for the above research project was granted on 16 December 2010 by the Psychology Ethics Committee under delegated authority from the Ethics Committee of the University of Hertfordshire.
The end date of your study is 30 September 2013

Signed:  **Date:** 16 December 2010

Professor Lia Kvavilashvili
Chair
Psychology Ethics Committee

STATEMENT OF THE SUPERVISOR:

From my discussions with the above student, as far as I can ascertain, s/he has followed the ethics protocol approved for this project.

Signed (supervisor): ....

Date: ....

13. Appendix – Ethics Approval (2)

UNIVERSITY OF HERTFORDSHIRE
SCIENCE AND TECHNOLOGY

M E M O R A N D U M

TO **Hock Chye Gan**

CC **Ray Frank**

FROM Dr Simon Trainis, Science and Technology ECDA Chairman

DATE **2/5/13**

Protocol number: COM/PG/UH/00006

Title of study: The use of non-hand held devices to support a cognitive test for Physically and Neurologically Impaired children

Your application for ethical approval has been accepted and approved by the ECDA for your school.

This approval is valid:

From: 2/5/13

To: 1/9/13

Please note:

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1 may need to be completed prior to the study being undertaken.

14. Appendix – Ethics Approval (3)

UNIVERSITY OF HERTFORDSHIRE
SCIENCE AND TECHNOLOGY

M E M O R A N D U M

TO **Hock Chye Gan**

CC **Ray Frank**

FROM Dr Simon Trainis, Science and Technology ECDA Chairman

DATE **9/7/13**

Protocol number: COM/PG/UH/00023

Title of study: The use of non-hand held devices to support a cognitive test for physically and Neurologically Impaired children

Your application for ethical approval has been accepted and approved by the ECDA for your school.

This approval is valid:

From: 9/7/13

To: 30/9/13

Please note:

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1 may need to be completed prior to the study being undertaken.

15. Appendix – Ethics approval (4)

UNIVERSITY OF HERTFORDSHIRE
SCIENCE AND TECHNOLOGY

M E M O R A N D U M

TO **Hock Chye Gan**

CC **Ray Frank**

FROM Dr Simon Trainis Science and Technology ECDA Chairman

DATE **23 July 2013**

Protocol number: aCOM/PG/UH/00006

Title of study: The use of non-hand held devices to support a cognitive test for Physically and Neurologically Impaired children

Your application to extend the existing protocol detailed above has been accepted and approved by the ECDA for your school.

This approval is valid:

From: 23 July 2013

To: 30 April 2014

Please note:

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1 may need to be completed prior to the study being undertaken.

16. Appendix – Ethics Approval (5)

UNIVERSITY OF HERTFORDSHIRE
SCIENCE AND TECHNOLOGY

M E M O R A N D U M

TO **Hock Chye Gan**

CC **Ray Frank**

FROM Dr Simon Trainis Science and Technology ECDA Chairman

DATE **23 July 2013**

Protocol number: aCOM/PG/UH/00023

Title of study: The use of non-hand held devices to support a cognitive test for Physically and Neurologically Impaired children

Your application to extend the existing protocol detailed above has been accepted and approved by the ECDA for your school.

This approval is valid:

From: 23 July 2013

To: 30 April 2014

Please note:

Approval applies specifically to the research study/methodology and timings as detailed in your Form EC1. Should you amend any aspect of your research, or wish to apply for an extension to your study, you will need your supervisor's approval and must complete and submit form EC2. In cases where the amendments to the original study are deemed to be substantial, a new Form EC1 may need to be completed prior to the study being undertaken.

17. Appendix – Pick-N-Drop System Design

17.1. Introduction

This document provides a brief description of the implementation of a device for interacting with a computer using electroencephalography (EEG) and eye-tracking. The hybrid device is able to construct words from a restricted alphabet; spelling out the picture content of images presented. The system is built using pc equipment and off-the-shelf products for EEG (Emotiv™) and eye-tracking (Tobii™). The software was implemented using software libraries of the vendor components and Microsoft™ to interact with the Windows operating system resident on the PC. Development work was carried out using the Microsoft™ Visual Studio Integrated Development Environment (IDE) in C++.

This document describes three different aspects of the hybrid system architecture; a system view, a functional view and a process view. The system view is a general overview describing the system topology, hardware and software components. The functional view is an overview of the functional software blocks that make up the system. The process view describes the run-time architecture in terms of a general overview of some processes and threads of interest. The build view provides an overview of the libraries used.

17.2. System overview

This section refers to figure below.

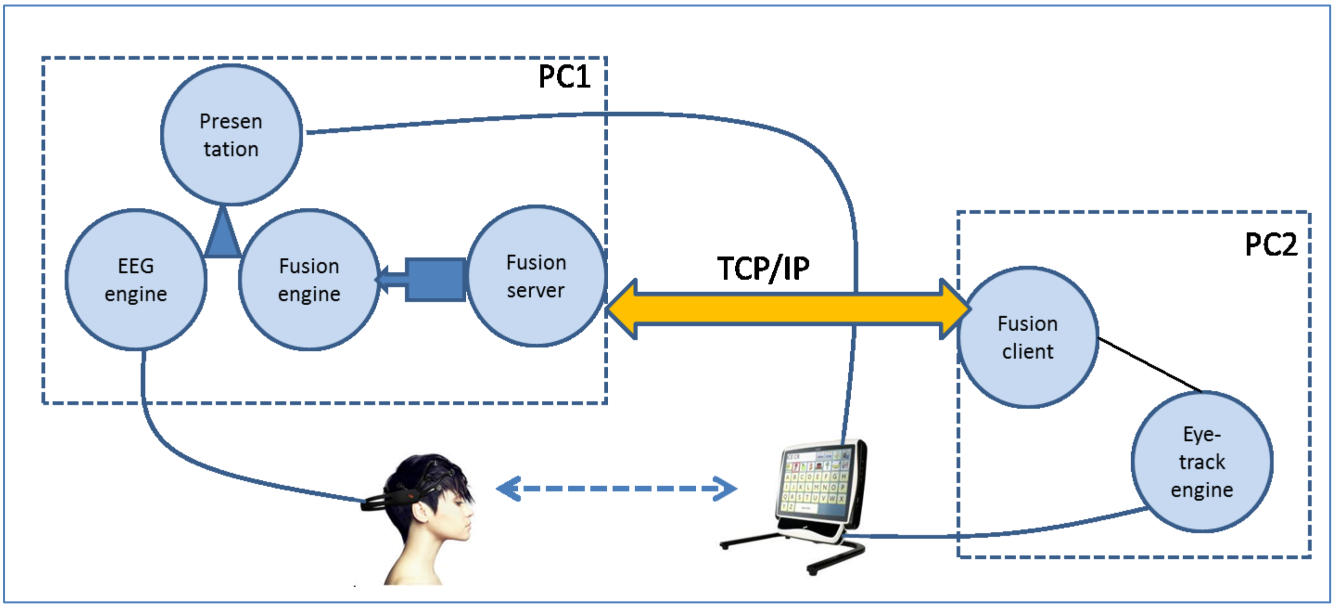


Figure 52 - Overview of EEG and eye-track fusion system

17.2.1. Hardware

A minimal system configuration consists of one Windows PC, the gaming EEG headset from Emotiv (EPOC™) and a remote eye-tracker from Tobii™ (Tobii X Light). If the PC does not have sufficient processing power to encapsulate all the functions, a configuration of two inter-connected PCs (using an Ethernet cable) may be used instead. This is shown in the figure above as PC1 and PC2 connected by the two-way orange arrow which represents a TCP/IP connection. A sufficient single PC configuration consists of a system with Intel™ Duo Core running at 2.13 GHz with 4GB memory using a 256 MB graphics card and 800 MHz DDR2 RAM.

The gaming headset captures the EEG from the head using saline damped pad sensors and translates it to a form that is more amenable for yes/no decisions. This document refers to the translation as Brain-Computer Interface (BCI) Control. The headset is connected to the PC using a blue-tooth wireless link. This is shown in the figure as a solid line from the headset to PC1.

The remote eye-tracker is mounted below a monitor and consists of cameras that pick up images of the eye and specific reflections of infra-red illuminators (also mounted below the monitor) from the cornea of the eye. The eye-tracker is connected to the PC via a USB connector and is shown in the figure as a solid line to PC2.

If a minimal system configuration of one PC is used, only one monitor is required but if two PCs are used, the monitor that is connected to the speller application provides the stimulus which both the eye-tracker and EEG must respond to. This monitor is shown connected to PC1 in the figure by a solid line. PC2 has its own monitor which is not shown that displays the dialog required to run the eye-tracker.

In terms of processing architectures, the gaming headset and eye-tracker can be considered to possess their own processors communicating with the PC. A PC may also have more than one processing core but no specific proprietary optimisation techniques have been employed. Also within the PC, the Graphics Processing Unit (GPU) plays a role in providing fast feedback to the user and researcher.

17.2.2. Software

The software described in this sub-section relates to the software that runs on a deployed machine running on the Windows™ operating system.

Redistributable software libraries are provided by the vendors of the EEG gaming headset (EPOC) and the eye-tracker (Tobii) for interacting with their hardware that runs on the PC. In the case of the EEG headset, the software includes the BCI Control function.

As the software is written in C++ using the Microsoft Visual Studio™ IDE, a redistributable library that supports the C++ structures used was required. The structures include a concurrent queue (describe in a section below) used for inter-task communication that uses a parallel library (structures that can work in a parallel processing environment).

The speller application was built using Microsoft DirectX™ libraries. Most of the DirectX libraries are present in Windows 7™ but some updates may be absent.

The TCP/IP connections for communicating between PCs were built using the Windows™ socket library which is normally present in a Windows distribution.

The proprietary software written takes inputs from both the EPOC and Tobii to control the speller application. Two executables were developed; one (gBrain) for the EEG, fusion function and speller application, the other (gEye) for the eye-track function. The former used a Windows programming model and the latter used the Microsoft Foundation Classes™ (MFC) which is present in the C++ redistributable. The reason for the different programming model is that the code for the eye-track function was modified from older Tobii sample code that used MFC. The speller application used the DirectX library for the media presentation and manipulation on the monitor. Batch scripts written using DOS™ commands was used to launch the two executables in turn.

An installer was also implemented for software deployment based on Flexera's InstallShield™.

17.3. Functional view

Figure 52 - Overview of EEG and eye-track fusion system shows the major software functional blocks of the system as circles. They are subsequently described in the sub-sections below.

17.3.1. EEG engine

The input to the EEG engine is provided by the gaming headset which detects the EEG signals and sends it to the EEG engine. When the speller application is engaged, the EEG Engine performs the BCI Control function on the received EEG signals and makes the translated data available to the Fusion engine. The EEG engine has a vendor part and a proprietary part. The proprietary part takes the output from BCI control, changes it to a yes/no decision and forwards it.

17.3.2. Eye-track engine

The input to the Eye-track engine is provided by the eye-tracker which detects the eye motion and the reflected infra-red and sends it to the Eye-track engine. When the speller application is engaged, the Eye-track engine translates the information from the eye-tracker to spatial positions on the screen. The translated data is passed to the Fusion client who takes care of communication to the Fusion engine which may be resident on another PC. The Eye-track engine has a vendor part and a proprietary part. The proprietary part takes the translated output from the vendor part, performs some data validation and rationalisation for binocular vision and sends the results which include x,y gaze data positions to the Fusion client.

17.3.3. Fusion client

The Fusion client is a TCP client of the Fusion server. The client is a blocking sender. The data is low (less than 200 bytes) so the anticipation is that more efficiency is gained from a blocking client.

The Device client takes care of TCP transport of the eye-track data opaquely by packaging it with a simple proprietary protocol.

The Device client is proprietary software.

17.3.4. Fusion server

The Fusion server is a TCP server. The TCP server is a non-blocking listener. The server creates a specific thread which is blocking to service a TCP client. The Fusion client is the only client for the current implementation.

The Fusion server forwards its data to the Fusion engine using a concurrent queue from the C++ parallel library. The queue bridges the thread from the Fusion server with the thread of the Fusion engine.

The Fusion server and client communicate using a simple proprietary protocol and the data that it manipulates is done in an opaque fashion.

The Fusion server is proprietary software.

17.3.5. Fusion engine

Figure 53 - Details of the Fusion engine show how the Fusion engine is further sub-divided as represented by the circle subset.

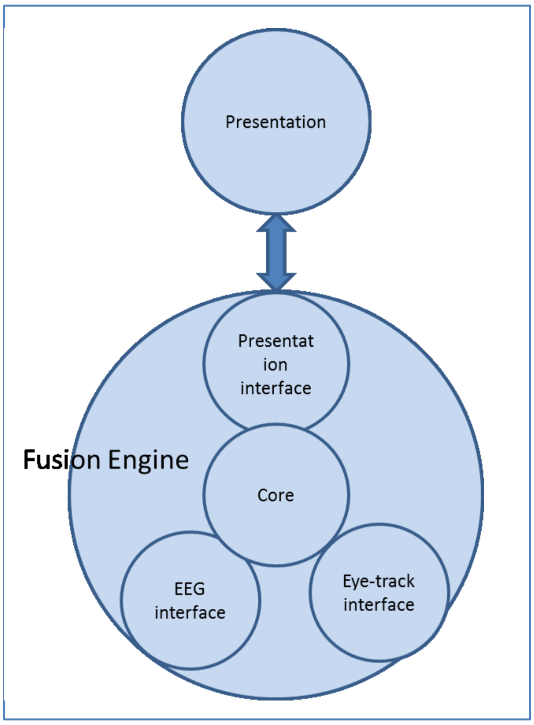


Figure 53 - Details of the Fusion engine

The EEG interface receives updates from the EEG engine. The rates at which the updates arrive depend on the data to be received. The reception rate of raw EEG data samples is the highest and occurs around every 10 ms.

The Eye-track interface is connected to the Fusion Server and receives updates from the Eye-track engine. The rates at which the updates arrive depend on the scan rate of the eye-tracker. The fastest used so far provided is capable at most of producing updates every 25 ms.

The Core polls the EEG and Eye-track interface very 2 ms for new updates. The data is then forwarded to the Presentation Interface to be sent to the Presentation function.

The Fusion engine is proprietary software.

17.3.6. Presentation

The Presentation function takes care of the graphics and audio manipulation for the speller application. In the experimenter's world, the speller application provides the stimulus for the experiment. The human body responds to the stimulus and the EEG and eye movement provides feedback which the hybrid device uses to generate controls to interact with the speller application.

The Presentation function displays some letters, a cursor that maps the eye gaze and some images that the letters can spell out.

<u>Figure 54 - Details of the Presentation function</u> show some of the main software objects that play a role in the speller application.

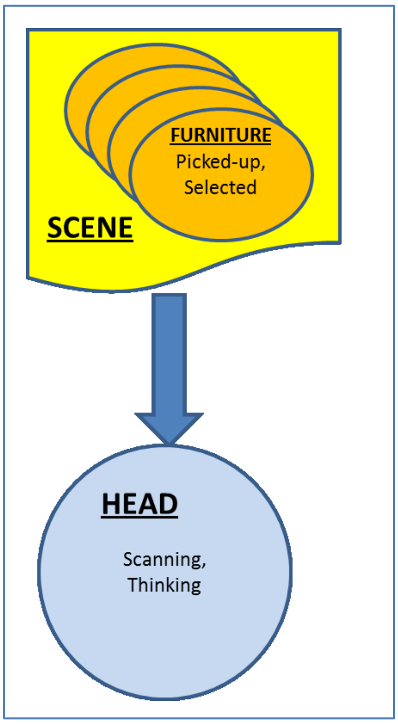


Figure 54 - Details of the Presentation function

The main software classes are the Head, the Scene (collection class) and the Furniture. The Head is a state machine and has two states; SCANNING and THINKING. The Scene is a list of Furniture items. Each Furniture item is also a state machine. The Furniture item has two Boolean state variables; SELECTED and PICKEDUP.

The speller application would first put the Head into a SCANNING state. Letters on the screen would be mapped as Furniture items that are initially not SELECTED and not PICKEDUP. Once the eye fixates on a Furniture item, it will change its state to SELECTED.

The Head then enters the THINKING state. The brain has to produce an EEG pattern that corresponds to picking up the letter (yes decision). If that happens, the letter is piggy-backed onto the eye gaze cursor and the Furniture item also enters the PICKEDUP state.

The Head then reverts back to a SCANNING state where the eye looks for a location to drop the letter. Once the eye fixates on a location, the Furniture item will change its state to not SELECTED.

The Head then re-enters the THINKING state. The brain has to produce an EEG pattern that corresponds to dropping the letter (yes decision). If that happens, the letter is dropped from the eye gaze cursor to the chosen location. The Furniture item then also enters the not PICKEDUP state and the process repeats itself until a word is formed.

The Presentation function is proprietary software.

17.4. Process view

This section describes the process view for the inter-process communications between the EEG (gBrain) and Eye-tracking (gEye) executables. The process view for vendor specific threads, window and console presentations and dialogs are not described.

Figure 55 - Processes and threads for TCP shows the process view for inter-task and inter-core communications. This section makes reference to the figure.

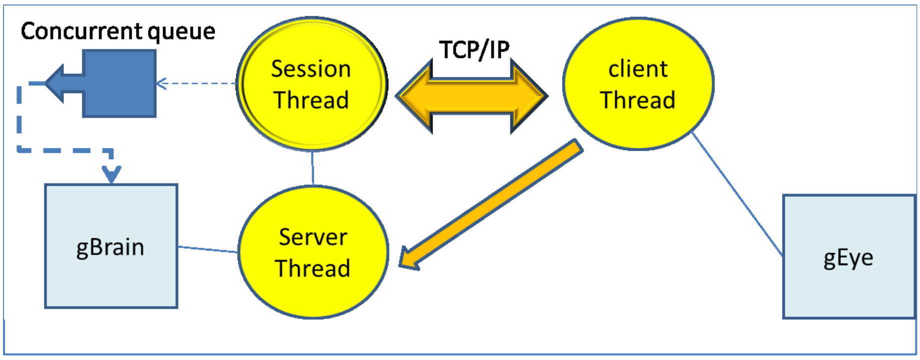


Figure 55 - Processes and threads for TCP

The squares in the figure represent the processes and the ellipses represent the threads.

The system runs by launching the gBrain executable first. The gBrain process will create a thread to run the TCP server. The TCP server listens for a connection from any client. Once the server is up, the client side (gEye) can be launched.

When gEye is launched, the process creates a TCP client thread. The TCP client establishes a connection with the Server thread which has a known IP address and port number. Once communication is established, the Server thread will create a Session thread and passes the data received from the client over to the Session thread. The Session train will respond to the Client thread with its own port address. The Client thread then establishes a dedicated connection with the Session thread when it receives the response.

The Client thread sends eye-track data to the Session thread using the new dedicated connection. The Session thread receives the data and forwards it to the gBrain process using a concurrent queue.

The gBrain process polls the queue regularly for new incoming messages.

18. Appendix – Probability of Streak

18.1. Chance of getting k or more successes in a Bernoulli trial of N throws

18.1.1. Problem

A coin is tossed N times. What is the probability that k or more consecutive heads are obtained within the N tosses. Assume that the coin is fair and the chances of getting a head is ½ (that is p = 0.5). The solution just needs to be an algorithm that can be implemented by a computer. The exact solution is not required.

If R(N,k) is the probability that k heads or more are obtained within N tosses then prove that:

$$R(N,k) = p^k + \sum_{j=1}^{k} p^{j-1}(1-p)R(N-j,k)$$

(Ask-a-mathematician, 2014)

18.1.2. Solution

This solution is meant to be implemented by a computer.

The conditions for getting a streak of k consecutive heads are first examined. The k heads can occur at any point within the N tosses.

18.1.3. Proof

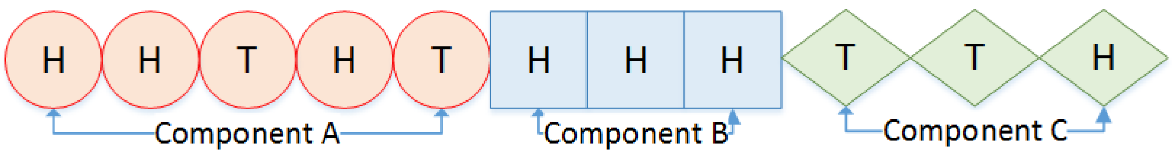


Figure 56 - Components of event sequences in N coin tosses

This section gives a general overview to the solution. Sections following this will go into the details of how the method works. The method consists of categorizing the combination of events in N throws, iterating through every combination of the categories with a view to calculating the probabilities of those categories. In the figure above, a sequence where a required streak of 3 consecutive heads (k = 3) within 11 throws (N = 11) is shown. The sequence is made up of a maximum of 3 components; A, B and C of which only component B is mandatory.

1. Component A: consists of sub-sequences of heads followed by a tail. The figure above shows two sub-sequences, { H, H, T } and { H, T }.
2. Component B: consists of the required consecutive heads. In this case k = 3.
3. Component C: consists of all the combinations that make up the rest of the throws given that the sequence is prefixed with other components. In the figure above, 3 throws remain after being prefixed with Component A and B. This gives 2^3 = 8 combinations that can occur in Component C.

If the number of throws is equal to the required number of consecutive heads (N = k), only Component B exists. If the number of throws exceed the required number of heads by 2 (N = k + 2), combinations of Component A, B and C will exist in all the sequences.

18.1.3.1. The first k heads

The simple case of getting the first k heads is considered. The condition for this is that any sequence before the first k heads contains (k – 1) heads at most, otherwise the event of the first k heads would occur earlier. Before k heads is reached, there must be at least one tail separating a previous sequence of continuous heads. To put this more precisely, consider the case of a sequence of (j – 1) heads followed by a tail at the jth position and then followed by another sequence of events that are a combination of both heads and tails. The basic sequence that occurs (up to getting a run of the first k heads) can be written as

$\{ H_1, H_2, H_3 \ldots H_{j-1}, T_j, H_1, H_2, \ldots H_k \}$ [E.1.1]

If j = 1, the sequence consists of a tail. The highest value that j can have is k. Any combination that precedes a run of the first k consecutive heads can be broken down into a sequence of j-1 heads followed by a tail. For example,

$\{ [H_1, H_2, T], [H_1, H_2, H_3, T], [T], [T], H_1, H_2, \ldots H_k \}$ [E.1.2]

Conversely, the solution to getting all the sequences for the first k heads lies in prefixing heads-tail combinations uniquely to k consecutive heads. The prefixing can be iterative in that established strings of unique heads-tail combinations are prefixed with new combinations.

Another consideration which will be helpful is the sequences formed when the number of trials starts from the smallest possible value. So this is the case where different values of N are considered and the smallest being N = k.

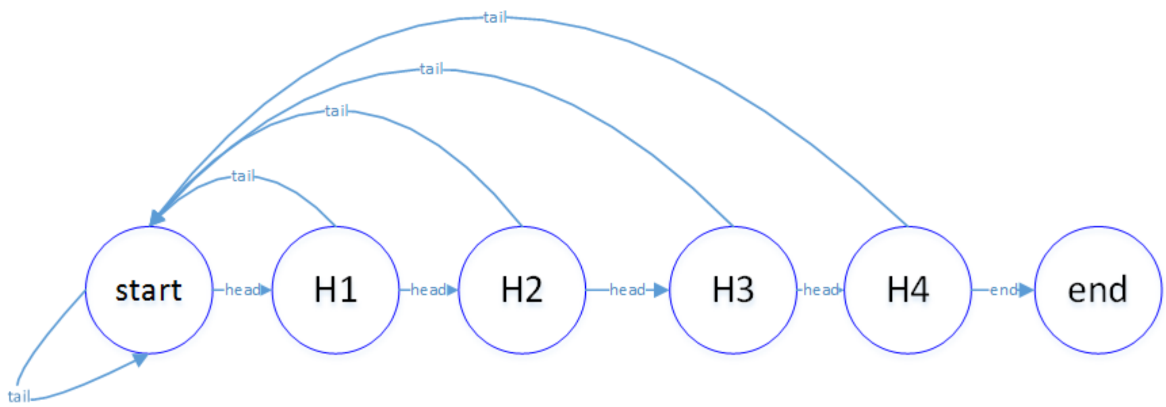


Figure 57 - Transition graph of getting 4 consecutive heads

The figure above shows a transition graph of getting 4 consecutive heads as an example. The occurrence of a tail brings the process to the Start state. Once 4 heads are thrown consecutively, the process ends. If N = 4, no tail can occur. If N = 5, one tail at the Start state is allowed in addition to the transitions for N = 4 to make up 5 throws. The transitions for N = 4 also apply when N =5 as it is possible to stop within 4 throws. If N = 6, the additional prefixes to the transitions of N = 5 are either a tail or a head and tail as the transitions run from the Start to H1 and back to the Start again. Similarly, for N = 7, the additional prefixes to the transitions of N = 6 are either a tail, a head and tail, or two heads and a tail as the transitions cycle through to H2 and back to the Start again. By re-using previous transitions, all the possible transitions to the first 4 consecutive heads can be obtained.

In general, if we let S_N be the sequence of a maximum of N events for getting the first k heads, if N = k, the very first k heads will have no tail preceding it;

$S_N = \{ H_1, H_2, H_3 \ldots H_k \}$.

If N = k + 1, the very first k heads will have in addition to the sequence for $S_{k,}$ one tail preceding it;

$S_N = T_1S_k = \{ T_1, H_1, H_2, H_3 \ldots H_k\}$

where T_1S_k represents the concatenation of a tail with the events in $S_{k,}$

or $S_N = S_k$.

If N = k + 2, the very first k heads will have in addition to the sequences for $S_{k,}$ a head and tail preceding it to make up for k trials in S_k, and in addition to the sequences for S_{k+1}, one tail preceding it to make up for k+1 trials in S_{k+1};

$S_N = T_1S_{k+1}$

or $S_N = H_1T_1S_k$ although strictly speaking this case, is already covered in the previous equation for S_N.

In general, if n represents all the possible values of N, for n - j ≥ 0,

$S_n = H_1H_2H_3 \ldots H_{j-1}T_j\, S_{n-j}$ [E.1.3]

In this way, we have a mechanism for generating every unique combination to get the first k heads. The next thing is to derive the probabilities of each of these combinations of events.

18.1.4. Probability of k heads

The expression for the probability of independent events is required. This states that the probability P(A,B) for two independent events A and B is the product of the individual probabilities:

$P(A,B) = p_A \times p_B$ [E.2.1]

If k heads are required, the probability of getting k heads in k throws is

$P(X = k \text{ heads}) = p^k$ [E.2.2]

There is only one combination of k heads possible and they therefore have to be consecutive.

If the number of throws is greater than k then the probability of getting k consecutive heads or more is also given by

$P(X \geq \text{consecutive } k \text{ heads}) = p^k$ [E.2.3]

This result is due to the requirement for k heads in N throws which include all combinations of throws that follow the k heads to make up the N throws. These combinations can include more heads. A detailed proof will be provided later. Basically the increase in combinations for the number of consecutive k heads is proportionally the same as the total increase of combinations within the N throws.

18.1.4.1. Probability of first k heads

The previous sections provide a basis for calculating the probability of getting first k heads within N throws R(N, k). We can use the law of independent events to calculate the probability of a chain of required events. We will also need the expression to calculate mutually exclusive events as mutually exclusive chain of events can occur.

The probability of each chain of events described by [E.1.3] can be calculated using the law of independent events. The probability of each chain must be added for all chains that make up the ways of getting the first k heads within N throws. It is sufficient to calculate the probability of getting the first k heads to get the probability of getting k heads or more.

From [E.1.3], for $n \geq j$,

$S_n = H_1H_2H_3 \ldots H_{j-1}T_j\, S_{n-j}$

The probability of S_n, $P(S_n)$, is thus the same as the probability of the sequence $H_1H_2H_3 \ldots H_{j-1}T_j$, $P(H_1 \ldots T_j)$ multiplied by the probability of S_{n-j}, $P(S_{n-j})$ from the law of independent events. This is the probability of one chain of events. To get R(N,k) we need to sum the probabilities of all the chains formed by j in the range 1 to k. We also need to consider the cases where n ranges from 0 to k - 1. These are cases where the number of throws is less than k (the number of consecutive heads required).

For n in the range 0 to k – 1, R(n,k) = 0 as the probability of getting k consecutive heads for throws less than k is impossible.

For n = k, $R(k,k) = p^k$

For n > k, the following applies.

The probability of a chain of events consisting of j – 1 heads and one tail using the law of independent events is

$P(H_1 \ldots T_j) = p^{j-1}(1-p)$

Since each chain of events is a mutually exclusive event, the probability of getting j – 1 heads, a tail and more than k consecutive heads is

$$R(k,N) = R(k,k) + \sum_{j=1}^{k} p^{j-1}\,(1-p)R(N-j,k)$$

$$= p^k + \sum_{j=1}^{k} p^{j-1}\,(1-p)R(N-j,k)$$

For example, the probability of getting more than 10 consecutive heads (k = 10) or more within 32 (N = 32) throws is the sum of the probability of

- Getting 10 consecutive heads or more starting with no tail
- Getting 10 consecutive heads or more starting with 1 tail. Since the first throw has one tail, the probability of getting 10 consecutive heads or more is limited to 31 throws
- Getting 10 consecutive heads or more starting with 1head and 1 tail. Since the first throw has 1 head and one tail, the probability of getting 10 consecutive heads or more is limited to 30 throws
- … and so on until there are 9 heads and 1 tail

To get the actual values for the entire sum, we have to start calculations from N = 10 and calculate the sums recursively until we reach N = 32.

- R (10, 10) = p^{10}
- R (11, 10) $= p^{10} + \sum_{j=1}^{10} p^{j-1}\,(1-p)R(11-j,10)$

  $= p^{10} + (1-p)R(10{,}10)$ since R for N < 10 is zero
- R (12, 10) $= p^{10} + \sum_{j=1}^{10} p^{j-1}\,(1-p)R(12-j,10)$

  $= p^{10} + (1-p)R(11{,}10) + p(1-p)R(10{,}10)$
- R(13, 10) $= p^{10} + \sum_{j=1}^{10} p^{j-1}\,(1-p)R(13-j,10)$

  $= p^{10} + (1-p)R(12{,}10) + p(1-p)R(11{,}10) + p^2(1-p)R(10{,}10)$
- … and so on until N = 32.

Note that the last chain of event for getting the first 10 heads within 32 throws (N = 32, k = 10) is { H, T, $H_{1,\ldots}H_{9,}$ T, $H_{1,\ldots}H_{9,}$ T, $H_{1,\ldots}H_{10}$ }.

18.1.5. Discussion

In a practical scenario, it is normally required to know the probability of getting a series of YES/NO responses correct after a set number of trials. A series of YES/NO questions could be asked and if the responder provides a number of consecutively correct responses, the responder would be deemed successful in the test. The solution provided shows that the probability of the responder achieving the results by chance is dependent on the number of trials undertaken and therefore it would depend upon when the responder is able to make the required number of consecutively correct responses. To provide the answer to how likely a responder to such a test succeeds by chance, a table and or a plot of the probabilities for a differing number of trials proves useful.

18.1.6. Plots of probability of up to 10 consecutive successes

As an example, the table below shows the probability P of event X of getting a streak of k heads or more in N throws of a coin. The coin is a fair coin and the probability of getting a heads is 0.5.

N	P(X≥10)	P(X≥9)	P(X≥8)	P(X≥7)	P(X≥6)	P(X≥5)	P(X≥4)	P(X≥3)	P(X≥2)	P(X≥1)
1										0.5
2									0.25	0.75
3								0.125	0.375	0.875
4							0.0625	0.1875	0.5	0.9375
5						0.0313	0.0938	0.25	0.5938	0.9688
6					0.0156	0.0469	0.125	0.3125	0.6719	0.9844
7				0.0078	0.0234	0.0625	0.1563	0.3672	0.7344	0.9922
8			0.0039	0.0117	0.0313	0.0781	0.1875	0.418	0.7852	0.9961
9		0.002	0.0059	0.0156	0.0391	0.0938	0.2168	0.4648	0.8262	0.998
10	0.001	0.0029	0.0078	0.0195	0.0469	0.1094	0.2451	0.5078	0.8594	0.999
11	0.0015	0.0039	0.0098	0.0234	0.0547	0.1245	0.2725	0.5474	0.8862	0.9995
12	0.002	0.0049	0.0117	0.0273	0.0625	0.1394	0.2988	0.5837	0.908	0.9998
13	0.0024	0.0059	0.0137	0.0313	0.0702	0.1541	0.3242	0.6172	0.9255	0.9999
14	0.0029	0.0068	0.0156	0.0352	0.0778	0.1685	0.3487	0.6479	0.9398	0.9999
15	0.0034	0.0078	0.0176	0.039	0.0854	0.1826	0.3723	0.6762	0.9513	1
16	0.0039	0.0088	0.0195	0.0429	0.0929	0.1965	0.395	0.7023	0.9606	1
17	0.0044	0.0098	0.0215	0.0467	0.1003	0.2102	0.4169	0.7262	0.9681	1
18	0.0049	0.0107	0.0234	0.0506	0.1077	0.2237	0.438	0.7482	0.9742	1
19	0.0054	0.0117	0.0254	0.0544	0.1151	0.2369	0.4584	0.7684	0.9791	1
20	0.0059	0.0127	0.0273	0.0582	0.1223	0.2499	0.478	0.787	0.9831	1
21	0.0063	0.0137	0.0292	0.062	0.1295	0.2626	0.4969	0.8041	0.9863	1
22	0.0068	0.0146	0.0311	0.0657	0.1367	0.2752	0.5151	0.8199	0.9889	1
23	0.0073	0.0156	0.0331	0.0695	0.1438	0.2875	0.5327	0.8344	0.9911	1
24	0.0078	0.0166	0.035	0.0732	0.1508	0.2997	0.5496	0.8477	0.9928	1
25	0.0083	0.0175	0.0369	0.077	0.1578	0.3116	0.5659	0.8599	0.9941	1
26	0.0088	0.0185	0.0388	0.0807	0.1647	0.3233	0.5817	0.8712	0.9953	1
27	0.0093	0.0195	0.0407	0.0844	0.1715	0.3348	0.5968	0.8815	0.9962	1
28	0.0098	0.0204	0.0426	0.088	0.1783	0.3462	0.6114	0.891	0.9969	1
29	0.0102	0.0214	0.0445	0.0917	0.1851	0.3573	0.6255	0.8998	0.9975	1
30	0.0107	0.0224	0.0464	0.0953	0.1918	0.3682	0.6391	0.9078	0.998	1
31	0.0112	0.0233	0.0483	0.099	0.1984	0.379	0.6521	0.9153	0.9984	1
32	0.0117	0.0243	0.0502	0.1026	0.205	0.3896	0.6647	0.9221	0.9987	1

Table 22 - Probability of k consecutive heads or more in N throws

Note that the values for P are corrected to 4 significant figures.

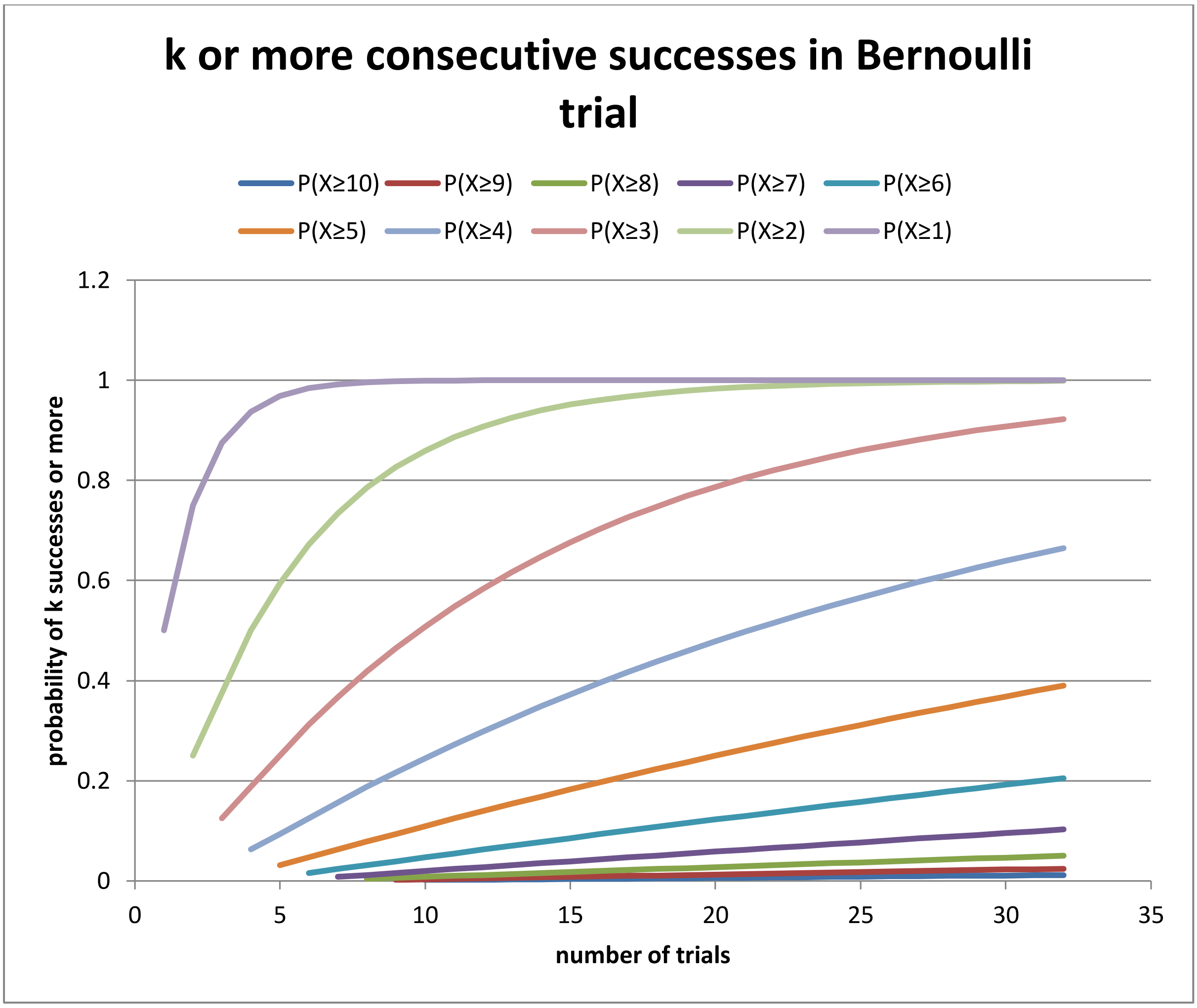


Figure 58 - Plots of probability of getting k or more consecutive successes within N throws

19. Appendix - Papers (1)

This paper was published in the 2014 Proceeding of International Conference on Advances in Computing, Electronics and Electrical Technology CEET:

Gan, H., Frank, R., Amirabdollahian, F., Sharp, R., & Rainer, A. (2014b). Development of the maximum-streak measure for evaluating the suitability of non-handheld devices in cognitive tests of Physically and Neurologically Impaired(PNI) children. *International Journal of Advances in Computer Science & Its Applications, 4*(4), 130 - 136.

Development of the maximum-streak measure for evaluating the suitability of non-handheld devices in cognitive tests of Physically and Neurologically Impaired(PNI) children

Hock C. Gan, Ray J. Frank, Farshid Amirabdollahian, Rob Sharp, Austen W. Rainer

***Abstract*—This study examines an objective measure for discriminating between different degrees of performance in a child's use of multiple non-hand-held devices when undertaking a simple competence test: the maximum streak of successfully, making a binary (YES / NO) response in the test. The paper reports encouraging results for the maximum consecutive success (streak) measure in its ability to inform on device suitability for a particular child. Our results show that different devices influence success streak, while performance of children vary when using different devices, based on severity of impairment. While this is in line with our main hypothesis, that a suitable match between device and impairment is necessary, our results provide a first candidate measure that can be potentially useful to personalizing a device to a particular user and his/her capabilities.**

***Keywords*— disabled children, non-hand held device, motor-skill test, cognitive test, streak**

I. Introduction

Higher life expectancies have seen a rise in average population ages. A huge potential exists in the market for new bio-modal devices that provide interaction with an aging population with a need for assisted living. This paper provides a re-think of the user-device-application triad in human-computer interaction with focus on inputs that are distorted by involuntary interactions that may be muscular or otherwise (e.g. unstable bio-modal systems) in origin.

Physically and Neurologically Impaired (PNI) children have special needs due to problems which result from brain injury. The effects of the brain impairment result in physical impairment and resulting in the need to address both problems. A large proportion of such children suffer from Cerebral Palsy (CP) [1], a condition affecting a combination of movement, speech, cognitive processing, and visual perception, as well as other difficulties. The involuntary actuation and inhibition of movements in CP children mean that such children have great difficulty controlling the traditional devices used in tests, the typical traditional device being a computer mouse.

Hock C. Gan, Ray J. Frank, Farshid Amirabdollahian, Austen W. Rainer
University of Hertfordshire
United Kingdom
h.c.gan@herts.ac.uk, r.j.frank@herts.ac.uk, f.amirabdollahian2@herts.ac.uk, a.w.rainer@herts.ac.uk

Rob Sharp
Hertfordshire County Council
United Kingdom
rob.sharp@hertfordshire.gov.uk

New bio-modal devices that do not depend on using the traditional motor skills are becoming available, and at a relatively low cost. These bio-modal devices provide new opportunities for CP children to undertake a test, but may also require that the children develop new skills in order to control those new devices. The availability of these bio-modal devices raises questions: Which bio-modal device or devices are suitable for which 'types' of impairment? What is the impact of each of the three components of the child-device-test triad on the outcome of the test? In order to begin to explore these potential answers to such questions, there is a need to be able to measure the degree of effective use of a device, by a child, when undertaking a test.

In this paper, we report the development of a measure, maximum-streak, to assess the degree to which a child with neurological impairment can use a bio-modal device when undertaking a motor skills test. Streaks are consecutive binary event sequences of the same type. Success streaks refer to a consecutive sequence of success events, of two or more. Maximum-streak is the highest streak size attained in one run of a test of several trials, each requiring a binary (YES/NO) response. The work reported here examines a number of different bio-modal devices, with a small sample of children and one motor skills test requiring very simple cognitive ability.

The following devices were chosen for this study:

1. A neuro-headset[2] based on EEG used to detect facial artefacts [3, 4]
2. An eye-tracker [5, 6]
3. A face/head-tracker[7, 8]
4. A hybrid device consisting of both EEG and eye-tracking(implemented for evaluation)
5. A hybrid device consisting of both EEG and head-tracking(implemented for evaluation)

This study also considers *Signing* as a method of input. Signing provides a contrast to bio-modal inputs and the typical physical inputs. For our research, Signing refers to a child who communicates using gestures to an interpreter. The gesture acts as the child's response to a test, and this response is then entered as a mouse input by the interpreter. Signing is a valuable contrasting 'input' because a child that successfully completes a test using the Signing 'device' most likely has the cognitive ability for being tested. Where that child then has difficulties completing the test successfully with a bio-modal device helps to expose the positive or negative impact of that device.

A previous project with a similar perspective was the "i-match" project which focused on comparing input devices to identify a best fit to individual's abilities [9, 10]. The general principles that were extracted from that project were the motor skills assessment usually based on the use of Fitts' test [11] and a multitude of benchmarks. A motor skills assessment was made in the case of the "i-match" project to evaluate the effectiveness of a number of set configurations for a single device for a sample population. In the case of this study, general configurations were found and the motor skills assessment made to determine the suitability of a number of set devices for a single individual.

Motor skill tests do not, in general, need to consider the failure to make a correct response. Devices used for such tests are often stable and tests are simple reaction tests not expected to have a high cognitive load. Participants chosen have enough developmental maturity to carry out the tests perfectly. The parameter of interest moves on to looking at performance times or accuracy, speed or error, and statistical results are generated by looking at a large enough sample [10, 12].

This study, when considering the problems of the severely impaired, uses success streaks instead of a simple number of successes and performance timing. The unintended inputs caused by the involuntary movements, cognitive impairment, fatigue of the child, together with mechanistic imperfections of the system, are treated as noise. As noted above, streaks are consecutive binary event sequences of the same type, and success streaks refer to a sequence of successes of two or more. Streaks have previously been the subject of studies in gambling behaviour [13, 14]. Streaks have been used to establish success in cognitive tests like the Wisconsin Card Sorting Test (WCST) [15] where reinforcement or learning is concluded after a certain number of consecutive successes and various other cognitive significances are drawn from other thresholds within the test. This study draws inspiration from those procedures to benchmark significant inputs from the motor skills test. So in the motor-skills test, a success threshold would exist as well as a threshold to separate intention from noise. It is notable that the term intention is used here to mean that the action is not the result of a reflex action but one arising from a conscious decision[16].

One advantage of a parameter similar to the number of successes used in standard tests is that it is simple measure that provides a graded scale of success. It is helpful in the case of this study, to also start by considering a parameter which is able to generally provide a measure of success in order to assess suitability of different bio-modal devices. This study uses the maximum streak size in that it represents the best attempt in a block of trials. The best attempt and other significant attempts would be the result of [17] intention and control largely affected by practice[18], talent[19] and noise. There is an indirect contribution to this noise caused by the motor-skills test which takes in a binary (YES/NO) response. If there are binary streaks of considerable length, the outcome would be a reflection of the algorithm that generates the sequence of stimuli presentation. This sequence is in general random but the algorithm can be improved to keep the streak noise low in those cases by ensuring that there are no more than 2 stimuli in sequence that require a YES or a NO response. These conceptual ideas and enhancements from WCST were implemented in a motor-skills test (called COMPTEST for Competence Test) that was used in this study.

II. Experimental method

A. *Participants*

Seven PNI children were tested, five have various forms of CP, one has methotrexate leucoencephalopathy and one has septo-optic dysplasia with autistic spectrum disorder. Approval from both school and parents were sought under the University of Hertfordshire's ethics protocol aCOM/PG/UH/00006. Fictitious names have been used for all the children in all publications.

The ages of the participants were rationalized using the British Picture Vocabulary Scale III (BPVS III) [20] to provide the developmental age as shown in Table I. As a result of impairments of the children, some BPVS results are best effort results as the case with Apollo and Thor. The table indicates three participants who are severely impaired in that they are wheelchair bound, have almost no speech, and have involuntary muscular problems and weak muscular control. The remaining four suffer some impairment to a lesser degree and are not wheelchair bound.

Geronimo was picked to provide an example of a person with CP and with mature developmental capability. The others were picked as examples of children with varying impairments between the developmental ages of 4 to 7 for compatibility with another study.

TABLE I. AGE EQUIVALENTS OF PARTICIPANTS

Name	***Age***	***Severe impairment***	***Developmental age (years: months)***
Apollo	14	Yes	04:10
Bacchus	12	No	04:07
Baldr	15	No	< 04:00
Geronimo	13	No	11:03
Lavender	12	Yes	< 04:00
Nimrod	13	No	04:05
Thor	12	Yes	07:03

B. *Procedure*

The participants are tested in a room (located in a school) equipped with a laptop, separate screen monitor and hand-held and non-hand held devices for interaction with the software tests running on the laptop. The screen monitor is arranged side to side with the laptop so that the participants with a view to the screen monitor are seated beside the researcher who has a view to the laptop. The eye-tracker is mounted below the screen of the monitor using magnetic mounts. The head-tracker uses a remote web camera enabling tracking from the monitor.

A set of stimulus is produced to which the participant must provide a positive or negative response. The positive response is an active response which involves the actuation of a device. The negative response is passive requiring no action. Success terminates the trials for each device determined by 20 consecutive correct responses. Otherwise, the trials terminate after a block of 32 trials. Sessions of several blocks of trials involving different devices and tests are conducted once a day and all tests are run within 2 sessions of an hour each. During the tests observations were made by the researcher regarding test response behaviour and notes compiled after testing. Informal feedback was sought regarding fatigue, comfort levels and preferences after tests from carers and participants. Devices were run in order of increasing complexity as observed in a pilot testing. In general, this would involve running tests with the mouse/switch first, followed by single mode devices and then hybrid multi-mode devices. Signing is used when it was determined that there were no other devices that seem to be reliable enough to be used.

When the EEG headset is used, participants provide a positive response by gritting their teeth for a period identified as the "bite-time". When an eye-tracker is used, participants have to move a mouse cursor with their eyes to an image of a switch and "dwell" the cursor over the virtual switch for a period identified as the "dwell-time". When a head-tracker is used, participants have to move a mouse cursor using movement of their head to the virtual switch for the dwell-time. The hybrid EEG and eye-track device uses the eyes to move the mouse cursor but instead of a dwell-time, participants have to grit their teeth for a bite-time to indicate a positive response. In a similar way, the EEG and head-tracker used the movement of the head to move the mouse cursor and the teeth-grit to actuate the virtual switch. When Signing is used, the participant is asked "Is this a dog?" to which they would use gestures that are specific to them to respond with YES or NO to a carer who will then interpret their responses.

The stimulus for the COMPTEST consists of either an image of a cuddly dog or scraggy cat. Participants have to provide a positive response when they see an image of the dog and negative response when they see the image of the cat. Participants are tested on all devices on the COMPTEST with each device forming a block with a maximum of 32 trials. Feedback is provided in the form of the sound of a bell tinkle when the virtual switch is actuated but not whether the response is correct or incorrect. Participants are familiarized with a few runs of the test to ensure they understand the test and are able to engage the devices.

C. *Design*

Fig.1 shows the test design. The experiment is a 6 x 32 within-subjects factorial design for the maximum number of trials.

- Device {(Signing), mouse/switch, eye-track, head-track, EEG, EEG and eye-track, EEG and head-track}. Signing is only used when no other devices can be used and does not increase the maximum count of devices for trials.
- Block {1 to a range between 20 and 32}
- Tests {COMPTEST}

There are 7 participants giving a maximum total number of 7 x (6 x 32) trials.

D. *Data capture*

COMPTEST results are represented as a 32-bit field, each bit representing an OK/NOK (not OK) outcome for a particular trial. The field can be represented by a list of success and failure streaks. Failure streaks are suffixed with x. For example, for a list of 17 successes followed by 3 failures, 10 successes and 2 failures, the field is represented as {17, 3x, 10, 2x}. The consecutive successes and failure are referred to as success and failure streaks respectively. The entire list which captures an entire block of trials is referred to as an outcome event sequence. The maximum number of consecutive successes in the example list is 17. Multiple maxima may exist in general.

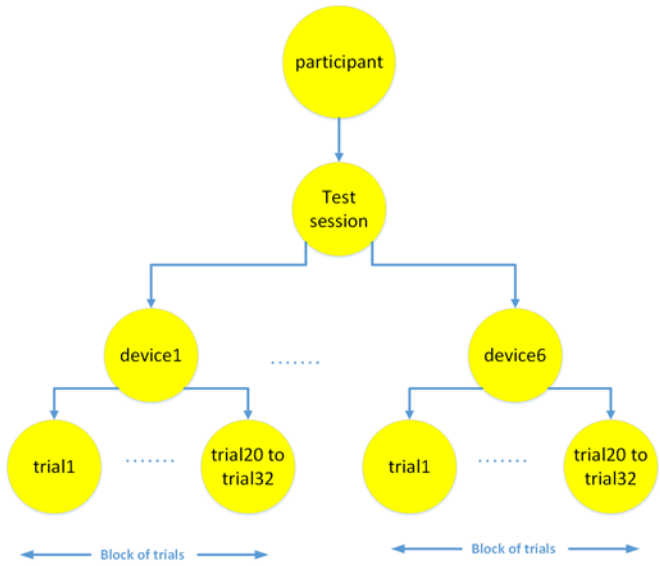


Figure 1. Test design

III. Results and discussions

Fig.2 provides an overview of the maximum streak size data for success streaks. Each maximum streak size represents a block of trials that a PNI child has had with a specific device. The count therefore shows the number of blocks that achieved a specific maximum.

We see that performances are clustered into three groups of achievements (2-7, 15-17, 20). From the observation notes of the tests, the higher maximum streaks were achieved by participants who had no problems with the test or device and the spread of maximum streaks was due to minor distractions. If they did not achieve the target maximum streak, they would most likely do so in a subsequent test. Looking at those who have achieved the low maximum streaks (of 2 to 7) we note that the participants have very severe impairments or impacting cognitive problems. The lowest maximum streaks of 2 and 3 were from tests that were aborted because the child had long periods of no response and showed an inability to use the novel device due to spasms or not understanding how to control the device.

The lowest maximum streaks hint that there is a level of low streak sizes which is noise naturally generated by the system from which no useful conclusion can be derived. At this point it is helpful to examine the distribution of streaks from the 2 to 7 maximum streak size group to explore the noise threshold.

Fig.3 shows for the low maximum streak size cluster, the number of occurrences of a particular streak size (taken for all the tests). Streak sizes of 1 and 2 are very common because the test program generates that noise. Streak sizes between 3 and 7 are much less likely.

Fig.4 provides a confidence level for the decision on choosing a streak size of less than 5 to represent the noise level provides about a 95% confidence level for the signals deemed as significant. At this point the assumption is made that all streak sizes that are 5 or above indicate intention resulting from control. The term intention is used to mean that the action is not the result of a reflex action but arising from a conscious decision[16].

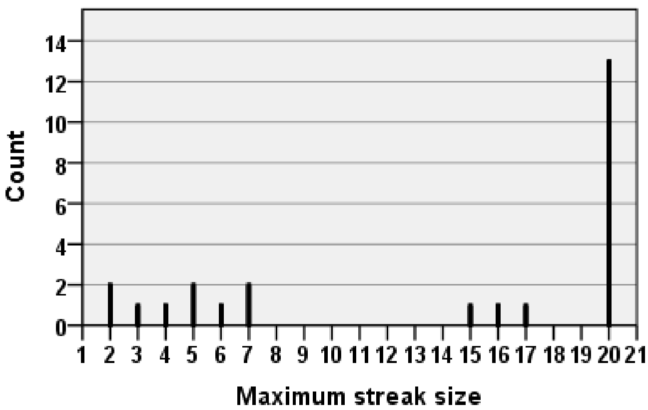


Figure 2. Maximum success streak size distribution

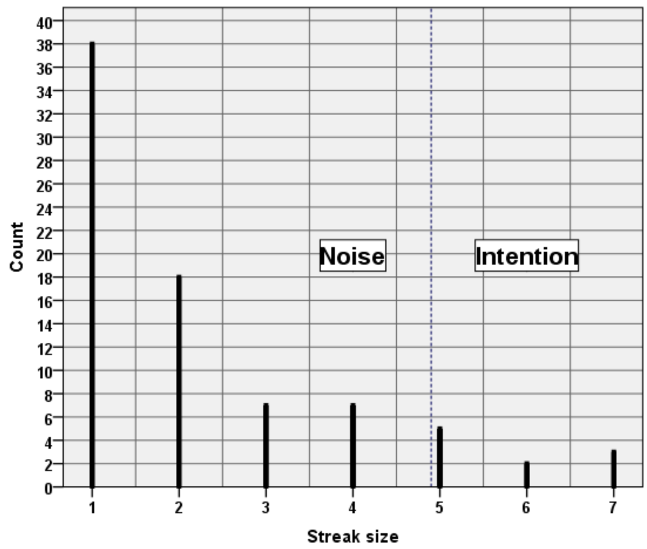


Figure 3. Success streak size distribution of severely impaired

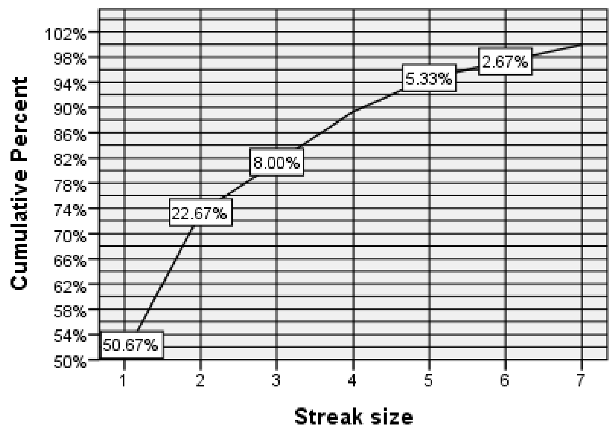


Figure 4. Success streak size cumulative distribution of severely impaired

Fig. 5 emphasizes the child's ability with different physical devices using maximum streak size as a measure. The children are sorted to show a general trend of ability starting from the left. As the comparison is done for ability with physical devices, we can examine the trend discounting Signing (blue). Signing was only carried out in cases where some validation was required due to the low scores with their best physical device to separate cognitive inadequacy from other problems (for example device problems). Signing is assumed to be more reliable as the child has had a longer exposure to its use with a carer who is attuned to subtle signals from the child.

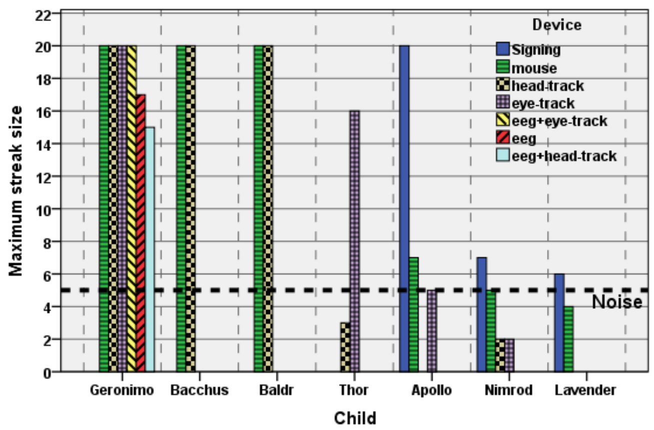


Figure 5. PNI child device ability. For physical device evaluation, Signing should be discounted.

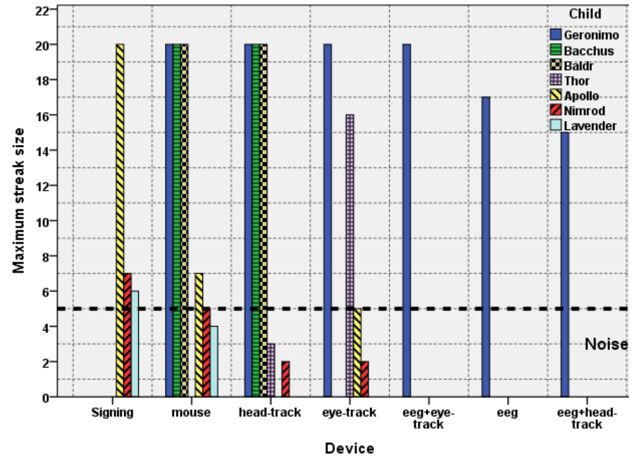


Figure 6. Device competency. For physical device evaluation, Signing should be discounted.

The maximum attainable score is the test target streak size of 20 and it decreases to 6. If we now look at just the highest maximum streak size attained by each child in general, the streak size decreases with the ability, with the last four children in this figure presenting children with most severe neurological impairments clustered to the right hand side (with the exception of Nimrod). We note that Thor was offered good affordances from his only physical device (thus Signing was not offered) but for the last three children, maximum streak size was only reached using Signing (with the physical devices offering poor affordances). Apollo had a good result with Signing but both Nimrod and Lavender also had poor results from Signing. The physical devices appear to be helpful for Thor, unhelpful for Apollo and correlate with the representations made by Nimrod and Lavender. The trend can also be examined using the number of devices each child is capable of using. This number decreases in the plot from left to right. This starts off with all devices being usable (Geronimo) to only one device being usable (Thor-Nimrod) to none (Lavender), discounting levels that are regarded as noise.

Fig. 6 is another viewpoint of the same data which clusters the children for each device. This figure makes it clear which devices tend to be the more successful ones. The figure was again arranged with the best results starting on the left, indicating which device tend to be more readily used by the children in the short space of time that the tests were conducted. To look for a trend for physical devices, we can ignore Signing. Again if we first consider the highest maximum streak size achieved for each device, as we go from the mouse to EEG and head-track, a decrease in streak size is found. However, this trend is mostly driven by one individual (Geronimo). If we now consider the number of children for each device (again ignoring Signing), we find a progression that ranks the mouse, the head-tracker and the eye-tracker in that respective order.

The rest of the devices have only one user.

Fig. 7 provides a comparison of a typical measure used in testing; the score (blue) with the maximum-streak measure (red). The score measures the number of successful trials in a block of tests and in the figure is expressed as a percentage of the number of trials executed in a block. The figure shows the results using the two measures and is clustered around children using various devices ordered by the child's ability.

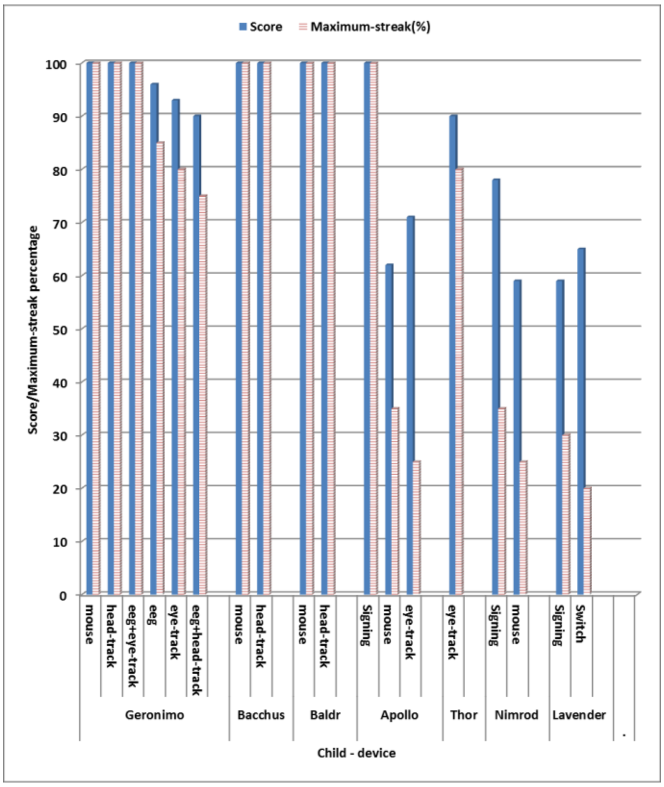


Figure 7. Comparison of maximum-streak measure with success score sorted by child and maximum-streak

The children with the best outcomes start on the left followed by the device outcomes using the maximum-streak measure. Among the children, Geronimo has the best outcomes followed by Bacchus, Baldr, Apollo, Thor, Nimrod and Lavender. The results differ for scores and maximum-streak with scores being always higher in general apart from when they both are 100%. The relative trends of both measures are not always aligned with the scores sometimes increasing when the maximum-streak shows an increase.

The differences come from the design of the measures. Scores are an averaging measure of the successes achieved spread among the total number of tries. Maximum-streak measure the best success achieved as a fraction of the maximum (20) that could be achieved. Scores do not identify events that are significant. For instance, if a child had 16 successive correct responses and 16 incorrect responses, the score would still be 50%. However, a maximum-streak measure would have indicated a significant achievement of 80 because a maximum-streak measure specifically distinguishes itself from average occurrences of which noise will be a strong component and acts more specifically as a measure of achievement. Scores are suited for cases where there are large numbers of re-attempts interspersed with occasional failures. Maximum-streaks show cases where singular strong attempts are made but there is possibly insufficient energy or opportunity to provide other significant outcomes.

The use of streak measures allows possible noise that arises from involuntary muscular outcomes to be quantified and isolated. Maximum-streaks are used as a first candidate measure as it is sensitive to users that suffer easily from fatigue due to the extra effort required to both mitigate impairment and drive the input device. The example above where a child had 16 correct and incorrect responses highlight two unusual events; high correct and incorrect responses and show that often one measure is not enough to reveal the complete story. Other supporting measures are required and are applicable candidates for future research.

IV. Conclusion

The maximum streak measure seems to be able to discriminate between at least three clusters of performance on a test: children that can relatively easily achieve the 'best' intended test outcome; children that are able to achieve close to the 'best' intended test outcome; and children that struggle to achieve any outcomes (this last group appearing to be those that are most severely impaired or have impacting cognitive problems). This paper speculates on the causes for these three clusters of performance; however these explanations remain a direction for further research. It is the final group that is the most relevant as we are seeking to develop devices that can support the assessment of just this group of PNI children. We have established a threshold value of 5, above which the maximum-streak measure appears to be reliable. Such a threshold value increases our confidence in the measure. The measure seems to be able to discriminate between children's performance on different devices. This is important because the larger research project is investigating a range of different non-hand-held devices.

In terms of devices, we have some that provide adequate representation of intention, some which were shown to be inadequate but no other option was available (Apollo using mouse) and some which may be adequate but other problems retard performance (Nimrod using mouse and Lavender using switch). Signing appears to be a good reference "device" and future work has the aim of matching or out performing it.

The limitations for this study are the small sample size that was used and the development of explanations. A larger sample would represent the opportunity for further research. The small sample size meant that it was difficult to gather sufficient information on the different combination of tests, devices and "types" of PNI. The ideal is that a large number of PNI children representative of the wide variation in their impairments (as a "natural control") are used. The wide variations and complexity of their impairments require a wider set of test variations to develop and establish explanations.

The future directions for this research include a number of directions. We intend to look at other measures as well as patterns of outcomes and not just maximum-streak. We will see if a combination of different measures and patterns can produce categorizations of test behaviour. The categorizations will be used to support the attempts to match devices to impairments.

A number of countries are projecting a rise in the mature population due to the average increase in life-expectancy. This mature population will form a huge market that needs to be addressed[21]. This paper proposes a measure for human-computer interaction with a user that has responses that will be distorted by involuntary muscular action by ignoring the responses that is possibly contaminated. By using this paradigm, the test application can also generate requirements for responses that are optimised to reduce the probability of specific noise levels to fit the receiving evaluation that only considers responses above a specific level. This indicates the need for expanding the user-device-application triad design to respond to patterns of impairment in order to find a best match device that can augment an individual's capabilities be it motor or cognitive.

Acknowledgment

We would like to thank Mr. I. Glasscock for his support. We are grateful to the schools that agreed to participate in this study especially the staff of the special needs schools for their time and effort, and to the participants and their care-givers/families in assisting us with this study.

References

[1] C. Cans, H. Dolk, M. J. Platt, A. Colver, A. Prasausk1Ene, and I. K. RÄGeloh-Mann, "Recommendations from the SCPE collaborative group for defining and classifying cerebral palsy," *Developmental Medicine & Child Neurology,* vol. 49, pp. 35-38, 2007.

[2] Emotiv. (2013, June). *Emotiv EEG Features. http://www.emotiv.com/eeg/features.php*.

[3] C. E. Reyes, J. L. C. Rugayan, C. Jason, G. Rullan, C. M. Oppus, and G. L. Tangonan, "A study on ocular and facial muscle artifacts in EEG si

signals for BCI applications," *TENCON 2012-2012 IEEE Region 10 Conference,* pp. 1-6, 2012 2012.
[4] A. B. Barreto, S. D. Scargle, and M. Adjouadi, "A practical EMG-based human-computer interface for users with motor disabilities," *Journal of Rehabilitation Research & Development,* vol. 37, pp. 53-64, 2000.
[5] Tobii Technology. (2013, June). *Portable lab - Tobii X2-30 Eye Tracker. http://www.tobii.com/fr/eye-tracking-research/global/products/hardware/tobii-x2-30-eye-tracker/.*
[6] C. Ware and H. H. Mikaelian, "An evaluation of an eye tracker as a device for computer input2," *SIGCHI Bull.,* vol. 18, pp. 183-188, 1986.
[7] Boston College. (2013, June). *CameraMouse. http://www.cameramouse.org/.*
[8] M. Betke, J. Gips, and P. Fleming, "The camera mouse: visual tracking of body features to provide computer access for people with severe disabilities," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on,* vol. 10, pp. 1-10, 2002.
[9] F. Amirabdollahian, M. Munih, F. Kouris, E. Laudanna, E. Stokes, and G. Johnson, "The I-match project: A VR based system to allow matching of an optimum interface to a user of assistive technology," *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on,* pp. 526-529, 2005 2005.
[10] F. Amirabdollahian, G. T. Gomes, and G. R. Johnson, "The peg-in-hole: a VR-based haptic assessment for quantifying upper limb performance and skills," *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on,* pp. 422-425, 2005 2005.
[11] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *Journal of experimental psychology,* vol. 47, pp. 381-391, 1954.
[12] D. W. Man and M. S. Wong, "Evaluation of computer-access solutions for students with quadriplegic athetoid cerebral palsy," *Am J Occup Ther,* vol. 61, pp. 355-64, May-Jun 2007.
[13] E. M. Altmann and B. D. Burns, "Streak biases in decision making: Data and a memory model," *Cognitive Systems Research,* vol. 6, pp. 5-16, 2005.
[14] P. Ayton and I. Fischer, "The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness?," *Memory & cognition,* vol. 32, pp. 1369-1378, 2004.
[15] D. A. Grant and E. Berg, "A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem," *Journal of experimental psychology,* vol. 38, p. 404, 1948.
[16] B. Libet, E. W. Wright Jr, B. Feinstein, and D. K. Pearl, "Subjective referral of the timing for a conscious sensory experience," in *Neurophysiology of Consciousness*, ed: Springer, 1993, pp. 164-195.
[17] A. T. Oskarsson, L. Van Boven, G. H. McClelland, and R. Hastie, "What's next? Judging sequences of binary events," *Psychological bulletin,* vol. 135, p. 262, 2009.
[18] K. A. Ericsson, R. T. Krampe, and C. Tesch-Römer, "The role of deliberate practice in the acquisition of expert performance," *Psychological review,* vol. 100, p. 363, 1993.
[19] G. Marcus, *The birth of the mind: How a tiny number of genes creates the complexities of human thought*: Basic Books, 2008.
[20] L. M. Dunn, *The British picture vocabulary scale*: GL Assessment Limited, 2009.
[21] R. Chapin and D. Dobbs-Kepper, "Aging in Place in Assisted Living Philosophy Versus Policy," *The Gerontologist,* vol. 41, pp. 43-50, 2001.

About Author (s):

Hock Gan: "Signing" points the way to the best device fit for any individual; the process being forged by evolution.

An interesting aspect of this work is that it enables almost all PNI children to demonstrate their motor/cognitive skills over a wider range of motor and cognitive development compared to simple pass/fail tests giving educational psychologists a far better assessment tool.

Farshid Amirabdollahian: Non hand-held devices are thought to be suitable as replacement interactive tools for assessing motor and cognitive capabilities. By devising simple motor or cognitive tests, we show that devices and tests should be carefully considered towards minimising additional motor and cognitive requirements.

Rob Sharp: Children who undergo learning tests express a range of emotions including surprise, frustration and satisfaction. I want a device that will tell me that.

Austen Rainer: It's interesting to observe that even neuro-typical, "normal" children can struggle to use some of the devices when undertaking the tests. My assumption was that all of these children would be capable of successfully passing the tests using the different devices.

20. Appendix - Papers (2)

This paper was published in "Human System Interactions (HSI), 2014 7th International Conference on":

Gan, H., Frank, R., Amirabdollahian, F., Sharp, R., & Rainer, A. (2014c). Use of re-attempts measure for evaluating device test results of children with neurological impairments. *Human System Interactions (HSI), 2014 7th International Conference on*, 206 - 211. doi: 10.1109/HSI.2014.6860476

Use of re-attempts measure for evaluating device test results of children with neurological impairments*

Hock C. Gan, Ray J. Frank, Farshid Amirabdollahian, Austen W. Rainer
School of Computer Science
University of Hertfordshire
Hatfield, UK
Emails: h.c.gan@herts.ac.uk, r.j.frank@herts.ac.uk, f.amirabdollahian2@herts.ac.uk, a.w.rainer@herts.ac.uk

Rob Sharp
Hertfordshire County Council
Welwyn Garden City, UK
Email: rob.sharp@hertfordshire.gov.uk

***Abstract*—Severely impaired children with Physical and Neurological Impairments (PNI) often have erratic test responses because of impairments. Very often even the binary (YES/NO) intention of a PNI child cannot be determined because responses are made at the wrong time and conflicting signals are sent. We propose that it is possible to determine intention using significant streaks of successful responses found in the noisy responses. We can use two measures among others to determine intention; the maximum streak size attained in a test run with a device, and the sum of the significant streaks in the test run. The maximum streak size measures consecutive successes and the sum of streaks gives an indication measure of re-attempts. This work is part of a larger study to increase accessibility of PNI children to a cognitive test through the use of new non-hand held devices that interact with a computer. Using the proposed two measures, we are able to compare more closely the performance of the less capable and the more capable PNI children. The results show that the children who are more capable re-attempt in that test when they fail to achieve a target. Conversely the less able children divide into two groups: those that do re-attempt the target and those that do not re-attempt the target dependent however on the device being used. These results provide two measures that are potentially useful for determining the intention of the child undertaking the cognitive test.**

Keywords—disabled children; streaks; cerebral palsy; motor skills test; non-hand held devices;

I. INTRODUCTION

Physically and Neurologically Impaired (PNI) children commonly suffer from disabilities as a result of brain trauma. A large proportion of such children suffer from Cerebral Palsy (CP) [1], an umbrella term for a range of disabilities that include a combination of movement, speech, cognitive processing, and visual perception. Education of children with disabilities that affect communication such as motor and speech is difficult. Education, which is a two-way process of disseminating and assessing knowledge, is hampered if the communication pathways are restricted or degraded. Received knowledge is distorted and assessment is inaccurate. Assessment in this case is challenging and perhaps even more necessary. Current assessments typically involve tests that are digital and require interaction with an input device, usually a mouse. Unfortunately, the involuntary actuation and inhibition of movements of a number of PNI children mean that such children have great difficulty controlling devices used in tests.

*Research supported by the Engineering and Physical Sciences Research Council (EPSRC UK) and BioDigitalHealth under iCASE 09001842.

New devices are available that involve other affordances that potentially requires skills that the PNI child may be more capable of. This resulted in the following devices being chosen for this study:

1. A neuro-headset[2] based on EEG used to detect facial artefacts [3, 4]
2. An eye-tracker [5, 6]
3. A face/head-tracker[7, 8]
4. A hybrid device consisting of both EEG and eye-tracking(implemented for evaluation)
5. A hybrid device consisting of both EEG and head-tracking(implemented for evaluation)

This study also considers Signing as a method of input, and Signing provides a contrast to bio-modal inputs and the typical physical inputs. Signing in this case refers to a child who communicates using gestures to an interpreter. The gesture acts as the child's response to a test, and this response is then entered as a mouse input via the interpreter.

This paper identifies some related work that was used to build this study upon in Section II, followed by a description of how the experimental study was set up in Section III, a description of the results obtained from the study in Section IV, a discussion of the significance of the results in Section V and concludes the paper with the significant points, critique and an idea of future possible directions.

II. RELATED WORK

The use of new devices necessitates a means of evaluation to rank and match the devices to the user. Similar projects in the past such as the "i-match" project had those particular aims [9, 10]. A motor skills test was set up based on Fitts' test [11] to evaluate the effectiveness of a number of set configurations for a single device for a sample population.

A difference in the approach used by this study is the use of streaks or consecutive successes instead of a simple count of total success for assessment of PNI children. Streaks have been the subject of study in several areas. Streaks have been the subject for mathematical publications[12], studies on gambling behaviour [13, 14] and used to establish success in cognitive tests like the Wisconsin Card Sorting Test (WCST) [15] where reinforcement or learning is concluded after a certain number of consecutive successes. In this study, inspiration is drawn from WCST in the use of the maximum streak size as a first candidate measure for intention. Intention is a reference made for an action that is not the result of a reflex action but one arising from a conscious decision[16]. The intention response was separated from the response cases where no conclusion could be drawn through the establishment of a threshold level. The threshold level is a streak size by which streaks of a smaller size would be regarded as noise and a larger size as significant streaks indicating intention. A different level (a target streak size) was set to indicate achievement of general required performance for a future cognitive test. Using the maximum-streak measure we obtained relative ranking of performances primarily of the capable children with a device and the less capable ones. The capable ones had results near the target streak size. The less capable ones tended to have results that clustered near the noise level such that some were larger in size (above the noise level) and some lower. The ones below the noise level were regarded to be disengaged from either the device or the test. The value of the noise threshold was determined empirically by a pilot study whereby streak sizes below 5 were regarded as noise. We will investigate the threshold further and results will be considered for future publication. The maximum-streak size only provides a single instance of achievement and that was seen to be an attribute that could be improved on.

There are several reasons for a measure that is able to indicate multiple instances of achievement (multiple streaks). Frequently the responses made by the severely impaired people contain errors because of the inability to control muscular movement that causes undesired activity or lack of activity. These unwanted movements manifest as noise in a test where consecutive success are broken up by uncontrolled impulses produces errors. An ideal test situation is one where these uncontrolled impulses were absent and one projection that can be made is to ignore the intermittent errors and consider the outcome as a sequence of consecutive successes without the errors. A measure of the multiple streaks can be seen as a projection of the true measure of achievement for a PNI child.

Another reason for a measure of multiple streaks is that the multiple streaks provide a greater confidence of intention than a single streak. The chance of getting multiple streaks of equivalent size in succession is less than just getting a single one. Moreover, multiple streaks show that the child has the stamina to use the device in the test. The other way of looking at it is that the physical and cognitive load imposed by both test and device are small enough to encourage re-attempts. Single attempts can then be possibly due to the lack of stamina implying fatigue.

Multiple streaks provide another means for the device test ranking. Although impairments disrupt the opportunity to reach the target streak size, it does not prevent the capability to reach a significant size several times. The inability to reach the target size may be the result of a combination of device mismatch and impairment. A child who displayed multiple streaks would be ranked higher with another who achieved the same streak size but with a single streak.

TABLE I. AGE EQUIVALENTS OF PARTICIPANTS

Name	*Age*	*Severe impairment*	*Developmental age (years: months)*
Apollo	14	Yes	04:10
Bacchus	12	No	04:07
Baldr	15	No	< 04:00
Geronimo	13	No	11:03
Lavender	12	Yes	< 04:00
Nimrod	13	No	04:05
Thor	12	Yes	07:03

The measure that was chosen to represent the multiple streaks (streak-sum) was one that measured the sum of significant (above noise level) streaks in a test run with a device. The streak-sum is a summary of both the magnitude and the frequency of the streaks resulting from the test. When compared with the maximum streak, the streak-sum differentiates between outcomes that were the result of a single streak and outcomes from the case where there were multiple streaks. The importance of being able to make the distinction is that a single attempt, unless it attains the target streak size, can be seen as due to a combination of device imperfections, unmatching device, fatigue and inattention.

III. EXPERIMENTAL METHOD

This section describes the how the experiment was constructed and run and starts with the choice of participants (sub-section A), followed by the set-up for an experiment (sub-section B), followed by an overview of the structure of the experimental design (sub-section C) and finally, a description of the format of data captured for analysis (sub-section D).

A. Participants

Seven PNI children were tested, five have various forms of CP, one has methotrexate leucoencephalopathy and one has septo-optic dysplasia with autistic spectrum disorder. Approval from both school and parents were sought under the University of Hertfordshire’s ethics protocol aCOM/PG/UH/00006. Fictitious names have been used for all the children in all publications.

The ages of the participants were rationalized using the British Picture Vocabulary Scale III (BPVS III) [17] to provide the developmental age as shown in TABLE 1. As a result of impairments of the children, some BPVS results are best effort results as the case with Apollo and Thor. The table indicates three participants who are severely impaired in that they wheelchair bound, have almost no speech, and have involuntary muscular problems and weak muscular control. The remaining four suffer some impairment to a lesser degree and are not wheelchair bound.

Geronimo was picked to provide an example of a person with CP and with mature developmental capability. The others were picked as examples of children with varying impairments between the developmental ages of 4 to 7 for compatibility with another study.

B. *Procedure*

The participants are tested in a room (located in a school) equipped with a laptop, separate screen monitor and hand-held and non-hand held devices for interaction with the software tests running on the laptop. The screen monitor is arranged side to side with the laptop so that the participants with a view to the screen monitor are seated beside the researcher who has a view to the laptop. The eye-tracker is mounted below the screen of the monitor using magnetic mounts. The head-tracker uses a remote web camera enabling tracking from the monitor.

A set of stimulus is produced to which the participant must provide a positive or negative response. The positive response is an active response which involves the actuation of a device. The negative response is passive requiring no action. Success terminates the trials for each device determined by 20 consecutive correct responses. Otherwise, the trials terminate after a block of 32 trials. The stimulus for the COMPTEST consists of either an image of a cuddly dog or scraggy cat. Participants have to provide a positive response when they see an image of the dog and negative response when they see the image of the cat. Participants are tested on all devices on the COMPTEST with each device forming a block with a maximum of 32 trials. Feedback is provided in the form of the sound of a bell tinkle when the virtual switch is actuated but not whether the response is correct or incorrect. Participants are familiarized with a few runs of the test to ensure they understand the test and are able to engage the devices.

Sessions of several blocks of trials involving different devices and tests are conducted once a day and all tests are run within 2 sessions of an hour each. During the tests observations were made by the researcher regarding test response behaviour and notes compiled after testing. Informal feedback was sought regarding fatigue, comfort levels and preferences after tests from carers and participants. Devices were run in order of increasing complexity as observed in a pilot testing. In general, this would involve running tests with the mouse/switch first, followed by single mode devices and then hybrid multi-mode devices. Signing is used when it was determined that there were no other devices that seem to be reliable enough to be used.

When the EEG headset is used, participants provide a positive response by gritting their teeth for a period identified as the "bite-time". When an eye-tracker is used, participants have to move a mouse cursor with their eyes to an image of a switch and "dwell" the cursor over the virtual switch for a period identified as the "dwell-time". When a head-tracker is used, participants have to move a mouse cursor using movement of their head to the virtual switch for the dwell-time. The hybrid EEG and eye-track device uses the eyes to move the mouse cursor but instead of a dwell-time, participants have to grit their teeth for a bite-time to indicate a positive response. In a similar way, the EEG and head-tracker used the movement of the head to move the mouse cursor and the teeth-grit to actuate the virtual switch. When Signing is used, the participant is asked "Is this a dog?" to which they would use gestures that are specific to them to respond with YES or NO to a carer who will then interpret their responses.

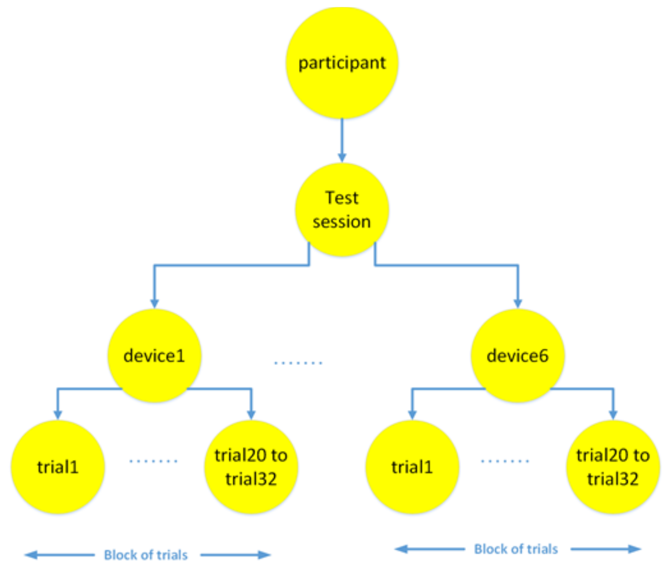


Fig. 1. Test design

C. *Design*

Fig.1 shows the test design. The experiment is a 6 x 32 within-subjects factorial design for the maximum number of trials.

- Device {(Signing), mouse/switch, eye-track, head-track, EEG, EEG and eye-track, EEG and head-track}. Signing is only used when no other devices can be used and does not increase the maximum count of devices for trials.
- Block {1 to a range between 20 and 32}
- Tests {COMPTEST}

20 consecutive successes are the minimum number of responses required to terminate test. A success streak of 20 was chosen to be compatible with a more challenging cognitive test subsequently conducted. If the target streak size of 20 was not attained, 32 responses terminated the test. The choice of 32 resulted from a pilot trial which explored a number of repetitions before participants felt fatigue.

There are 7 participants giving a maximum total number of 7 x (6 x 32) trials.

D. *Data capture*

COMPTEST results are represented as a 32-bit field, each bit representing an OK/NOK (not OK) outcome for a particular trial. The field can be represented by a list of success and failure streaks. Failure streaks are suffixed with x. For example, for a list of 17 successes followed by 3 failures, 10 successes and 2 failures, the list is represented as {17, 3x, 10, 2x}. The consecutive successes and failure are referred to as success and failure streaks respectively. The entire list which captures an entire block of trials is referred to as an outcome event sequence. The maximum number of consecutive successes in the example list is 17. Multiple maxima may exist in general.

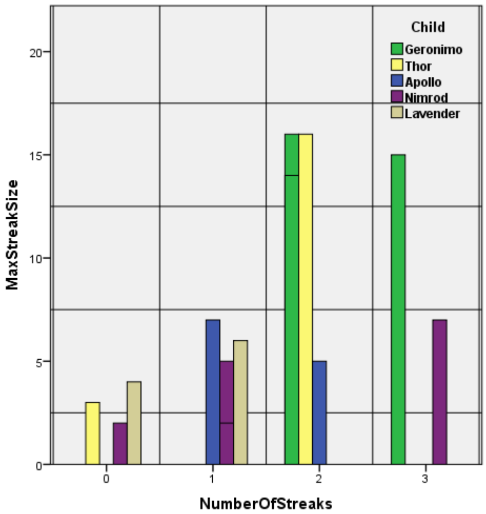


Fig. 2. Variation of maximum streak size with number of significant streaks

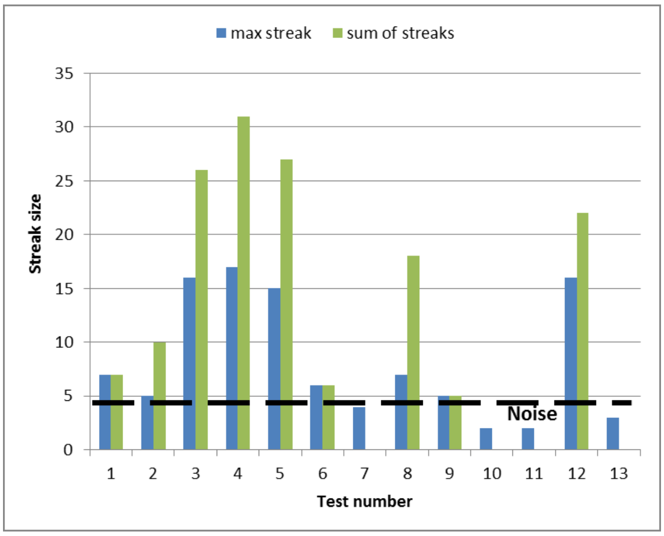


Fig. 3. Comparison of maximum-streak and streak-sum measure

IV. RESULTS

Fig. 2 shows the distribution of the maximum number of streaks with respect to the number of streaks that are significant for each block of test trials with a device. Only the tests with test outcomes that did not end with getting the target score of 20 were chosen because a test condition where the length of the test was not shortened was of interest.

Fig.2 also shows the distribution of children with respect to the number of significant streaks. We find that Geronimo always has more than one significant streak. Apollo and Nimrod are the children of interest who displayed more than one significant streak but not high maximum streak sizes.

Fig. 3 compares the maximum-streak measure with the streak-sum measure. The plot shows the streak size on the vertical axis and a test number on the horizontal axis representing a particular test run (block of 32 trials). The maximum-streak provides the maximum number of

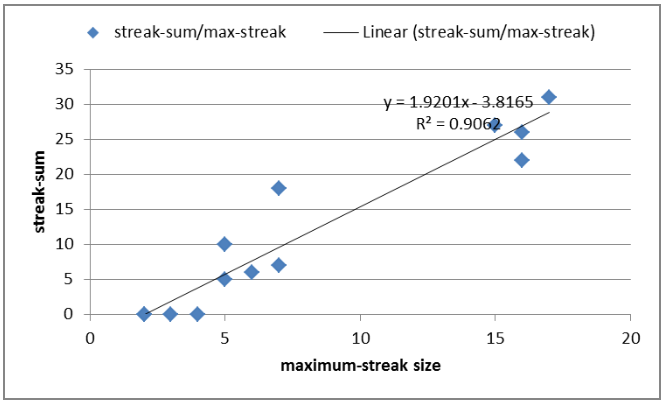


Fig. 4. High positive correlation: maximum streak size by sum of significant streaks

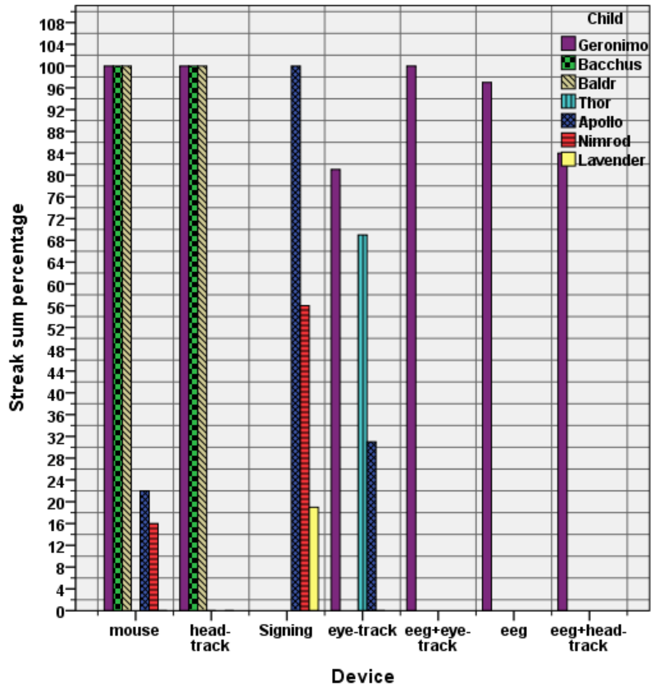


Fig. 5. Comparison of devices ranked by sum of intentional streaks (percentage) measure. Signing should be the best device but in this case, is only used when the children cannot use other devices or have very poor outcomes.

consecutive successes in a test run. The streak-sum is the sum of all the significant streaks (streaks with sizes above the noise level) in a test run. The test runs are not arranged in any particular order and each test run is performed with a child using a device. Each child will use a different device for each test run.

We find maximum streaks of all sizes. The streak sizes (15-20) that are close to the target streak size of 20 tend to distance themselves from the group that has steak sizes (5-7) close to the noise level.

When the streak-sum is present (above the noise level), it always tracks (is equal to) (Test number 1, 6 and 9) the maximum-streak or exceeds it (Test number 2 to 5, 8 and 12). The streak-sum is absent for Test number 7, 10, 11 and 13.

The relationship between the maximum-streak size and the streak-sum size was investigated with Pearson's product-moment correlation coefficient as shown in Fig.4. There were no violations of normality. The two sets of scores correlated strongly and positively ($r(11) = 0.952$, $p \le 0.000$).

Fig. 5 shows the use of the streak-sum measure to compare test results of children centered on each device. The results are ordered from left to right according to the number of children successfully accessing the device first, followed by decreasing values of streak-sum percentage. A score of 100% means that the child responded correctly all the time (a target of 20 terminates the test); other values indicate the proportion of streak-sum to the number of trials (32 trials being the maximum). For a comparison of physical devices, Signing should be ignored as it is not a true device. The mouse is the most accessible in this study (5 children), followed by the head and eye-tracker (3 children). Finally, the ones with just one child using them are the EEG and eye-track hybrid, followed by the EEG headset and the EEG and head-track hybrid.

V. Discussions

Fig. 2 identifies the control participant (Geronimo), chosen for his maturity in developmental ability with high maximum-streak sizes. In addition, he also shows multiple re-attempts with the number of streaks above 1. The other child with a high maximum-streak size also has done the test run with 2 attempts. This is the expectation that participants who engage with the test will usually re-attempt to achieve the target streak size.

Fig. 3 demonstrates that the streak-sum parameter provides more information than the maximum-streak parameter. Where no bar is presented for streak-sum is an indication that the outcomes have fallen below the noise level. The streak-sum combines both the sizes of the maximum-streak and the other passable streak sizes when there are re-attempts. Streak-sum on its own can be used to replace the maximum-streak parameter. Streak-sum like the maximum streak parameter ranks outcomes by consecutive successes except that it uses the total number of significant successes per test run. The bounds of streak-sum therefore range from 0 to the number of trials per test run (32 in this case). There are no streak-sum values below the noise level as only significant streaks (5 and above) are added up. There is no need for streak-sum to consider a noise level because noise has been filtered out (all values < 5). However, streak-sum can still be evaluated relative to the noise level as the streaks that are summed do not have the noise level removed. By placing streak-sum beside the maximum-streak parameter, the single attempt results are those which have equal heights for both parameters. The interest is for those test runs (test number 2, 8) where the maximum-parameter is close to the noise level but the streak-sum is much higher, indicating that although each attempt is limited in size, multiple attempts have been made.

Test number 2 was executed by Apollo using an eye-tracker. Apollo is a child that has severe impairments, being wheelchair bound and suffering from involuntary muscular activity. The impairments prevent Apollo from operating physical devices appropriately and over prolonged duration. However, the results show Apollo producing multiple streaks in the operation of a physical device similar to the children who displayed better control of their physical devices. Apollo also achieved the target streak size with Signing (not shown in plots) indicating that he is cognitively adequate to take the test. Apollo appears to have reached a ceiling with the physical device far below the target at a streak size of 5.

Test number 8 was executed by Nimrod using Signing. Nimrod is a child that achieved a low maximum streak size using Signing. However, he is also exhibiting multiple attempts in the test. In this particular case Nimrod may have reached a ceiling or the build-up time required is beyond the 32 trials afforded by the test. The ceiling may be due to the nature of the Signing (and so behaves similar to cases where physical devices have problems) or it may be due to a cognitive problem.

Test number 12 was executed by Thor using an eye-tracker. Thor like Apollo also has severe impairments, is wheelchair bound and suffers from involuntary muscular activity. However, in this case Thor has been provided with good affordance from his eye-tracker device and exhibits characteristics similar to a developmentally able child (Geronimo) in displaying both high maximum streak size and a number-of-streaks above 1 (re-attempts).

Fig. 5 highlights the dependency between a device and a child. Each PNI child has a set of impairments that effects access to different devices. Able-bodied people on the other hand should be able to access all devices (because the devices are designed for the general population) and ranking of devices can take place on the basis of measures such as performance timings. For PNI children, the focus is more on fitting an appropriate device and very often none is found because the effort to fully customize devices to very specific impairments does not exist. Even if a child is able to access a device, two separate effects impact children with cognitive impairments specially (and probably developing children also); the cognitive load imposed by the device they are using and the cognitive load imposed by the test. The culmination of these impacts point to a strong dependency between a triad consisting of the user, the device and the test; all of which have a strong bearing on the results of the specific test.

VI. Conclusions

The study suggests that children who have the ability to use a device well in COMPTEST, display their competence with high maximum streak sizes as well as typically re-attempting to build up to the target streak size with multiple attempts. Children who have displayed poor maximum streak sizes may or may not re-attempt. The re-tries provide a higher confidence for intention. Using a streak-sum parameter allows the re-attempts to be captured within a single parameter as well as the information provided by the maximum-streak. A comparison of streak-sum and maximum streak enables us to identify cases where the streak outcomes may be restricted to low values but re-attempts have been made. This is important because impairments tend to cluster maximum streak values to a range of values that are close together. This measure allows us to rank the devices further to produce a match in a larger research project investigating a range of different non-hand held devices and their suitability to intended users. We speculate that single attempts are made due to a combination of device imperfections, unmatching device affordances, fatigue and inattention unless the single attempt achieves the target streak size. The limitations for this study are the small sample size that was used. Based on this limitation, it was difficult to gather sufficient information on the different combination of devices and "types" of PNI with other variations in levels of impairment.

Another limitation was the absence of non-PNI children. Our next study explores the same experiment with the non-PNI children. The presence of non-PNI children as participants would also act as "controls" for the study.

The future directions for this research include a number of directions. We intend to look at even more measures as well as patterns of outcomes and not just maximum-streak and streak-sum. A specific measure of interest is the position in the event sequence in which a single streak occurs. An early position can be indicative of fatigue and a late position could indicate inattention. We will explore if a combination of different measures and patterns can produce categorizations of test behaviour. The categorizations will be used to support the attempts to match devices to impairments.

ACKNOWLEDGMENT

We would like to thank Mr. I. Glasscock of BioDigitalHealth (www.BioDigitalHealth.com) for his support.

We are grateful to the schools that agreed to participate in this study especially the staff of the special needs schools for their time and effort, and to the participants and their care-givers/families in assisting us with this study.

REFERENCES

[1] C. Cans *et al.*, "Recommendations from the SCPE collaborative group for defining and classifying cerebral palsy," *Developmental Medicine & Child Neurology,* vol. 49, pp. 35-38, 2007.

[2] Emotiv. "Emotiv EEG Features. http://www.emotiv.com/eeg/features.php," June, 2013.

[3] A. B. Barreto, S. D. Scargle, and M. Adjouadi, "A practical EMG-based human-computer interface for users with motor disabilities," vol. 37, no. 1, pp. 53-64, 2000.

[4] C. E. Reyes *et al.*, "A study on ocular and facial muscle artifacts in EEG signals for BCI applications," *TENCON 2012-2012 IEEE Region 10 Conference*, pp. 1-6, 2012, 2012.

[5] Tobii Technology. "Portable lab - Tobii X2-30 Eye Tracker. http://www.tobii.com/fr/eye-tracking-research/global/products/hardware/tobii-x2-30-eye-tracker/," June, 2013.

[6] C. Ware, and H. H. Mikaelian, "An evaluation of an eye tracker as a device for computer input2," *SIGCHI Bull.,* vol. 18, no. 4, pp. 183-188, 1986.

[7] Boston College. "CameraMouse. http://www.cameramouse.org/," June, 2013.

[8] M. Betke, J. Gips, and P. Fleming, "The camera mouse: visual tracking of body features to provide computer access for people with severe disabilities," vol. 10, no. 1, pp. 1-10, 2002.

[9] F. Amirabdollahian *et al.*, "The I-match project: A VR based system to allow matching of an optimum interface to a user of assistive technology," *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, pp. 526-529, 2005, 2005.

[10] F. Amirabdollahian, G. T. Gomes, and G. R. Johnson, "The peg-in-hole: a VR-based haptic assessment for quantifying upper limb performance and skills," *Rehabilitation Robotics, 2005. ICORR 2005. 9th International Conference on*, pp. 422-425, 2005, 2005.

[11] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *Journal of experimental psychology,* vol. 47, no. 6, pp. 381-391, 1954.

[12] F. S. Makri, A. N. Philippou, and Z. M. Psillakis, "Shortest and longest length of success runs in binary sequences," *Journal of Statistical Planning and Inference,* vol. 137, no. 7, pp. 2226-2239, 2007.

[13] E. M. Altmann, and B. D. Burns, "Streak biases in decision making: Data and a memory model," *Cognitive Systems Research,* vol. 6, no. 1, pp. 5-16, 2005.

[14] P. Ayton, and I. Fischer, "The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness?," *Memory & cognition,* vol. 32, no. 8, pp. 1369-1378, 2004.

[15] D. A. Grant, and E. Berg, "A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem," *Journal of experimental psychology,* vol. 38, no. 4, pp. 404, 1948.

[16] B. Libet *et al.*, "Subjective referral of the timing for a conscious sensory experience," *Neurophysiology of Consciousness*, pp. 164-195: Springer, 1993.

[17] L. M. Dunn, *The British picture vocabulary scale*: GL Assessment Limited, 2009.

21. Appendix - Papers (3)

This paper was published in the 2014 Proceeding of International Conference on Advances in Computing, Electronics and Electrical Technology CEET:

Gan, H., Frank, R., Amirabdollahian, F., Sharp, R., & Rainer, A. (2014a). Bio-digital device impact on a constant load cognitive test of children with physical and neurological impairments. *International Journal of Advances in Computer Science & Its Applications, 4*(4), 99 - 105.

Bio-digital device impact on a constant load cognitive test of children with physical and neurological impairments

Hock C. Gan, Ray J. Frank, Farshid Amirabdollahian, Rob Sharp, Austen W. Rainer

Abstract*— In this paper we extend on our earlier work utilizing maximum consecutive success as a measure used in the context of testing motor skills, for children with physical and neurological impairment. Here we verify a change in the "ceiling" of performance that takes place as the cognitive load is increased. In the results we see two groups establish themselves in terms of the ones capable of achieving the test objectives and the ones that engage but are not successful. In this study, we compare results for two different tests, one of motor skills and the second of categorization assessment and find agreement between the two using our approach. Working on our main hypothesis that it is possible to provide better fit of devices to PNI children, the results provide initial suggestions that cognitive test success depend on a combination of motor skill and cognitive ability which is not completely separable. Ranking of bio-digital devices ideally need to be done using the target application or the use of direct communication which is some form of Signing. It is the intention in a future project to further show that for some children their cognitive capacity can deal with just the motor skill for a physical device or the cognitive load for a test but not both. The ideal fit of a physical device for the child depend on both the demands of the application and device and it is not possible to draw conclusions on a fit without considering a triad which involves the user, the device and the test. The role of novel non-hand held devices in complex working environments are a potential market but fatigue is a serious problem in critical environments. This study provides a start to examining the fatigue levels imposed by in a user-device-application triad as a result of different levels of cognitive loads.

***Keywords*—disabled children, non-hand held device, motor-skill test, cognitive test, streak**

I. Introduction

Physically and Neurologically Impaired (PNI) children have special needs due to problems which result from brain injury. The effects of the brain impairment result in physical impairment and resulting in the need to address both problems.

Hock C. Gan, Ray J. Frank, Farshid Amirabdollahian, Austen W. Rainer
University of Hertfordshire
United Kingdom
h.c.gan@herts.ac.uk, r.j.frank@herts.ac.uk, f.amirabdollahian2@herts.ac.uk, a.w.rainer@herts.ac.uk

Rob Sharp
Hertfordshire County Council
United Kingdom
rob.sharp@hertfordshire.gov.uk

A large proportion of such children suffer from Cerebral Palsy (CP) [1], an umbrella term for a range of disabilities that include a combination of movement, speech, cognitive processing, and visual perception. Education of children with disabilities that affect communication such as motor and speech is difficult.

Education, which is a two-way process of disseminating and assessing knowledge, is hampered if the communication pathways are restricted or degraded. Received knowledge is distorted and assessment is inaccurate. Assessment in this case is challenging and perhaps even more necessary. Current assessments typically involve tests that are digital and require interaction with an input device, usually a mouse. Unfortunately, the involuntary actuation and inhibition of movements of a number of PNI children mean that such children have great difficulty controlling devices used in tests.

New devices are available that involve other affordances that potentially requires skills that the PNI child may be more capable of. This resulted in the following devices being chosen for this study:

1. An eye-tracker [1, 2]
2. A face/head-tracker [3, 4]

This study also considers Signing as a method of input, and Signing provides a contrast to bio-modal inputs and the typical physical inputs. Signing in this case refers to a child who communicates using gestures to an interpreter. The gesture acts as the child's response to a test, and this response is then entered as a mouse input via the interpreter. Signing was used to confirm that a child was able to understand the test without the impact of dealing with a physical device in cases where there was uncertainty with the results. Signing was deemed to impose the least cognitive load for manipulating a device as the child has had years of using and developing it to use the parts of the body which were functionally capable for purpose and a carer who also was able to pick up on subtleties of the communication. The replacement of Signing using a physical device loses some of those advantages.

The use of new devices necessitates a means of evaluation to rank and match the devices to the user. An accompanying paper described work which establishes the procedure and a measure (maximum-streak) used to evaluate the results. A difference in the approach used by the paper is the use of streaks or consecutive successes instead of a simple count of total success for assessment of PNI children. Streaks have cognitive tests like the Wisconsin Card Sorting Test (WCST) [8] where reinforcement or learning is concluded after a certain number of consecutive successes. The maximum streak size was regarded as a measure for intention, representing a "ceiling" of performance in a block of trials. Intention is a reference made for an action that is not the result of a reflex action but one arising from a conscious decision[9]. The intention response was separated from the response cases where no conclusion could be drawn through the establishment of a threshold level. The threshold level is a streak size by which streaks of a smaller size would be regarded as noise and a larger size as significant streaks indicating intention. A different level (a target streak size) was set to indicate achievement of general required performance for a future cognitive test. Using the maximum-streak measure the paper obtained relative ranking of performances primarily of the capable children with a device and the less capable ones using a simple motor skills test (COMPTEST) designed proprietarily for cognitive assessment of PNI children. Motor skills have to be learnt and require a cognitive component in execution. This cognitive component is reduced as we train the skill[10]. COMPTEST as a motor skills test was designed with a low cognitive load from the viewpoint that the cognitive component required does not go much beyond responding to a stimulus. The use of novel devices provided in COMPTEST required new motor skills. COMPTEST ranks the ability of the child using a device and the ranking includes a measure of the cognitive ability required primarily to operate the device. A portion of the cognitive ability is still required to understand the stimulus and decide to respond to it. If we assume that all tests generally have these two basic cognitive quantities then all tests can be differentiated by the relative amounts of cognitive effort required for the operation of the device and the specific cognitive demands of the test. Another way of looking at this is that in any test, there will always be a portion of cognitive effort required that is not related specifically to the test and the opposite applies; that in any motor skills test, there will always be a portion of effort not specifically related to driving the device. This paper extends the work done previously by introducing a new test that has more cognitive complexity than COMPTEST. The aim is to use the same measure (maximum-streak) to quantify changes and make comparisons and thus verify maximum-streak as a first candidate measure.

A Categorization test (CATTEST) is therefore introduced by this paper which serves to provide more cognitive complexity than COMPTEST, increasing the cognitive component required to pass the test. According to Piaget and his 4 stages of cognitive development, classification is a feature that occurs consistently in the Pre-operational stage (ages 2 to 7) within limits [11]. The test involves recognising birds and fruits and being able to work out that the stimulus consists of all birds or fruits or not. The cognitive component that is increased in this case has no particular focus on device operation although there may be parts of that component that is also required for device operation. CATTEST is applied on a child using his best devices. A change in results is therefore expected to be mainly due to the increase in the cognitive component. CATTEST is not designed to test cognitive limits but to introduce a general load and therefore a ranking of cognitive ability is not expected. The maximum-streak as a measure can be viewed as a "ceiling" where the best effort that can be made in a test run is achieved. The test run is designed so that it is of sufficient length where the ceiling can be reached for some children. A target is set where achieving a maximum-streak of 20 ends the test and if a child attains this, their ceiling can be higher. Introducing an additional load to a test is expected to lower the ceiling for a child. The maximum-streak measure can thus be viewed as a measure of a child's capacity to handle the demands of both device and cognitive load. Setting the cognitive load to a specific level allows for the demands of the device to be estimated and conversely, setting the device load at a specific level allows for the demands of the cognitive load to be estimated. COMPTEST does the former and both COMPTEST and CATTEST does the latter. The differentiation of device load and cognitive load identifies stimuli that are tightly coupled in the test environment. In a system which is divided into a user, device, test triad, all three components influence the generation of an outcome. The user who is influenced by their impairments (which diminish the capacity to handle a cognitive load and a device load), the input device (which generates a device specific cognitive load and modifies user responses due to its imperfections) and the test (which generates a specific cognitive load in the form of stimuli and evaluates device responses depending on design) all play a role in the final outcome which is generally regarded as a response from the user to a test.

II. Method

A. *Participants*

Seven PNI children were tested, five have various forms of CP, one has methotrexate leucoencephalopathy and one has septo-optic dysplasia with autistic spectrum disorder. Approval from both school and parents were sought under the University of Hertfordshire's ethics protocol aCOM/PG/UH/00006. Fictitious names have been used for all the children in all publications.

The ages of the participants were rationalized using the British Picture Vocabulary Scale III (BPVS III) [12] to provide the developmental age as shown in TABLE I. The rationalization groups children into developmental levels according to their verbal age. As a result of impairments

TABLE I. TABLE 1. AGE EQUIVALENTS OF PARTICIPANTS

Name	*Age*	*Severe impairment*	*Developmental age (years: months)*
Apollo	14	Yes	04:10
Bacchus	12	No	04:07
Baldr	15	No	< 04:00
Geronimo	13	No	11:03
Lavender	12	Yes	< 04:00
Nimrod	13	No	04:05
Thor	12	Yes	07:03

of the children, some BPVS results are best effort results as the case with Apollo and Thor. The table indicates three participants who are severely impaired in that they wheelchair bound, have almost no speech, and have involuntary muscular problems and weak muscular control. The remaining four suffer some impairment to a lesser degree and are not wheelchair bound.

Geronimo was picked to provide an example of a person with CP and with mature developmental capability. The others were picked as examples of children with varying impairments between the developmental ages of 4 to 7 for compatibility with another study.

B. *Procedure*

The participants are tested in a room (located in a school) equipped with a laptop, separate screen monitor and hand-held and non-hand held devices for interaction with the software tests running on the laptop. Each participant was tested on 2 of their best devices; the usual device that the participant is most familiar with (typically a mouse) and one other device which testing with COMPTEST showed as appropriate, especially if the device provided better results than the usual device. COMPTEST results can be accessed in an accompanying paper [13, 14]. Some participants could only use one device because of impairments.

The screen monitor is arranged side to side with the laptop so that the participants with a view to the screen monitor are seated beside the researcher who has a view to the laptop. The eye-tracker is mounted below the screen of the monitor using magnetic mounts. The face-tracker uses a remote web camera enabling tracking from the monitor. The order of devices tested was arranged to minimize anticipated boredom and fatigue so that the easiest would be done first.

A set of stimulus is produced which the participant must provide a positive or negative response. The positive response is an active response which involves the actuation of a physical or virtual device. The negative response is passive requiring no action. Success terminates the trials for each device determined by 20 consecutive correct responses. Otherwise, the trials terminate after a block of 32 trials. A single session of two blocks of trials involving different devices and tests are conducted within an hour. During the tests observations were made by the researcher regarding test response behaviour and notes compiled after testing. Devices were run in order of increasing complexity. The stimulus for the CATTEST consists of images of birds and fruits [15]. Participants are expected to provide a positive response when they see a bird go into a bird cage consisting of 2 different birds or a fruit placed onto a fruit bowl holding a few different fruits. A negative response is required when a bird is placed in the fruit bowl or a fruit is placed in the bird cage. No feedback is provided to indicate if the response is correct or incorrect but actuation of the virtual switch produces a click. Participants are familiarized with a different set of birds and fruits which does provide feedback if the response is correct or incorrect.

When an eye-tracker is used, participants have to move a mouse cursor with their eyes to an image of a switch and "dwell" the cursor over the virtual switch for a period identified as the "dwell-time". When a head-tracker is used, participants have to move a mouse cursor using movement of their head to the virtual switch for the dwell-time.

C. *Design*

Fig. 1 shows the test design. The experiment is a 2 x 32 within-subjects factorial design for the maximum number of trials. For the participants who can only use 1 device due to impairments the design is 1 x 32. For each participant the following are components of the test:

- Device {(Signing), mouse/switch, eye-track, head-track}. Signing is only used when no other devices can be used and does not increase the maximum count of devices for trials.
- Block {1 to a range between 20 and 32 trials}
- Tests {CATTEST}

There are 7 participants giving a maximum total number of 7 x (2 x 32) trials. This is only the maximum total as it would depend on whether there was more than one appropriate device (1 or 2) or more than 20 trials (20 to 32).

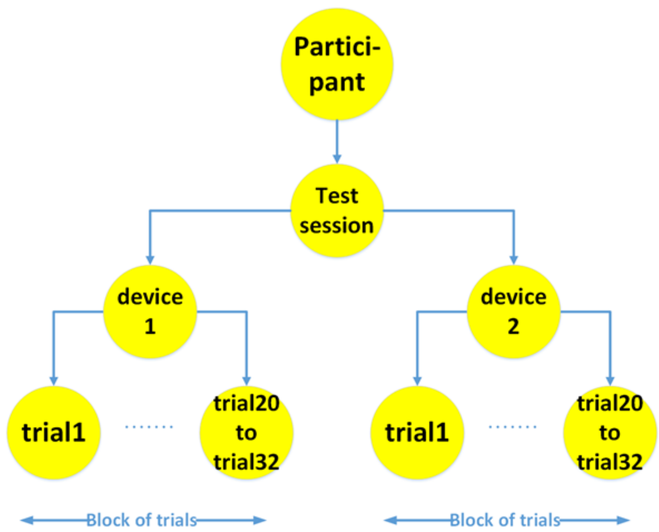


Figure 1. Test design

D. *Data capture*

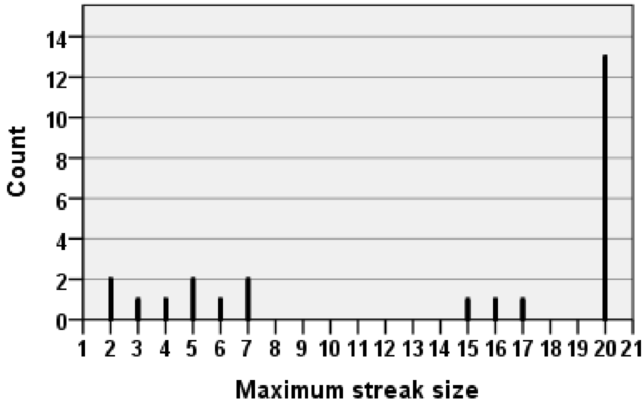
CATTEST results are represented as a 32-bit field, each bit representing an OK/NOK (not OK) outcome for a particular trial. The field can be represented by a list of success and failure streaks. Failure streaks are suffixed with x. For example, for a list of 17 successes followed by 3 failures, 10 successes and 2 failures, the list is represented as {17, 3x, 10, 2x}. The consecutive successes and failure are referred to as success and failure streaks respectively. The entire list which captures an entire block of trials is referred to as an outcome event sequence. The maximum number of consecutive successes in the example list is 17.

III. Results

Fig. 2 (top) gives a view of the number of trial blocks having specific attainment levels in CATTEST. Each maximum streak size is generated from a block of trials that a PNI child has undertaken using a specific device. Each bar represents the number of blocks that had achieved a specific maximum streak size. 7 participants were tested with a maximum of 2 devices (giving 7 x 2 blocks) but 4 were unable to engage with either the device or test, leaving 10 blocks. Streak sizes of 0 (from children who do not engage with the test) are ignored as they provided no inputs.

Fig. 2 (top and bottom) shows a comparison of the maximum streak distribution for both CATTEST and COMPTEST. The distribution for COMPTEST shows the results clustering by the low (2-7) and high (15-20) end. The CATTEST results have an entry in the mid-range (11) as well as the low (2-6) and high end (18-20).

Fig. 3 shows the variation of frequency of the success-streak size of the population of low performers for CATTEST with the variable.



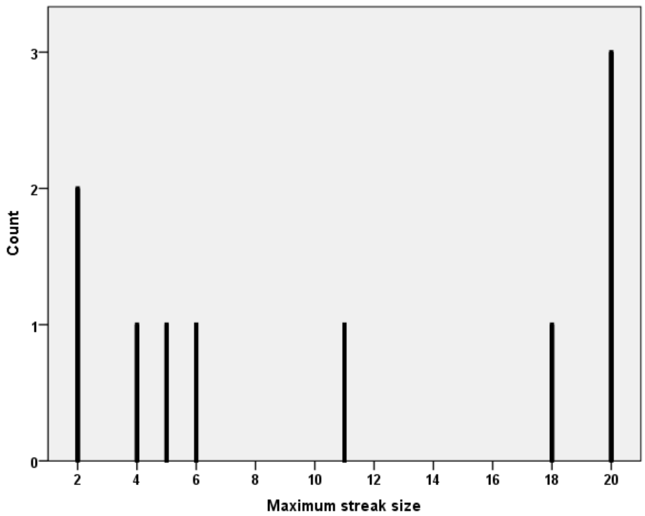


Figure 2. Maximum success streak size distribution; Streak sizes of 0 are ignored as the participants did not engage with the test – CATTEST(top), COMPTEST(bottom) from a parallel study [13]

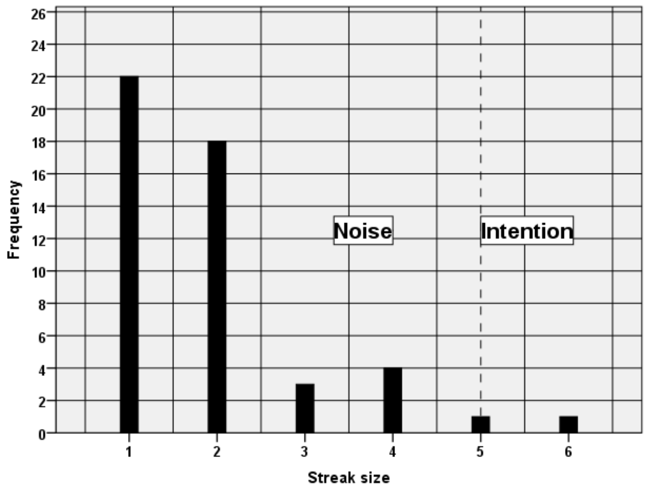


Figure 3. Success-streak size distribution of low performers

TABLE II shows the extent of the general decrease in frequency. The lower sizes (1 and 2) are the most common, occupying 83% of the streaks with the next two sizes (3 and 4) taking up more than 12% of the streaks. The rest occupy the remaining count (under 5%).

TABLE II. SUCCESS-STREAK SIZE CUMULATIVE DISTRIBUTION OF LOW PERFORMERS

Streak size	Frequency	Percent	Cumulative Percent
1	21	44.7	44.7
2	18	38.3	83.0
3	2	4.3	87.2
4	4	8.5	95.7
5	1	2.1	97.9
6	1	2.1	100.0
Total	47	100	

Fig. 4 shows the maximum streak size achieved for each run of CATTEST using a maximum of two different devices. The devices were among the best devices (determined from COMPTEST Fig.5) available for the child. The children are ordered in the plot in order of ability from left to right. Geronimo who showed good developmental ability was chosen to be a basis for reference for this test. We see Bacchus (20, 20) outperforming Geronimo (18, 20) in terms of maximum streak size in this test run by a small margin. The results then drop to around the mid-way value (11) for Thor. The level then drops further with Nimrod (5, 6) making it just above the noise threshold. Apollo falls just below the noise level with 4. Baldr falls far below the noise threshold with 2. Lavender (0) did not carry out the test. Signing was used for Nimrod, Apollo and Lavender because results for the mouse were low for them in COMPTEST. In this test we see the Signing results verifying the mouse results for Nimrod (5, 6) but Apollo did not have an alternate choice. In terms of devices, we see that the best devices provide similar results. There is not much difference between using one or the other.

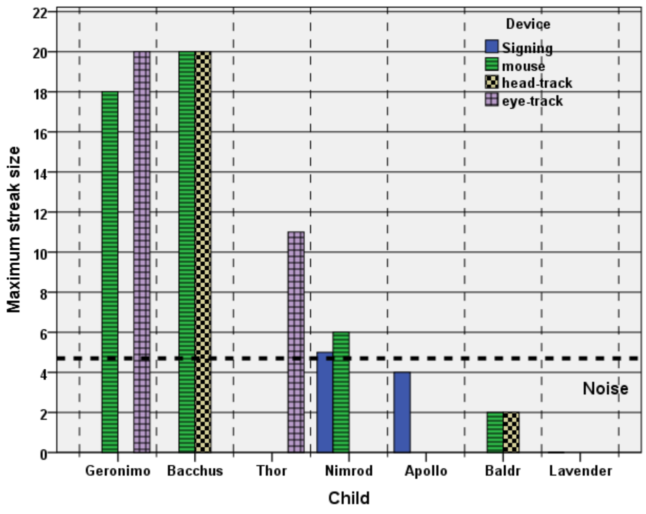


Figure 4. CATTEST maximum streak results for a PNI child using his/her best devices

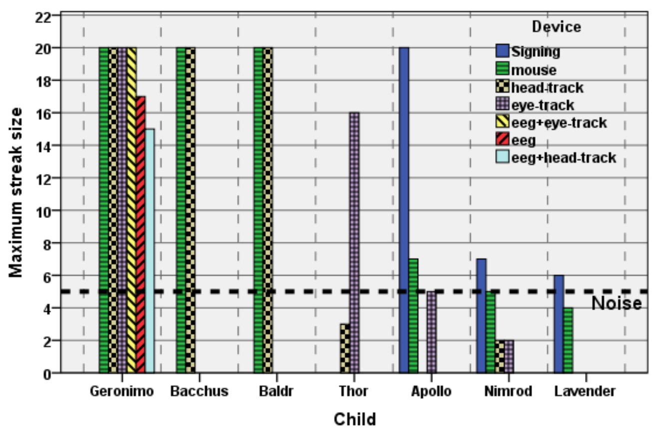


Figure 5. COMPTEST results for devices to be used in CATTEST

Fig.6 shows a comparison of the maximum streak sizes achieved by the children while taking the COMPTEST and CATTEST. The devices that best represented the children were used. Signing was used when the results using physical devices were poor. The results are arranged in order of children achieving the best performances in COMPTEST. Half the number of children (Bacchus – Apollo) managed to reach the required target streak size (20) when taking COMPTEST. This was followed by Thor who achieved a maximum streak size (16) well above the noise threshold. The remainder of the children (Nimrod, Lavender) achieved a maximum streak size close to the noise threshold (7, 6). If we now look at a comparison with the CATTEST results, we find that in general the CATTEST results are lower than the COMPTEST results with the exception of Bacchus who achieved the same results. Proceeding onwards from left to right after Bacchus, we find Geronimo attaining a CATTEST result of 18. Geronimo's CATTEST result is close to the COMPTEST high result. Next we find Baldr and Apollo (CATTEST result of 4 and 2) who managed the target in COMPTEST but failed to attain a CATTEST result above the noise level. Thor attained a fair result in both CATTEST (11) and COMPTEST (16) but the CATTEST result being significantly poorer. Nimrod had poor results in both CATTEST and COMPTEST (5 and 7). Lavender was unable to understand the familiarisation exercise with CATTEST using Signing and the actual test was not carried out.

The mean maximum-streak size in CATTEST ($M=8.57$, $SD=7.913$) was lower than the mean maximum-streak size in COMPTEST ($M=15.57$, $SD=6.373$) resulting in a mean decrease ($M=7$, $SD=7.141$) in the ceiling (maximum-streak size). This decrease was statistically significant, $t(6)=2.593$, $p<0.041$, two-tailed. However, the sample size is small.

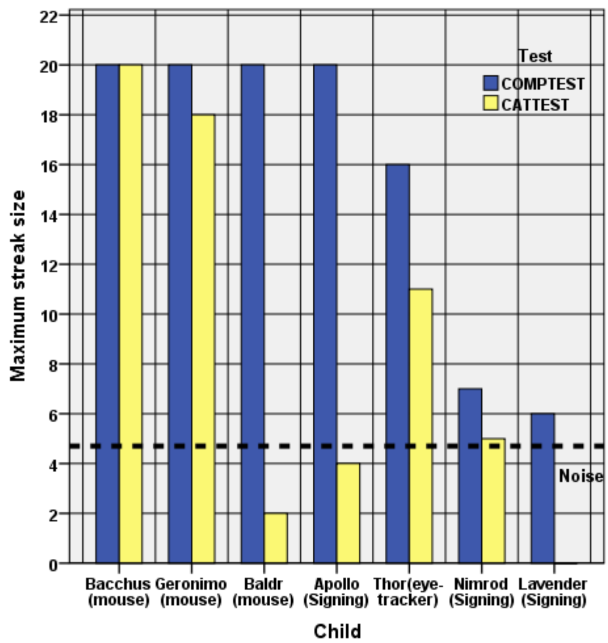


Figure 6. Comparison of motor skills test (COMPTEST) and cognitive test (CATTEST) using the best device

IV. Discussion

In general, the nature of the both COMPTEST and CATTEST encourage outcomes that are binary (pass/fail) which is demonstrated by Fig.2. Using maximum-streak as a criterion would mean that the measure favours high outcomes. The tests are straight-forward and at the high end of the results the target is sometimes not achieved due to minor distractions. The low-end of the results signify a ceiling that the child is unable to exceed due to problems with motor operation or something of a more cognitive nature. However, the CATTEST maximum-streak distribution also show an intermediate value appearing near the mid-range (11) towards the high end. A possibility is that the change in cognitive requirement of CATTEST has imposed a new ceiling on a child. Where a child was performing at a higher level before with a device, the increase in cognitive load of CATTEST has imposed a lower ceiling. Fig.5 identifies the children at the high end as Geronimo and Bacchus, the mid-range as Thor and the lower range as Apollo, Nimrod, Baldr and Lavender.

Fig.6 shows children who have no change in the results between CATTEST and COMPTEST at the high end (Geronimo and Bacchus) as their ability is able to cope with the demands of the additional cognitive load imposed by CATTEST. A proportion of the children (Apollo, Baldr, Lavender) who achieved outcomes above the noise threshold in COMPTEST and having outcomes that were below the noise threshold in CATTEST were not able to cope with the additional demands of CATTEST. Apollo and was tested on CATTEST using Signing so an initial assumption was that there was a problem with CATTEST. Baldr used the mouse and was not tested using Signing. Baldr kept clicking the mouse for all stimuli and was also assumed to have a problem with CATTEST but requires checking using Signing. Lavender was also tested using Signing but had problems of inattention. The two remaining children (Nimrod and Thor) have their CATTEST results above the noise threshold and somewhat below the target size. Thor has involuntary muscle activity that causes instability in the motion of the torso. The additional cognitive load appears to have lowered the ceiling he operates on in that the maximum-streak has decreased between COMPTEST and CATTEST. Nimrod does not have the physical problems that Thor has but does have cognitive problems. Nimrod too appears to have his ceiling lowered by a small amount but Signing was used.

V. Conclusion

The CATTEST results provide a view of a child's ability that is focussed on a cognitive function, namely categorization. However, the results reflect not only the cognitive ability for the test but also the effort required to manipulate the input device.

The maximum-streak distribution of CATTEST mirror those of COMPTEST quite well in that there are clusters of good performances and clusters of low performance. The increase in cognitive load in CATTEST produced a noticeable change in that the CATTEST maximum-streak results would be equal to or be less than the COMPTEST results. With the increase in cognitive load, there has been a decrease in the "ceiling" of performance. CATTEST had 3 out of 7 failures compared to COMPTEST and is an extreme reduction of the ceiling which does not provide conclusions on the loading as it could be due to total disengagement from the test. However, two results within CATTEST were able to show a lowering of the ceiling within the scale of the test which lends support to the suggestion that the issue is due to performance rather than total disengagement. The interest in this case is that the performance may be linked to fatigue and some form of evaluation is then possible using relatively similar procedures and measures.

There is a need to validate the results but the basis of communication of the severely impaired is an input device that is not perfectly reliable in many cases. The use of direct communication using Signing is one way to get round the issue but is by no means perfect. Signing was used initially in this study for children who do not have a device that can represent them adequately. However, it can be used to show that the child can succeed in the test without the use of a physical device especially in the extreme case where the results fall below the noise threshold. This supports the case that failure is not due to the inability to overcome the test of cognitivity but failure to overcome the load of both the device and the cognitive function tested. The separation of the device and the cognitive component tested is not achieved even in the use of Signing because there are still intermediate functions that perform the input and output functions for the child. The implication is that we can approach a limit where the translation functions (which require cognitive effort) are minimal but never absent until an accurate mind-reader is found.

Finally, it is recognised that this study requires a larger number of participants to enable more results that do not fall out to the extreme ranges of ones that are near the target objective or below the noise threshold. However, a fatigue test could be run where the number of trials in a block could be extended beyond 32 to capture more intermediate results.

There are many application areas for every-day use of non-hand held devices. Earlier examples of these are related to the aviation industry[16] but with the rise of embedded reality and its inclusion in 3D home entertainment[17], there comes a need to identify mechanisms where fatigue can be measured and minimized. Within this context, understanding the contribution of the triad of user-device-application is an important and elemental key to the success of these devices and furthermore, their integration in their intended application domains.

Acknowledgment

We would like to thank Mr. I. Glasscock of BioDigitalHealth (www.BioDigitalHealth.com) for his support. We are grateful to the schools that agreed to participate in this study especially their care-givers/families in assisting us with this study. This research is supported by the Engineering and Physical Sciences Research Council (EPSRC UK) under iCASE 09001842.

References

[1] Tobii Technology. "Portable lab - Tobii X2-30 Eye Tracker. http://www.tobii.com/fr/eye-tracking-research/global/products/hardware/tobii-x2-30-eye-tracker/," June, 2013.

[2] C. Ware, and H. H. Mikaelian, "An evaluation of an eye tracker as a device for computer input2," *SIGCHI Bull.,* vol. 18, no. 4, pp. 183-188, 1986.

[3] M. Betke, J. Gips, and P. Fleming, "The camera mouse: visual tracking of body features to provide computer access for people with severe disabilities," *Neural Systems and Rehabilitation Engineering, IEEE Transactions on,* vol. 10, no. 1, pp. 1-10, 2002.

[4] Boston College. "CameraMouse. http://www.cameramouse.org/," June, 2013.

[5] F. S. Makri, A. N. Philippou, and Z. M. Psillakis, "Shortest and longest length of success runs in binary sequences," *Journal of Statistical Planning and Inference,* vol. 137, no. 7, pp. 2226-2239, 2007.

[6] E. M. Altmann, and B. D. Burns, "Streak biases in decision making: Data and a memory model," *Cognitive Systems Research,* vol. 6, no. 1, pp. 5-16, 2005.

[7] P. Ayton, and I. Fischer, "The hot hand fallacy and the gambler's fallacy: Two faces of subjective randomness?," *Memory & cognition,* vol. 32, no. 8, pp. 1369-1378, 2004.

[8] D. A. Grant, and E. Berg, "A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem," *Journal of experimental psychology,* vol. 38, no. 4, pp. 404, 1948.

[9] B. Libet *et al.*, "Subjective referral of the timing for a conscious sensory experience," *Neurophysiology of Consciousness*, pp. 164-195: Springer, 1993.

[10] P. M. Fitts, and M. I. Posner, "Human performance," 1967.

[11] L. A. Sroufe *et al.*, *Child development: Its nature and course*: Mcgraw-Hill Book Company, 1992.

[12] L. M. Dunn, *The British picture vocabulary scale*: GL Assessment Limited, 2009.

[13] *H. Gan et al.*, "Development of the maximum-streak measure for evaluating the suitability of non-handheld devices in cognitive tests of PNI children," *International Conference on Advances in Computing, Electronics and Electrical Technology,* vol. (submitted), 2014.

[14] *H. Gan et al.*, "Use of re-attempts measure for evaluating device test results of children with neurological impairments," *7th International Conference on Human System Interaction*, 2014.

[15] P. C. Quinn, J. Oates, and J. O. A. Grayson, "Early category representation and concepts," *Cognitive and language development in children*, pp. 21-60, Maidenhead, BRK, England: Open University Press, 2004.

[16] L. Prinzel, and M. Risser, "Head-up displays and attention capture," *NASA Technical Memorandum,* vol. 213000, pp. 2004, 2004.

[17] S. Henderson, and S. Feiner, "Opportunistic tangible user interfaces for augmented reality," *Visualization and Computer Graphics, IEEE Transactions on,* vol. 16, no. 1, pp. 4-16, 2010.

About Author (s):

Hock Gan: At any point in time, our attention is necessarily divided and the result is that we are often judged by what we appear to achieve but not what we can achieve.

Ray Frank: An interesting aspect of this work is that it enables almost all PNI children to demonstrate their motor/ cognitive skills over a wider range of motor and cognitive development compared to simple pass/fail tests giving educational psychologists a far better assessment tool.

Farshid Amirabdollahian: In the context of assistive technology and devices assigned to children with cognitive and motor impairments, it is important to consider complexities offered by each device and each test, with regards to motor and cognitive loads.

Rob Sharp: Children who undergo learning tests express a range of emotions including surprise, frustration and satisfaction. I want a device that will tell me that.

Austen Rainer: It's interesting to observe that even neuro-typical, "normal" children can struggle to use some of the devices when undertaking the tests. My assumption was that all of these children would be capable of successfully passing the tests using the different devices.