International Journal of Advances in Computer Science & Its Applications – IJCSIA
Volume 4: Issue 4 [ISSN 2250-3765]
Publication Date : 27 December,2014

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.

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

Hock C. Gan, Ray J. Frank, Farshid Amirabdollahian, Austen W. Rainer
University of Hertfordshire
United Kingdom

Rob Sharp
Hertfordshire County Council
United Kingdom

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.

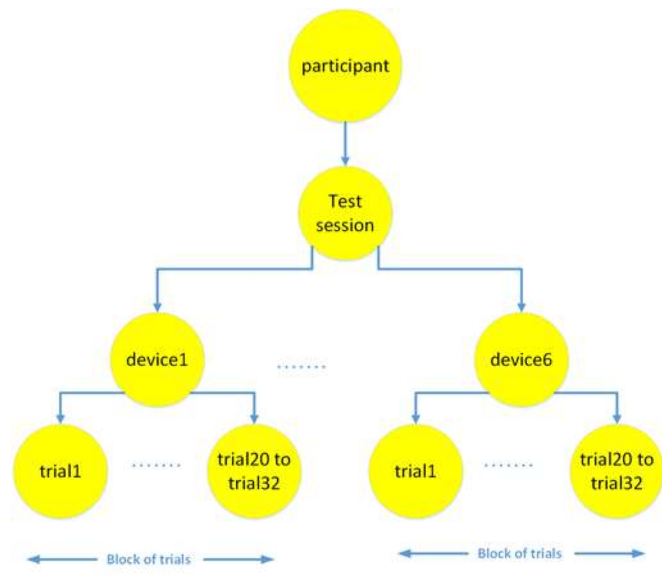

Figure 1. Test design

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].

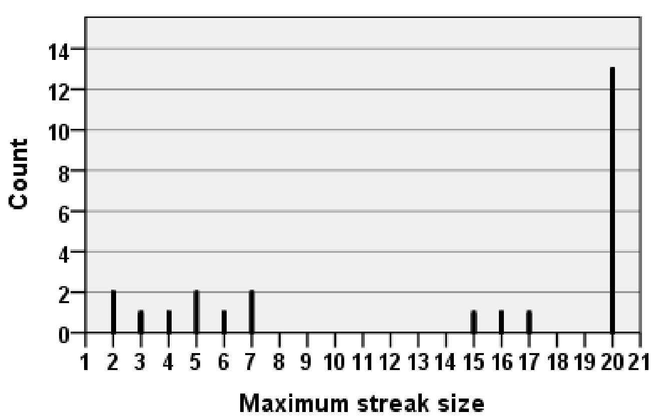

Figure 2. Maximum success streak size distribution

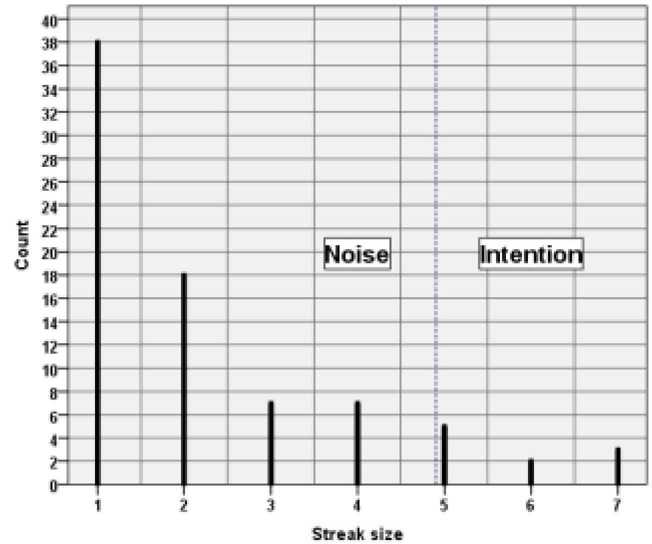

Figure 3. Success streak size distribution of severely impaired

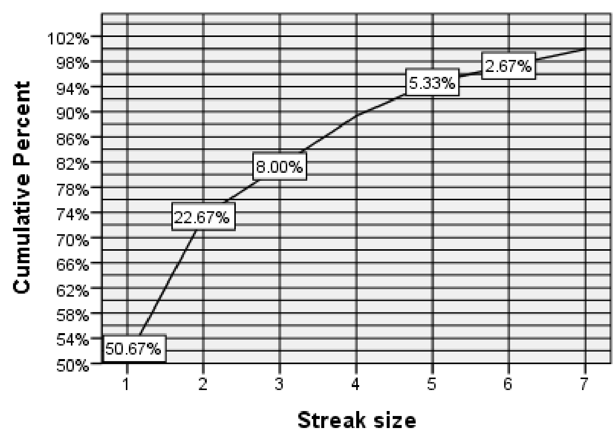

Figure 4. Success streak size cumulative distribution of severely impaired

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.

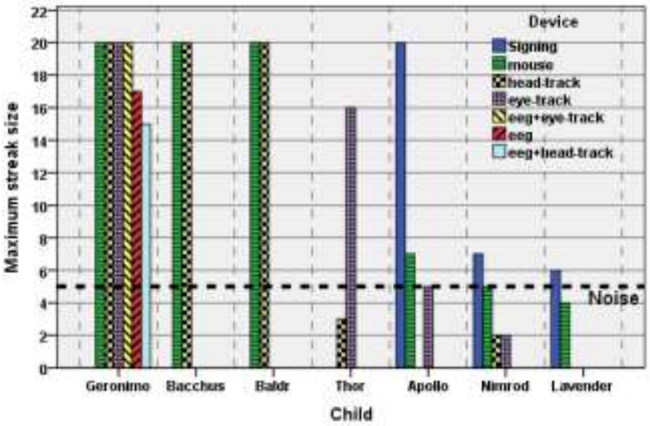

Figure 5. PNI child device ability. For physical device evaluation, Signing should be discounted.

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.

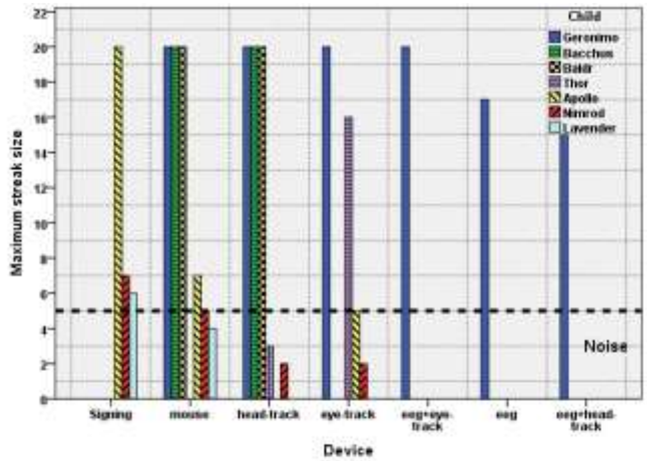

Figure 6. Device competency. For physical device evaluation, Signing should be discounted.

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.

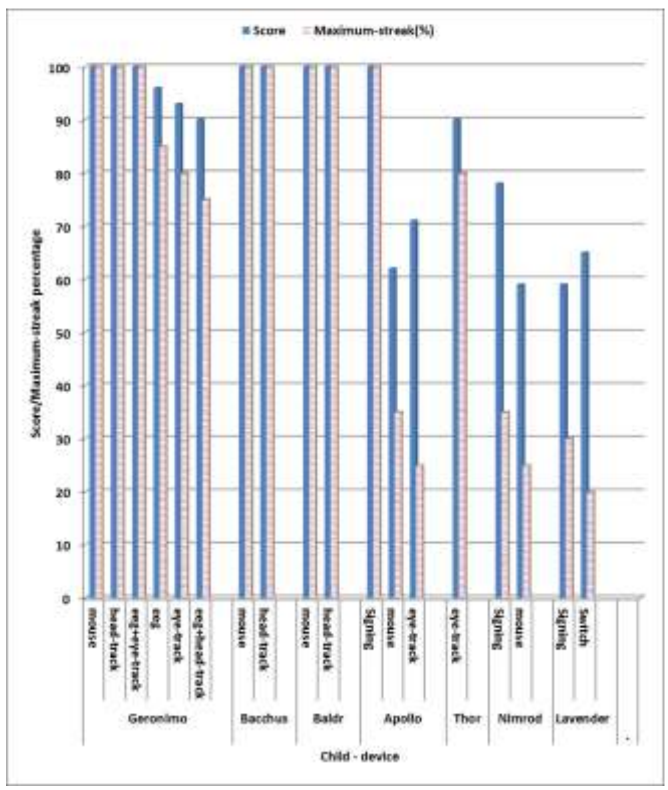

Figure 7. Comparison of maximum-streak measure with success score sorted by child and maximum-streak

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

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.