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 nonhand 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, motorskill 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

Rob Sharp Hertfordshire County Council United Kingdom

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 been the subject for mathematical publications[5], studies on gambling behaviour [6, 7] and used to establish success in 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 (maximumstreak) 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 maximumstreak 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

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, headtrack}. 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 (2×32) 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.

ш. 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 successstreak 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%).

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	

 TABLE II.
 SUCCESS-STREAK SIZE CUMULATIVE DISTRIBUTION OF LOW PERFORMERS



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 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. 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/tobiix2-30-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.

