A computer-based selective visual attention test for first-grade school children: design, development and psychometric properties

Farzaneh Yazdani¹, Malahat Akbarfahimi², Afsoon Hassani Mehraban³
Shohreh Jalaei⁴, Mohammad Torabini-masi⁵

Received: 20 May 2014    Accepted: 27 July 2014    Published: 23 February 2015

Abstract

Background: Visual attention is known as a critical base for learning. The purpose of the present study was to design, develop and evaluate the test-retest and internal consistency reliability as well as face, content and convergent validity of the computer-based selective visual attention test (SeVAT) for healthy first-grade school children.

Methods: In the first phase of this study, the computer-based SeVAT was developed in two versions of original and parallel. Ten experts in occupational therapy helped to measure the content validity using the CVR and CVI methods. Face validity was measured through opinions collected from 10 first-grade children. The convergent validity of the test was examined using the Spearman correlation between the SeVAT and Stroop test. In addition, test-retest reliability was determined by measuring the intra-class correlation (ICC) between the original and parallel versions of the SeVAT in a single session. The internal consistency was calculated by Cronbach's alpha coefficients. Sixty first grade children (30 girls/30 boys) participated in this study.

Results: The developed test was found to have good content and face validity. The SeVAT showed an excellent test-retest reliability (ICC= 0.778, p<0.001) and internal consistency (Cronbach's Alpha of original and parallel tests were 0.857 and 0.831, respectively). SeVAT and Stroop test demonstrated a positive correlation upon the convergent validity testing.

Conclusion: Our results suggested an acceptable reliability and validity for the computer-based SeVAT in the assessment of selective attention in children. Further research may warrant the differential validity of such a test in other age groups and neuro-cognitively disordered populations.

Keywords: Children, Computer-based, Test, Reliability, Validity, Attention.


Introduction

Based on a definition from the psychologist and philosopher Williams James, attention is “taking possession of the mind in clear and vivid form, of one out of what may seem several simultaneously possible objects or trains of thought s… It implies with withdrawal from some things in order to deal effectively with others” (1). Along these lines, visual attention as a cognitive property is known to provide a critical base for learning (2), working memory (3), self-regulation (4) and word learning (5). According to multiple models, attention has

References:

1. MSc of Occupational Therapy, Department of Occupational Therapy, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran. Farzane.yazdani@yahoo.com
2. (Corresponding author) Assistant Professor, Department of Occupational Therapy, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran. Akbarfahimi.m@iums.ac.ir
3. Associate Professor, Department of Occupational Therapy, Rehabilitation Research Center, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran. mehraban.a@iums.ac.ir
4. Assistant Professor, Department of Physical Therapy, Tehran University of Medical Sciences, Tehran, Iran. jalaiesh@hotmail.com
5. Assistant Professor, Department of Neuroscience, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran. torabini-masi@sums.ac.ir
Selective visual attention test for children

different classifications (6). One of the most important types of attention with a significant role in learning and enhancing school achievement is selective or focused attention (7). Selective attention is defined as a preferential allocation of limited processing resources to events that have become behaviorally relevant (8) and depends on working memory capacity (9, 10). Overall, sustained selective attention has an important role in academic performance (11, 12) such as reading efficiency (13-15) and mathematical skills (16). According to Posner and Rothbart, the stimulation of brain networks which involve in attention mechanism could enhance this skill in early childhood (17). Therefore, it is worthwhile to investigate the possible effect of interventions in this critical period. Results underscore the need for research on the course of attention problems and the necessity to test interventions on children’s early attention problems and their effects on subsequent academic achievement (18).

Different tools are available to assess various types of attention. Some currently validated neuropsychological tests for attention in children include the $d^2$ test (19), Test of Everyday Attention for Children (TEA-Ch) (20), Stroop (21) and Trail Making Test (part B) (22) which allow measuring the visual selective attention, and there is also the Persian version of the Sustained Auditory Attention Capacity Test (23) which is designed to assess auditory sustain attention. Since recently, there exist some computerized instruments to test cognitive skills such as attention which provide abundant proof. These instruments include Test of Variables of Attention (TOVA) (24, 25), the Integrated Visual and Auditory Continuous Performance Test (IVA plus) (26) and the Connors’ Continuous Performance Test (CPT) (27) which assesses sustained attention. Such computer-based assessments have two major benefits: first, they can be used to score the tests promptly, they are able to keep proper record of reaction time and accurate responses and can also generate an interpretive profile based on the normative data and provide concurrent stimuli (28); and second, they seem to be quite interesting for children and can increase their motivation to have more cooperation and participation during the evaluation (29). Although a variety of computer-based assessments for auditory selective/sustained attention, such as Paced Auditory Serial Addition Test (PASAT) (30) and Test of Sustained Selective Attention (TOSSA) (31), are available, it seems there is a limitation in computer-based assessment of visual selective attention.

The computer-based version of the Stroop test is currently available. Stroop Interference Test was originally developed as a paper-based tool to measure selective attention and cognitive flexibility (32), ability to set shifting (33), inhibition (34) and extinction (35). The original Stroop has been translated into numerous languages such as Chinese, Czechoslovakian, German, Hebrew, Swedish, Japanese (35), Spanish (36) and Persian (37). There are a variation of Stroop tasks i.e., word-color and number-color. Some limitations of this test are that the tool may not be administered to illiterate subjects such as preschoolers and is not specific for selective attention as it assesses set shifting and executive function as well. While attention difficulties are widespread among preschool children (7), development of a tool to evaluate the selective visual attention which fits into children’s conditions and requirements seems crucial.

To achieve this goal and upon designing the selective attention test complying with children’s requirements, four important factors were considered as essential and they are as follows: 1) When children are illiterate, using the pictorial stimuli is the preferred method as elementary school children can grasp visualized concepts more easily and they can reply to picture-based questions as instantly as written concepts (38); 2) Tasks were preferred to be designed in a game format because games are considered the most significant means of communication in elementary school chil-

http://mjiri.iums.ac.ir
The objectives of the current study were to design and develop a computer-based visual selective attention test (SeVAT) with stimuli requiring the attention level in daily life, and secondly, to evaluate the psychometric properties including content, face and convergent validity and test-retest and internal consistency reliability of the developed test. The test was designed as a game in order to engage children in a play and prevent them from feeling bored.

Methods
The study involved the assessment of original and parallel forms of the Computer-based Selective Visual Attention Test and was carried out in two phases:

- Phase 1: The Development of Computer-Based Selective Visual Attention Test (SeVAT): The parallel form of the Computer-based SeVAT was designed to improve the test reliability and eliminate the carry over and learned effect. The reliability for the parallel test was measured by comparing two different tests with similar but not fully equal content. This was basically done by creating a bank of questions or stimuli which measure the same quality after which stimuli are randomly divided into two separate tests. Both tests will then be simultaneously administered to the same subjects. This is considered a common method to assess the test reliability. Absent or insufficient consistency in each test (for instance difference in question items or the content) or measurement error diminishes the test reliability (40).

To create the preliminary design of this test, all variables involved in the computer-based selective attention test were taken into account. This was based on the available evidence and inputs provided from our experts panel which has been described in our previous study (41). The first phase of this study (the development of SeVAT) is outlined below.

**Stimuli Selection:** Based on our previous record, 20 experts from various related disciplines including cognitive science, rehabilitation and computer-based game designer (9 occupational therapists, two cognitive neuroscientists, two psychologists, six computer game designers and one pediatric psychologist) validated the selected stimuli (n = 200) for both the original and parallel versions. The stimuli were selected from a picture bank consisting of 600 pictures in different categories including clothes, familiar cartoons characters (which are popular in media), fruits, foods, animals, toys, geometric shapes and signs, letters and numbers (42-44). Based on the consensus reached by the experts panel, pictures from clothes, familiar cartoons characters, fruits, foods, animals and toys categories were selected. These pictures were shown to children while they were being asked to decide whether the pictures were attractive, familiar and simple to recognize. Results demonstrated that the clothing subgroup was the least attractive (97% agreements) while the food and fruit subgroups (100% agreements) were the most attractive items. Eventually, the test contained 5 subgroups which were attractive to children and acceptable based on the view points of the experts.

**Task:** The test comprised of 20 trials (10 right-10 left visual fields in order to assess...
the visual field effect on attention allocation). In each trial, five pictures of each subtest were framed. They were shown in the column of the right or left side of the screen, and one of them was taken as the target. In each trial, first, the target picture was largely displayed at the middle of the screen and was then paused for 1500 milliseconds based on the experts' opinion (41). Next, the target picture shrunk to small size and moved above a column with five cells in the right (or left) side of the screen and stayed there till the end of trial presentation (in order to decrease the memory effect). Each cell of the columns was differently colored. Then, five pictures moved horizontally from the left side of the screen towards this column. Children were expected to click on the color opposite to the target picture. After the first 10 replications, the position of the column and direction of the pictures were changed in order to compare the asymmetric effect of the attention network.

Software: The software codes were defined in Action script 3 and the test was executed in Adobe Flash Professional CS 5, Adobe Air 3, Adobe Photoshop CS 5.5, Sublime Text 2 and Notepad++ 9.8. The animation platform was designed using the Green sock Tween Engine software. The attribute of this test were reaction time in millisecond, accuracy and false reaction.

- Phase 2: The Psychometric Properties of Computer-Based SeVAT

Content Validity: Ten experts (occupational therapists) who did not take part in the first stage endorsed the appropriateness of the quality and quantity of each trial of the test. The inclusion criteria were as follows: 1) being the author of at least one relevant article and 2) acquiring 10-years academic work experience as a faculty member. The experts assessed the amount of essentiality of each trial in the test. The content validity ratio (CVR) and content validity index (CVI) were calculated based on the Lawshe method (45). Responses from experts were pooled and the number indicated the "essentiality" for each item.

Face Validity: After valid at ing each stimulus, ten children (7 years old) who did not take part in the first stage assessed the degree of attractiveness, the presentation time and degree of simplicity of each trial. Besides, they identified how much each task was interesting using the visual analog scale (VAS) where 0 indicated boring and 10 indicated very interesting. In addition, once the test was over, all participants (60 children) answered the open question of "Do you want to play it again?"

Convergent Validity: Sixty first-grade students were enrolled to measure the test-retest reliability and convergent validity of the test. The participants were all selected through convenience sampling method from two elementary schools located in the 1st district of Tehran during April-May 2013. Participants were 30 girls with the mean age of 80.83 months and standard deviation of 3.26 and 33 boys with the mean age of 81.50 months and standard deviation of 2.58. The inclusion criteria were age 78-84 months years old, normal visual acuity and hearing, no color blindness confirmed by the Ishihara test, normal visual field documented by visual confrontation test performed by an occupational therapist, normal intelligence score IQ ≥ 90 based on Raven Intelligence Questionnaire for Children and fluency in reading color names ability. The participants were excluded from the study if they revealed any history of a neurological disorder, loss of consciousness due to head injury, any medical condition that might affect cerebral functioning and epilepsy based on their medical record and interview with their parent(s). Lack of ADHD signs was confirmed using the teacher version of Conners' Rating Scales for children in elementary schools (46) which was confirmed by their teachers. Three boys were excluded from the study due to fever and lack of proper cooperation.

In order to evaluate the convergent validity, the correlation between the average of reaction time and accuracy of Computer-based SeVAT and the Persian version of
Computerized Stroop Color-Word Test was determined. The Persian version of Computerized Stroop Color-Word Test (Ravan-Sina Inc, Iran) includes two stages. In the first stage, the training phase, the participant should choose the color of the circle which is shown on the monitor in four possible colors of blue, red, yellow and green and press the keys which are covered with colorful labels (V (blue), B (red), N (yellow), M (green)] on the keyboard. The score of this stage has no influence on the final result. The main part of the test consists of 96 colorful words - 48 colorful congruent words (the meaning of the word complies with the ink color in which the word is written) and 48 colorful incongruent words (the meaning of the word does not comply with the ink color in which the word is written) - which was displayed in pseudo-randomly order on the middle of the monitor for 2000 milliseconds (ms) with 800 ms inter-stimulus interval (ISI). The participants were asked to identify the color of the words regardless of their meaning (47). The Persian version of the Stroop test has good validity and reliability (48).

Reliability: To examine the test-retest reliability, the parallel version was designed and all the 60 children were administered the original and parallel versions in a single session.

Procedure: All children participating in this study completed the two versions (original and parallel) of the computer-based SeVAT and the Persian version of Computerized Stroop Color-Word Test in a random order. All children had a snack before the experiment with 5-10 minutes resting time between the tests. Participants were seated comfortably on a chair in a quiet room at their schools in the morning during 8-12 am. The distance from monitor was 50 cm. The participants were initially briefed about the overall procedure (by training in practice block) and clicked on the correct picture using the mouse. The ethical protocol of this study was based on the approval from the Ethic Committee of Iran University of Medical sciences (IUMS) and was signed by all the participants and one of their parents. Written informed consent was received prior to enrollment.

Statistical Analyses
Result of the Kolomogorov-Smirnov test determined the non-normal distribution of the Stroop test data and normal distribution of the original and parallel versions of Se-VAT. As such, the Spearman correlation coefficients, intra-class Correlation Coefficient (ICC) and Cronbach’s alpha coefficients were used to examine the convergent validity, test-retest reliability and internal consistency, respectively. The analysis was done in SPSS 17.0. (The Statistical Package for the Social Sciences). Throughout the experiment, p<0.05 was considered significant. The reliability correlation coefficients less than 0.4, between 0.4 and 0.7 and more than 0.7 were considered as weak, tolerable to fine and great reliability, respectively (49).

Results
Content Validity: The content validity ratio (CVR) of the SeVAT’s 20 trials was determined by 10 occupational therapists. While the experts marked 19 trials as "essential", the CVR for all trials was1 except for one trial (animals' pictures CVR=0.8). The content validity index (CVI) was 0.995.

Face Validity: Based on the inputs col-

<table>
<thead>
<tr>
<th>Computer-based SeVAT</th>
<th>Original Version</th>
<th>Parallel Version</th>
<th>ICC</th>
</tr>
</thead>
<tbody>
<tr>
<td>CN</td>
<td>SEM</td>
<td>SD</td>
<td>M</td>
</tr>
<tr>
<td>CN</td>
<td>1.92</td>
<td>5.09</td>
<td>10.08</td>
</tr>
</tbody>
</table>

SEM: Standard Error of Measurement, SD: Standard Deviation, M: Mean, UL: Upper Limit, LL: Lower Limit, CN: Number of correct answers

Table 1. Reliability of Test-Reseeating using the ICC (n=60)
selected from 10 first-grade participants, the test was perceived to be attractive (n=10, 100%), the time duration for each trial was sufficient (n=8, 80%) and the task was simple (n=8, 80%). Besides, they acknowledged that the task was interesting (mean VAS= 9.5). The answers of all the participants to the question "Do you want to play it again?" were yes.

**Test-Retest Reliability**: With regards to the test-retest reliability using the ICC, there was a correlation between the original and parallel version of SeVAT (p<0.001, r=0.778) (Table 1).

**Internal Consistency**: The internal consistencies (Cronbach's alpha) of scores in the original and parallel tests were 0.857 and 0.831, respectively. Cronbach's alpha of each subgroup revealed that no item needed to be deleted, since all were less than the total score (Table 2).

The correlation between the subgroups and the total scores in both versions are demonstrated in Table 2. There was a significant relationship between the subgroups and the total scores in the original (0.759< r <0.818) as well as the parallel (0.714< r <0.802) versions of the SeVAT.

The Correlation between the subgroups and the total score for the original version of SeVAT was shown to be tolerable to fine (0.434< r <0.601), and it was weak to tolerable (0.394< r <0.575) for the parallel version of SeVAT.

**Convergent Validity**: The Spearman's correlation coefficient demonstrated a significant positive relationship between the number of correct answers and reaction time to correct answers in the original/parallel versions of SeVAT with the congruent/incongruent stimuli of the Stroop test (Table 3).

Repeated-measure ANOVA demonstrated neither a significant main/interaction effects of the right and left visual field in the original (F (1, 53) = 0.003, p= 0.956) and parallel (F (1, 57) = 0.379, p= 0.541) tests on the reaction time to correct answers nor on the number of correct answer of these two versions of the computer-based original (F (1, 57) = 0.970, p= 0.329) and parallel (F (1, 59) = 0.506, p= 0.480) SeVAT.
Discussion

The aim of this study was to develop and validate an instrument, which is interesting to children, to test their selective attention. To do so, a computer-based SeVAT was designed and developed. Following the administration of the test to first-grade children, the validity (content, face and convergent) and reliability (test-retest and internal consistency) of the test was assessed.

There are several models of attention for adults and fewer for children (50). One limitation of applying the attentional models on younger children was the greater overlap with other developing skills; for instance, executive function, language, visuospatial skills (51). Therefore, attentional tests in preschools may be influenced by development of other skills. Hence, the design of the current test was based on the perceptual matching tasks method in which children watched the target pictures in the monitor and chose the picture which was the same as the target picture among the 5 other pictures. In other words, the participants were asked to track the target picture among the other stimuli which were moving from the right to the left and vice versa (based on the visual field) with the task being somehow similar to visual tracking. Basically, these two methods are similar to other tasks for selective attention testing as described by Mahone and Schneider in 2012. They introduced “perceptual matching tasks, central- incidental learning tasks and visual search tasks” as common methods to measure the selective attention (7).

All the selected stimuli in each trial were chromatic pictures in warm colors. Hayakawa and colleagues recorded the responses of 111 children to colorful pictures to be more rapid than the gray scale regardless of the content as compared to the adults (52). The stimuli were selected from familiar pictures for children. Task designers believed that illiterate children could grasp concepts in pictures faster than letters, numbers and words as confirmed by evidence (53) as well, and they could understand concepts in familiar pictures more quickly than the unfamiliar ones.

The target stimuli in the SeVAT was shown at the middle part of the monitor and in a larger size than other stimuli and then became equal to others in the SeVAT based on the experts’ opinion which was recorded in our previous study (41). This might have rooted in the fact that concentration and fixation in visual field relies on the center and larger objects are processed in shorter period of time which is in agreement with the results found by Yao and colleagues (2011). They declared that categorization of the complex natural images may occur within a limited area of the visual field, referred to as "field of attention" (FA). The FA is limited to 20° x 24° of visual field within almost 0.1 second without eye movement. They recorded the accuracy rate of more than 90% in FA (54). As such, the target stimuli in the Stroop test are in the middle of the visual field (48).

In the SeVAT, each of the 5 stimuli in the trial moves horizontally in a straight line from the right to the left and vice versa on the monitor. Then children should track the stimulus which is found as same as the target. Chosing this design for SeVAT was done based on the agreements of the experts panel (41). Firstly, the visual tracking is considered as an essential skill for reading (55). Secondly, based on Galera et al. (56) and Maunsell et al. (57) studies, moving objects and orientation changes may attract more attention capacity. Based on the Lieberman et al. study, moving objects are more interesting for children and cause more excitement (58).

The hierarchical visual perception processing which begins with eye movement is followed by visual attention and is completed with visual memory (59). As stated by Burnham et al. (2014), parallel with load theory, the visual and spatial working memory may influence selective attention (9). Consequently, test designers are suggested to reduce the impact of working memory by selecting 5 stimuli as distracters since normal working memory capacity for
adult is 7 ± 2, or by letting the target stimulus remain displayed on the screen until the end of the replication. This finding is consistent with that of Tricket al. studies. They showed that the responses of children in the Multiple-Object Tracking Task (parts tracking and recognizing the moving position) were affected by their age (60). Thus, 6-year-old children can track the positions of more than four moving items because of their working memory capacity (61, 62).

The computer-based SeVAT measures selective attention in the right and the left visual fields. No significant difference was found between the two visual fields with regards to the reaction time and the number of correct answers when the two versions were compared. Such finding was not in line with that of Michael and colleagues in 2005. They found a significant difference in responses from left vs. right visual fields and confirmed the existence of hemispheric asymmetry in selective attention and concluded that stimuli similarity may play a critical role in this asymmetry (63). Although such result should be validated by examining the test in ADHD or patients with unilateral neglect, some possible reasons for this disparity may be small sample size, unrestricted head movement during the test, lack of sensitivity of the test (for which some electrophysiology modalities such as ERP or eye tracking are preferred) and presenting 10 trials for each right or left visual field orderly instead of randomly; to justify this, conducting further research is suggested.

**Validity:** One of the most important factors in developing tests is their validity. Regarding the content validity stages of the computer-based SeVAT, the value of both versions of the test was approved by the experts panel. According to the experts’ analysis, the content validity of the whole test, stimuli, psychophysics properties and the homogeneity of computer-based SeVAT subgroups were quite favorable for assessing selective attention.

With regards to the face validity, the values found were considerably high. All children found this test to be interesting and they wanted to play it again. However, they indicated that the Stroop test was boring for them and made them feel as anxious as at the time of the examination.

With respect to the convergent validity, there was a positive and significant correlation between raw scores in the original and parallel versions of the SeVAT test and Stroop test. Therefore, the SeVAT may be a proper tool to measure selective attention in 7 year-old children.

**Reliability:** Both versions of the computer-based SeVAT showed a great internal consistency. This means that there were very strong relationships between different subgroups of the SeVAT with the total scores. In addition, the different subgroups of the computer-based SeVAT were strongly interrelated, indicating a good internal consistency for the test. Meanwhile, exclusion of each subgroup decreased the internal consistency of the scale which confirmed the SeVAT measuring a specific area.

Concerning the test-retest reliability of the computer-based SeVAT, findings on the repeatability of the scores, using the parallel version to examine test retest in a single session, demonstrated a great relationship. This suggested that the original and parallel versions of the computer-based SeVAT may be conducted interchangeably.

Further studies should be conducted to assess the differentiated validity of the SeVAT in other age groups of preschool children and in other disorders such as ADHD, learning disorders and PDD. However, the main limitation of this study was the lack of other computer-based visual selective attention tests to measure convergent validity in children, and expectedly lack of relevant evidence to compare the results. The sample size of this study was small and narrow. Moreover, lack of any alternative test rather than the Stroop test made us to wait till the end of the first grade in which kids acquire an acceptable literacy level. This was considered a practical limitation.
Conclusion
The computer-based Selective Visual Attention Test is an easily administered instrument to assess selective attention in children who were not literate. Considering the good validity and reliability of the SeVAT, it can be used as a test besides other children’s cognitive assessment toolbox. Children were found to perceive it as “a game they like to play” and they demonstrated very good cooperation during the test. The applicability of such a test alone needs to be examined in other age groups (3-7 year-old kids) and in different subgroups with neuro-developmental predicaments.

Acknowledgements
This paper was a part of a Master’s degree thesis of the first author in Occupational Therapy in Iran University of Medical Sciences (IUMS) entitled “Design and Development of Computer-based Selective and Sustain Attention Test for First-Grade Kids: A Validity and Reliability Assessment”. The study was supported by the Research Deputy of Iran University of Medical Sciences. We are grateful to all the experts and children for their participation in this study.

References
20. Manly T, Robertson, I.H, Anderson V, &
Selective visual attention test for children


45. Lawshe CH. A quantitative approach to content validity personnel psychology 1975; 28:563-75.


