© Journal of the Indian Academy of Applied Psychology,
January 2008, Vol. 34, No.1, 127-136.

Assessing Validity of Web-Based Computer Adaptive Training Modules

Debdulal Dutta Roy
Indian Statistical Institute, Kolkata

With rapid change in web technology, training method gradually shifts from manual to web based computer adaptive training. Here, each trainee is not exposed to same stimuli. Stimulus exposure depends upon individual differences in proficiency level accounted by the computer. Therefore, usual validity estimating statistical tools (ANOVA with repeated measurement) are not applicable. Aim of the current research is to examine the criterion related validity of one web based computer adaptive training program (Fast ForWord Language Gateway Edition Product). The program is based on the adaptive training techniques such as frequency, reward, intensity and motivation allowed for more rapid learning. It includes seven modules. All modules were classified into two as sound (Circus sequence, old McDonald flying farm, Phoneme identification) and Word (phonic match, phonic words, language comprehension builder and block commander) exercises. This internet based training program helps children in building oral language comprehension and in other critical skills necessary for learning to read or becoming a better reader. 25 randomly selected students of grades four and five were trained with Fast ForWord program for more than 20 days. Results of each individual trainee by treatments and by module in the form of percentage of success were downloaded from the computer. In considering the typical data structure, Box whisker plot analysis was made to compare the results across treatments or trials. Several parameters of box plots (Outliers, box size, location of median in the box, location of upper and lower whisker, location of hinges, and fluctuation of box size after achieving the target) were identified as useful to estimate the validity of training modules. It provides insight about different psychological changes occurred during the different training programs. Box plot results suggest possible ability contamination in changing performance level. Therefore, in order to control contamination of ability on validity assessment, high and low ability groups were identified based on initial training performance. Learning graph profiles of both groups were compared finally. Results identified very poor validity in two modules out of seven. Finally role of box whisker plot analysis on validity assessment was discussed.

With the rapid change in computer network system, attention is gradually shifting from manual mode of training to web based training exercises. Web based training is useful as the programme can be imparted to large number of trainees at a time across different countries of the globe with minimum cost. It replicates manual operation on a scale which is beyond the capability of human organization. Besides, it maintains vast storage facilities to record each operation of trainees which may be useful in cross country

comparison.

Web based training focused on three things – (a) computer simulation technique (b) computer adaptive problem assignment and (c) computer adaptive feedback system.

Computer simulation

Computer simulation, which have for years been used by the military and airlines, are increasingly finding their way into profession such as teaching, policing, sales and other fields that depend more on interpersonal skills than teaching proficiency. This provides trainees to enter into the virtual world spontaneously. In computer simulation, attention is paid to imitation of some real thing, state of affairs or process. It is the modeling of natural systems or human systems in order to gain insight into their functioning. Simulation can be used to show the eventual effects of alternative conditions and course of actions.

Computer adaptive problem assignment

Computer adaptive methodologies have the potential to accommodate individual differences within required time limits to avoid instructional failures, ensure minimal proficiency outcome and the opportunity for maximizing the outcome proficiency for each individual. Current advances in conceptual modeling and in artificial intelligence technologies may provide the basic components of machine-based instruction that adapts to and evolves with the individual. This instructional systems reduce training costs by avoiding unnecessary planning and management activities and by reducing elimination and setback rates. They can increase force readiness by producing more proficient practitioners and by returning subject matter expert instructional staff to operational positions.

Computer adaptive feedback

Computer systems can be programmed to provide performance feedback to the learner on a variety of reinforcement schedules. Feedback timing and frequency can be selected that is most appropriate to the training session. For example, immediate feedback is used for perceptual-motor tasks and at the early stages for other learning. However, as mastery is gained, immediate feedback can be disruptive and a longer interval is more appropriate. Computers can also be programmed to provide positive reinforcement which enhances learning, retention, and motivation. Control over the pace of learning can be provided for the trainee, as well as, infinite opportunities for material review and practice—until mastery.

To make training module more trustworthy, verification and validation are of crucial importance. In training, validity refers to what the training wants to achieve or the outcome. Usually validity of training is assessed by comparing pre and during or post training performance. But availability of same students in both pre and post training is difficult issue. Other technique is comparison among different success rates across trials. Purpose of the current study was to determine validity of one web based computer adaptive training product.

Validity

Test validity concerns what the test measures and how well it does so. It tells us what can be inferred from test scores (Anastasi, 1990). Following the analogy of test validity, validity of training programme may be defined as what the training wants to develop and how well it develops. This definition indicates assessment of criterion related validity. That is the level of difficulty and how well trainees improve their performance with successive treatments or trials.

Method

Participants

Participants in this study were 25 students of one English medium school. Here English is the first language. Since, all the instructions

of the exercises are in English, it was assumed that students could follow the instruction of the program properly. School authorities agreed to fulfill the set of criteria for selection of school. The criteria were: (i) Arrangement of broadband system, (ii) Arrangement of continuous internet access to the 4 computers and the server, (iii) Arrangement of experimental lab, (iv) Providing 4 computers and server, (v) Providing infrastructure facilities and (vi) Arrangement of boys at the age of 9 to 11 years for the training program.

Random sampling was followed in selection of sample. Initially, the list of names of students for the age group of 10 to 11 years was prepared. Students of even number in the list were treated as samples for this study. Therefore, each student had equal chance to be selected for the training program.

Training module

Fast ForWord® is a CD-ROM and Internet-based training program that helps children rapidly build oral language comprehension and other critical skills necessary for learning to read or becoming a better reader. Fast ForWord® evolved from the work of noted research scientists Michael Merzenich and Bill Jenkins from the University of California at San Francisco, and Paula Tallal and Steven Miller, experts on the neurological basis of language at Rutgers University. The product is based on the adaptive training techniques such as frequency, reward, intensity and motivation allow for more rapid learning. It includes seven modules. All modules were classified into two as sound (Circus sequence, old McDonald flying farm) and word (phonic match, phonic words, language comprehension builder and block commander) exercises (Table 1).

Table 1 Description of exercises

Modules/Exe rcises	Tasks	Benefits
Sound Exercises Circus Sequence (CS)	Students hear a series of short, nonverbal tones. Each tone represents a different fragment of the frequency spectrum used in spoken language. Students are asked to differentiate between these tones	It helps to improve listening accuracy.
Old MacDonald's Flying Farm (OM)	Students use the computer mouse to catch and hold a flying animal. The animal repeats a single syllable several times, and students must release the animal when they hear a change in the syllable	It helps to improve phonological awareness, listening accuracy and working memory skills.

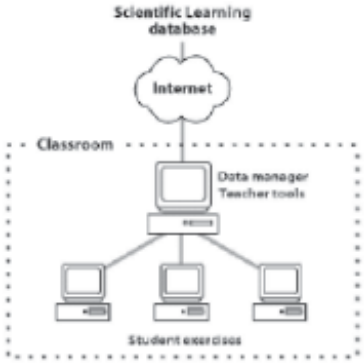
Phoneme Identification (PI)	First, students listen as one animal character utters a phoneme, and then two new animals utter similar phonemes. The students identify which of the latter two sounds was identical to the first phoneme	It helps to improve phonological awareness skills, listening accuracy, and working memory.
Word Exercises Phonic Match (PM)	Students choose a square on a grid and hear a sound or word. Each sound or word has a match somewhere within the grid. The goal is to find each square's match and clear the grid.	It helps to improve working memory, listening accuracy, phonological awareness and auditory word recognition.
Phonic Words (PW)	Students see two pictures representing two similar words that differ only by initial or final consonant ("tack" versus "tag"). When students hear the word representing one of the pictures, they	It helps to improve phonological awareness, listening accuracy, and auditory word recognition.

Procedure

After sample selection, parents meeting were made to establish rapport with the parents and the students. They were informed of the specific requirements of the product, i.e., non stop participation into the programme for 40 days (5 days X 8 weeks) minimum. Each student will be allowed to attend at least 3 programs per day. When the students arrived, trained teacher asked them to sit on the specific chair and turned on the computer. Figure 1 shows Computer network system used for collection of data.

One black screen was placed around each computer in order to control attention to unwanted things. The room was airy, well illuminated and no one was allowed to make any sound. Each student was given instruction through the computer. They were asked to raise help stand when they needed. After completion of the training, the results were downloaded from the progress tracker. Progress tracker is an on-line service to download the results of each trainee. The total arrangement for collection of data was made by the trainers of the Step-one foundation in Kolkata.

Figure 1 Computer network system



Analysis of data

Since the downloaded data were in percentage in stead of scores, box plot analysis was used to assess the validity of the training programme.

Box Whisker plot

The boxplot was invented in 1977 by American statistician John Tuckey. A box plot

or box and whiskers plot is a graphical way of representing the salient features of a distribution. It can be used with either Gaussian or non-Gaussian distributions. The box plot shows a rectangle stretching from the first to the third quartile of the distribution, these quartiles, the edges of the box, are called “hinges”. The box displays in a pictoral fashion the variability in the data. A line inside the box shows the approximate position of the median. From the median, one can determine the central tendency or location. From the length of box, one can determine spread, or variability of observation. If the median is not in the middle of the box the distribution is skewed. The further the median is from the middle, the more skewed is the distribution. If the median is closer to the bottom of the box than to the top, the data are positively skewed. If the median is closer to the top of the box than to the bottom, the distribution is negatively skewed. Cases with values that are between 1.5 and 3 box lengths from the upper or lower edge of the box are called outliers and are designated with circle (Pyzdek, 1976).

Results

Results of the study were presented into two ways – box plot analysis and comparison of high and low ability groups in means and SDs across treatments.

BOX-PLOT analysis

Outliers

No outliers were noted in any box plots (Figures 2-8), therefore, all data were accounted for analysis.

Box size reduction

Box sizes were systematically reduced across trials in case of CS (Figure 2), OM(Figure 3), and PM (Figure 5),LCB (Figure 7) suggesting gradual increase in proficiency level by trials in corresponding modules. In case of BC (Figure 8), box sizes were large and were similar for few trials after crossing reference point later on the sizes were reduced suggesting plateau stage in proficiency before achieving the target.

Location of median

Good training provides systematic change in performance. It should not be too difficult or very easy. Out of 7 modules, median always was below the 50% efficiency level or reference point in case of PI (Figure 4) exercise. And in case of PW (Figure 6), median was above the reference point at the end of total trials. This suggests high difficulty experienced by the students in those 2 exercises. On the other hand, students experienced less difficulty in solving problems of 3 exercises namely - BC(Figure 8), OM(Figure 3) and PM (Figure 5).

Length of whisker

Length of whisker was large after the lower hinge of box in case of CS(Figure 2), OM (Figure 3),PM (Figure 5),LCB (Figure 7) suggesting failure of few students to show success in those exercises like most of the students. However, this was small in case of BC (Figure 8) exercise suggesting less difficulty experienced by all most all of the trainees. In case of PI (Figure 4), length of whisker above the upper hinge of the box was high suggesting better performance of few students in comparison with most of the students. But no one reached at the target as no whisker reached at the 100% success

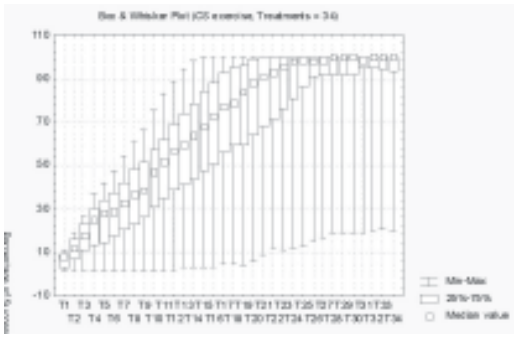


Figure 2 Box and Whisker Plot of CS exercise

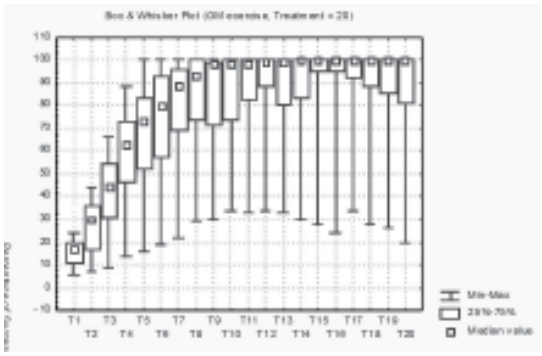


Figure 3 Box and Whisker Plot of OM exercise

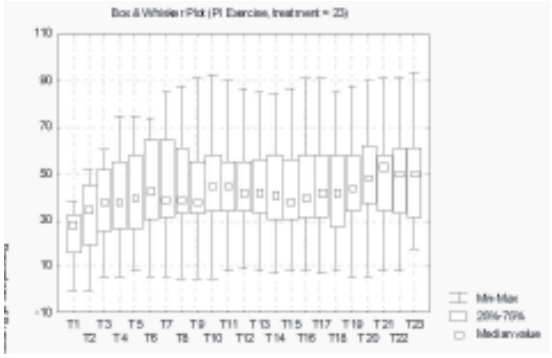


Figure 6 Box and Whisker Plot of PW exercise

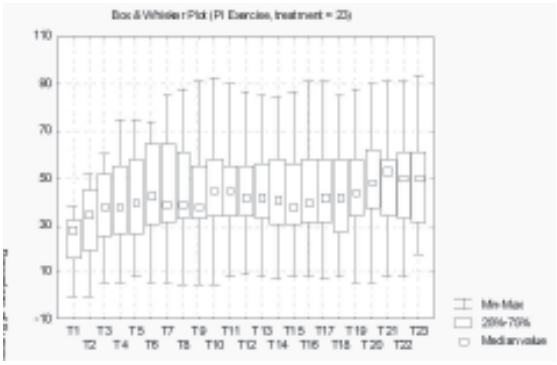


Figure 4 Box and Whisker Plot of PI exercise

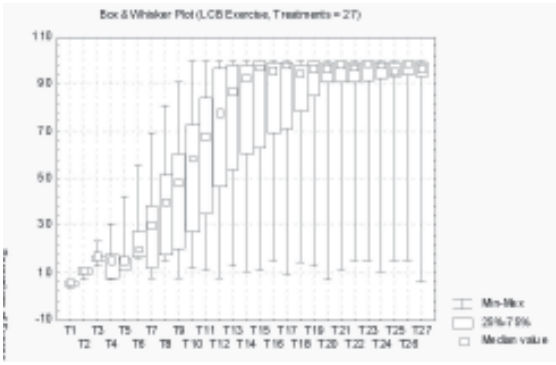


Figure 7 Box and Whisker Plot of LCB exercise

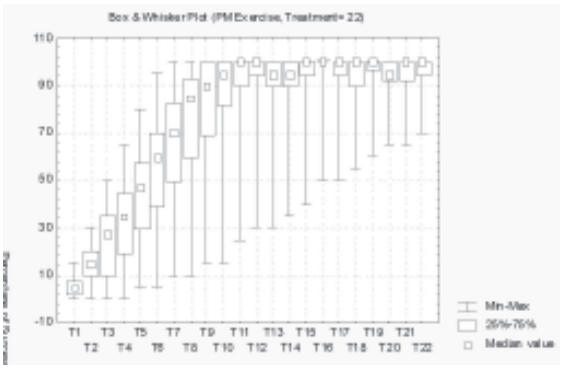


Figure 5 Box and Whisker Plot of PM exercise

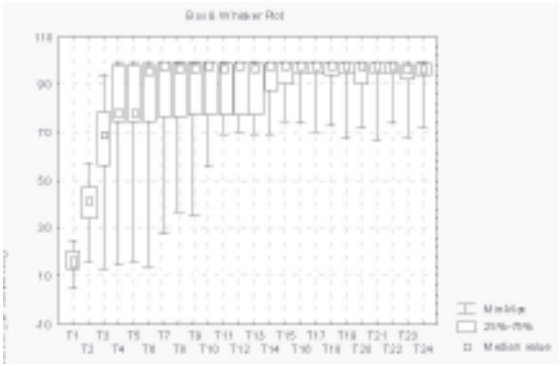


Figure 8 Box and Whisker Plot of BC exercise

level. Again, length of whisker below the lower hinge of box was high suggesting very poor performance of few students in comparison with most of the students in PI exercise. In case of PW(Figure 6), few whiskers above the upper hinge of box reached the target suggesting excellent performance of few students. Like, PI, whiskers below the lower hinge of box did not lie below the 10% success level suggesting relatively less difficulty experienced by the few students in solving problems of PW module.

Fluctuation in box size

After achieving the target, sizes of the box varied suggesting possible concentration difficulty in solving problems of same module. This was noted in OM (Figure 3)and PM (Figure 5) exercises.

Differences between high and low ability groups

Variation in different parameters of box plot analysis revealed possible within group variability in performance of the trainees under study. Therefore, it may be assumed that individual difference in ability plays critical role in changing responses from one level to another. This will lead to criteria contamination. In considering such limitation, high and low ability groups were compared using means and SD wise differences.

Mean differences

Individual difference in ability plays critical role in changing responses from one level to another. Box plot provides insight about distribution of data irrespective of ability wise differences. Figures 9, 11,13,15,17, 19, and 21 represent graphical distribution of average performance across trials by high and low ability groups. In all the exercises, mean difference between the groups was noted suggesting appropriateness about the categorization. At the initial stage, mean difference was high and gradually it was reduced. However, this reduction was high in case of PM exercise suggesting better proficiency among the low ability group with training. In case of PW exercise, the difference was high at the later treatment suggesting experience of difficulty by the students of low ability group in development of proficiency.

SD wise differences

Mean differences provide average distribution pattern but it fails to provide insight about pattern of errors committed by individuals. Figures 10, 12, 14, 16, 18, 20, 22 provide insight about within group variability or the pattern of errors experienced by each ability group across treatments.

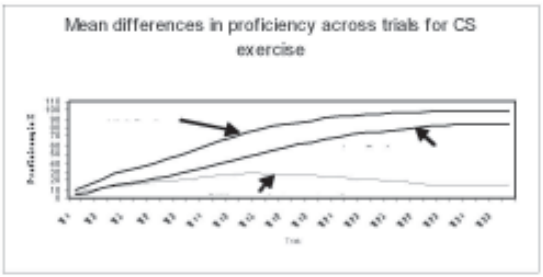


Figure 9 Mean differences in proficiency across trials for CS exercise

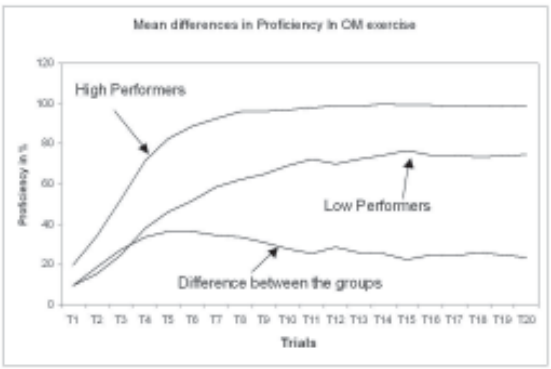


Figure 11 Mean differences in proficiency across trials for OM exercise

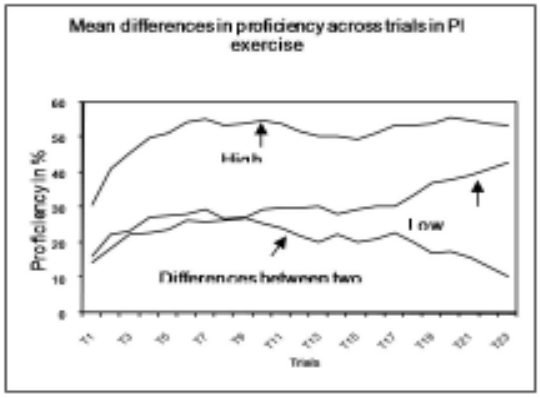


Figure 13 Mean differences in proficiency across trials for PI exercise

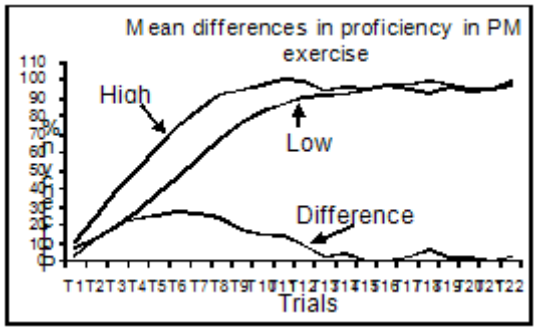


Figure 15 Mean differences in proficiency across trials for PM exercise

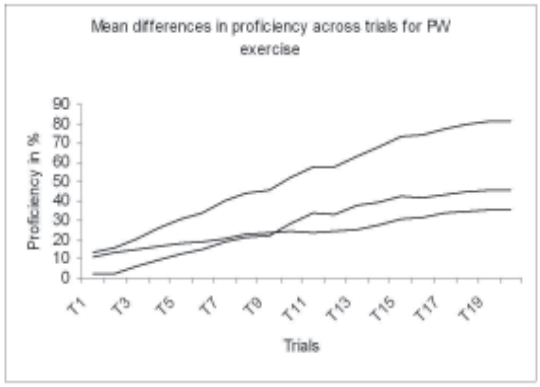


Figure 17 Mean differences in proficiency across trials for PW exercise

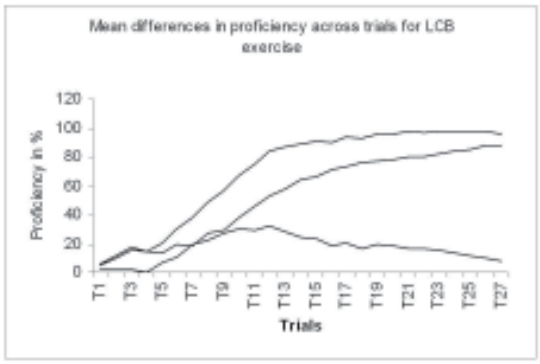


Figure 19 Mean differences in proficiency across trials for LCB exercise

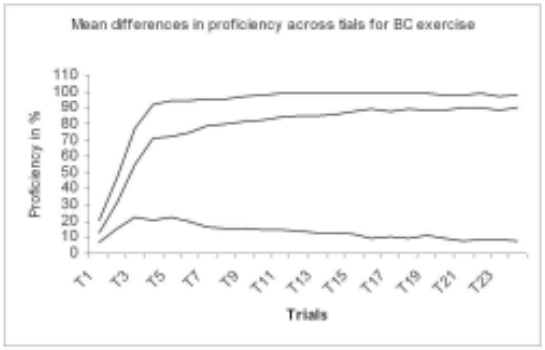


Figure 21 Mean differences in proficiency across trials for BC exercise

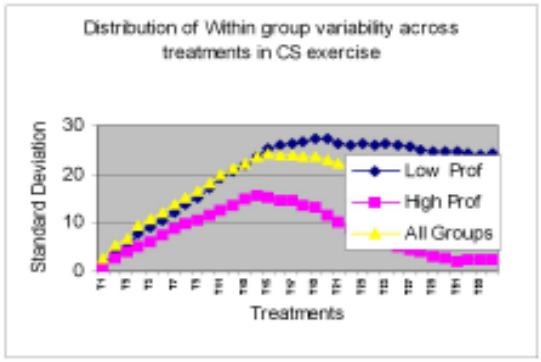


Figure 10 Within group variability wise differences in proficiency across trials for CS exercise

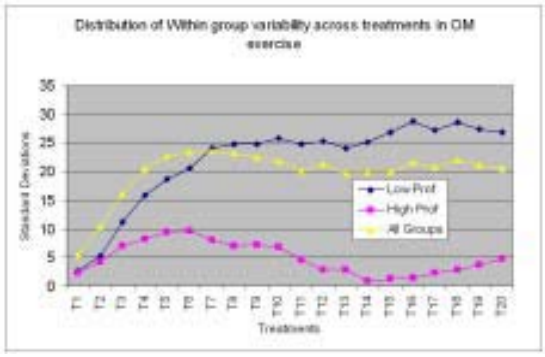


Figure 12 Within group variability wise differences in proficiency across trials for OM exercise

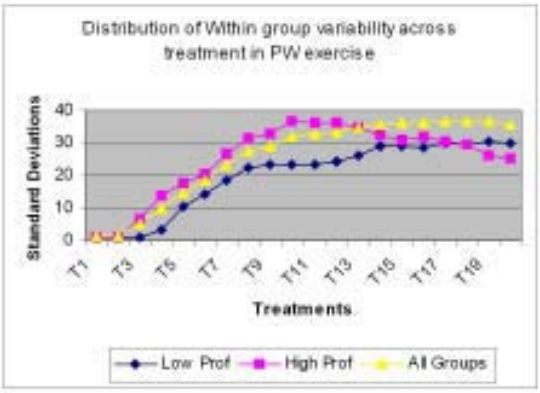


Figure 18 Within group variability wise differences in proficiency across trials for PW exercise

Figure 14 Within group variability wise differences in proficiency across trials for PI exercise

Figure 20 Within group variability wise differences in proficiency across trials for LCB exercise

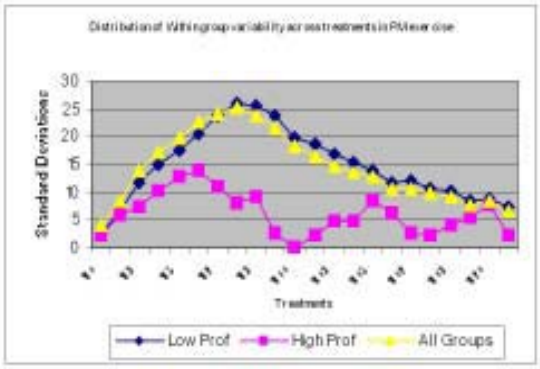


Figure 16 Within group variability wise differences in proficiency across trials for PM exercise

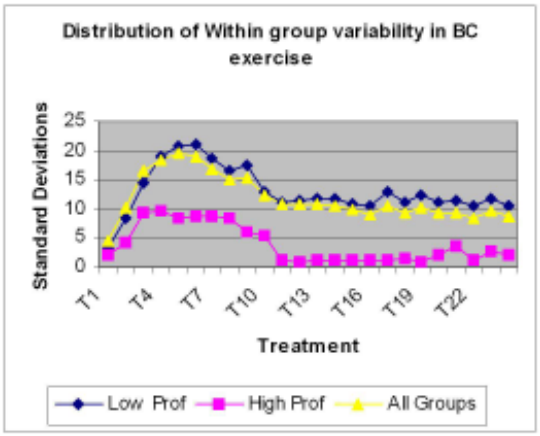


Figure 22 Within group variability wise differences in proficiency across trials for BC exercise

Figures 14 (PI) and 18(PW) show that SDs of both high and low ability group remained almost similar across trials suggesting more difficulty experienced by the students of two groups.

Discussion

Results noted that box-plot analysis is useful in analyzing various parameters of web based computer adaptive training module. By using box plot current study explored pattern of increase in proficiency level, plateau stage in proficiency, level of difficulty, target or criteria achievement, possible concentration difficulty in different training modules. Out of seven modules, two modules were identified as least valid in achieving the criteria. These are PI and PW. Further more different parameters of box plot analysis suggest possible within group variability across trials due to individual differences in initial ability level. Finally mean and SD wise line graphs were plotted for high and low ability groups and identified committing more errors by both high and low ability group in case of above two exercises. By scrutinizing the contents of these two modules, it was noted that both are meant for developing proficiency in western pronunciation and accent of English words. Though the participants of the study were English speaking but they could not follow the English accent used in the module as revealed in the results.

To sum up, in considering typical data structure of the web based computer adaptive training module current study explored role of box plot analysis in estimating validity.

References

Anastasi, A. (1990). Psychological testing. New York :Maxwell Macmillan International Editions.

Tukey, J.W. (1977). Exploratory data analysis. Addison-Wesley, Reading, MA.

Pyzdek, T. (1976). The six sigma handbook. New York:McGraw-Hill.

Received: September 17, 2007
Accepted: December 3, 2007

Acknowledgement: I am grateful to Scientific Learning Corporation, USA and to Step-One Foundation, Kolkata and St. Mary's Orphanage and Day School, Kolkata for assisting me in collection of data

Debdulal Dutta Roy, PhD, Psychology Research Unit, Indian Statistical Institute, 203, B.T. Road, Kolkata – 700 108, E-mail: ddroy@isical.ac.in