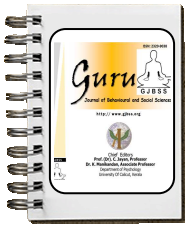
Guru Journal of Behavioral and Social Sciences
Volume 3 Issue 3 (July–Sept, 2015)
ISSN: 2320-9038 www.gjbss.org



Test Session Behaviors of Children with Academic Problems

Venkatesan, S.
Professor in Clinical Psychology, All India Institute of Speech and Hearing, Mysore, Indi

Received: 7 Aug 2015
Revised: 11 Sept 2015
Accepted: 15 Sept 2015

Keywords:

Academic Problems, Functional utilitarian, Test Session Behavior.

Abstract

This exploratory cross sectional random survey enumerates 67 types of behavioral reactions observed during psychological testing on 228 children with academic problems. The behaviors are empirically validated and grouped under speech, voice, emotions, motor and social domains. Gender, grade, and type of school emerge as significant variable in influencing during test behaviors that range from nervousness, shyness, timidity, trembling, and fear of making mistakes to teary eyes. Their general demeanor during testing is one of opposition, restless-over activity, avoidance, reluctance to speak or engage in test tasks, withdrawal, poor goal directedness, frequent seeking for reassurance or redirection which are interpreted as cues from the child's inner mental state indicating a cry for help.

Behavioral observations during testing are a routine part of every psychological report (Stapel-Wax, 2011). The systematic description of what subjects do during psychological testing is the foundation of clinical assessment. They provide objective information and insights to enable diagnosis, intervention planning and programming for the target persons on whom they are attempted. As an assessment strategy, observing test behaviors may look deceptively simple and straight forward task. Actually, they require special efforts on the part of examiners to be able to gather, scrutinize, organize, record and interpret such details (Pellegrini, Symons, & Hoch, 2004).

During psychological testing, observations typically cover what, when or how behaviors happen before attempting to derive clues on or about why it happens (Gabel, Oster, & Butnik, 1986). Formal assessments of test session behaviors in children have been carried out using standardized tools, such as, 'Test Behavior Observation Guide' (TBOG; Caldwell, 1951), 'Standard Observation for Psychological Evaluation' (SOPE; Jahoda & Goldfarb, 1957), 'Test Behavior Checklist' (TBC; Aylward & Mac Gruder, 1986), 'Guide to Assessment of Test Session Behavior' (GATSB; Daleiden, Drabman, & Benton, 2002; Glutting & Oakland, 1993), 'Test Observation Form' (TOF; McConaughty & Achenbach, 2004), and Children's Test Anxiety Scale (CTAS; Wren & Benson, 2004).

Most of these measures are Likert type scales focusing on different aspects of during test behaviors. They cover motor skills, handedness, rate of performance, activity level, attention-concentration, distractibility, impulsivity, receptive expressive language, interest, motivation, mood, fear of failure, mental set, frustration tolerance, consistency in performance, cooperation, persistence, anxiety, need for redirection, reactions to praise, encouragement or failure, self-criticism, etc. Others have used semi-structured observation formats. These scales vary by their number of items, domains, administration, scoring or rating procedures, and/or the target age groups they address. Sometimes, the item descriptions are alleged to be subjective, non-observable or non-measurable. Although, the use of such standardized tools greatly increase reliability and validity of test session observations, the time required to complete and score them has been pointed out as problematic by few examiners (Cohen & Spencer, 2009; Maller, Konold & Glutting, 1998).

That children undergo stress as test takers is well documented (Oakland, Gulek, & Glutting, 1996; Fleege et al. 1992; Perrone, 1991; Frisbie & Andrews, 1990). The test situation is likely to bring forth a variety of behavior reactions like fear, anticipatory anxiety, varying levels of confidence, agitation, aggression, or even refusal (Mantynen et al. 2001; Ounsted, Cockburn & Moar, 1983). Such during test taking behaviors have been even traced to their preterm status (Wocaldo & Rieger, 2000; Roth, Eisenberg & Sell, 1984). Despite all this, observations of test takers during assessment are highly relevant. They provide stable and clinically meaningful data about the child's behavioral tendencies (Heinonen et al. 2011). However, there are also many unanswered questions on this theme.

On their part, the child test taker is likely to have apprehensions on how it might be to undergo a psychological test. They may have questions like: Will it mean a bout of bitter pills and injections? Will it mean confronting doctors in white aprons or visiting hospitals with medical odor? Will it be a barrage of interrogations with shameful consequences for not answering them correctly? Will it mean that my test performance will be made known to my school, friends and teachers? Such assailing questions in the young test taker can be indeed a shattering experience, which may influence their test behaviors (Zatz & Chassin, 1985).

Against this scenario, it would be worthwhile to explore how children behave during test situations. It is the generic aim of this study to investigate the various types, spread, intensity and/or extensity of during test behaviors seen in contemporary children with academic problems in our country. The specific objective of this inquiry is to observe and enumerate the test behaviors across sessions of psychological assessment in children with academic problems and examine their distribution in relation to certain socio-demographic correlates.

Method

An exploratory cross sectional randomized survey design was used in this study. The key terms used in this enquiry are: 'children with academic problems' and 'test behaviors'. The term 'children with academic problems' denote referred clinical cases of students with scholastic problems from regular schools, with no apparent sensory impairments, physical, multiple or intellectual disabilities, autism, chronic health problems, missed schooling, change of school or medium of instruction, poor study habits or examination taking skills, impoverished or unsupportive educational environments, first generation learners, absent teaching, transient or long standing emotional/behavior problems, parent discord, or such other intervening factors. The term 'test behaviors' refer to 'observable measurable actions' that happen in the context of psychological testing.

The terms 'testing' and 'assessment' are also distinguished. Assessment refers to 'systematic collection, organization and interpretation of information about a person and his situations' (Sundberg and Taylor, 1962), to which is added, 'and the prediction of his behaviors in new situations' (Jones, 1970). Testing involves exposing an individual to a particular set of questions under specified and structured conditions in order to obtain a score. Testing carries a set of tasks or questions intended to elicit particular types of behaviors when presented under standard conditions and yield scores that have observable psychometric properties. In short, a psychological test is part of the larger process of psychological assessment (Cohen & Swerdik, 2002). It includes clinical interviews, use of paper-pencil or computer devices, or instruments.

Participants

The study covered 228 children with academic problems seeking diagnostic and remedial services at a national level institute which offers consultancy activities through a multi-disciplinary team by including medical and sensory screening. The institution is

reckoned as the official agency for certifying cases of children with clinical conditions, such as, mental retardation, autism, and learning disabilities. This becomes an occasion and opportunity to collect data from respondents. The children were in age range between 5-15 years (Mean Age: 10.7; SD: 4.4) and/or grades 1-10 (Mean Grade: 5; SD: 4.3) representing government (N: 52), private (N: 118) as well as partly aided (N: 58) schools pursuing their education with English (N: 154) or Kannada (N: 74) as their proclaimed medium of instruction. The period of study covered cases between 1st October, 2014 and 31st March, 2015.

Instruments

Two Instruments were prepared for this study: (a) Demographic Data Sheet; and, (b) Observation Protocol. The former was designed to obtain personal and background details of the children. The second instrument was meant to secure data on a child's behavior reactions during psychological testing. This tool was intentionally kept open-ended and semi-structured to elicit an exhaustive over-inclusive response list of during test behaviors seen in the children. The format was prepared to be handy, easy-to-tick, expand or add on notes taking, in such a manner that one could insert new entries into it while making tally ticks of existing behavior reactions and/or fresh observations if any during test situations. Both these tools were field tested on a sample of 15 children during a pilot study before their use on a fresh sample in the final study.

Procedure

Data collection involved individualized observations of each child across or during testing situations. Recording the observations was carried out immediately and contiguous to the psycho-educational testing process in the open-ended 'Observation Protocol'. Details of observations made covered included notes on how they behaved, what they commented, or otherwise during the test situation. The testing was carried out in separate places designated for that purpose and which was free from disturbances or distraction. The examiners/observers were at least post graduate degree in psychology or upwards. For example, during a test situation, if a child was noticed to 'cry', 'sullen' as well as 'silent', it was immediately recorded as such with three tallies against him or her. The scoring of all observations was carried out on all-or-none basis. The protocol carried a 'no response' item to indicate a theoretically demarcated anchor for children who showed no reactions at all during the testing situations. However, there was no on ground occasion when this category was actually used.

The testing for any given child takes place in the department sequentially on three occasions and at three levels. At the first level in 'Out Patient Section', examiners merely screen the child through a brief case history format and carry out a present status examination. In the second level at 'Detailed Assessment Section', often over 2-3 sittings, the same child undergoes an in-depth individualized testing on various parameters typically covering intelligence, adaptive behavior, academic achievement, grade level performance, adjustment, aptitude, interest, and/or problem behavior mapping. This is followed in the next level by a confirmatory final testing by the consultant not below the rank of senior faculty in clinical psychology. This authentication exercise is carried out individually through a period of at least half an hour along with the test protocols already in place. The commonly used psychological tests included administration of an intelligence scale (such as, Malin's Intelligence Scale for Indian Children, Raven's Progressive Matrices, Gesell's Drawing Test, or Seguin Form Board), adaptive behavior measures (such as, Vineland Social Maturity Scale, Indian Adaptation), achievement tests (such as, Graded Math, Reading and Spelling List), etc. All these procedures covering no less than 3-5 hours per child were carried out in the physical presence of parents, caregivers or escorts who accompany the child. The escorts were firmly instructed to remain merely as passive observers to provide moral supports and NOT to offer any verbal or non-

verbal assistance to the child during the test situations. Informed consent was taken and response anonymity was assured as mandated (Venkatesan, 2009).

All testing was undertaken in amiable, non-threatening and reassuring milieu. The testing rooms had adequate lighting, furniture, temperature, and the ready-to-use test materials in place. Pre-testing preparations typically covered rapport building, orienting the child and parent escort on what is in store during test sessions. The implication was conveyed that they were not being deemed as 'crazy'. It was assured that 'failed' answers would not end up into their being 'taken to task'. They had to simply put up their best performance. Rapport was given prime importance. Time breaks were provided when required. Simple and positively stated instructions like 'Please listen to what I say', or 'please begin when I say start!' were preferred to direct commands like 'Sit Down!' Small courtesies and praise statements were profusely used, such as, 'Good Work!', 'Great Attempt!', 'Thank you!'

Records on test observations and findings were frequently perused and exchanged between the examiners and the consultant. To determine the nature or extent of agreement or disagreement between observers on or about the child's behavior reactions during testing, Fliess Kappa for multiple observers (contrasting Cohen's Kappa applicable only for two raters) (Fliess, 1981; Fliess & Cohen, 1973) was used. All analysis was done on SPSS/PC (Sarma, 2010).

Results

The core data generated in this study was a corpus of 67 types of during test session behavior reactions as manifested through observation of children with academic problems by examiners. The observed and listed behavioral data was categorized into five heads: (a) Speech; (b) Voice; (c) Emotional; (d) Motor; and, (e) Social respectively.

For the **overall** sample (N: 228), nearly 2150 test behavior reactions are recorded (Mean: 9.43; SD: 2.11). Most of these reactions fall in the domain of 'social' (N: 774 out of 2150; 36.0%). This is followed by 'emotional' (N: 531 out of 2150; 24.7%), 'motor' (N: 506 out of 2150; 23.5%), 'speech' (N: 222 out of 2150; 10.3%), and 'voice' (N: 117 out of 2150; 5.4%) related test behavior reactions (Table 1).

Table 1

Summarized Frequency distribution of during test session behaviors

Sno.	*Behavioral Reactions*	*N*	*Gender**		*Grade/Class***			*School****		*Curriculum*****		
			Boys	*Girls*	*1-4*	*5-7*	*8-10*	*Govt*	*Pvt.*	*Aided*	*State*	*Central*
	N	*228*	*118*	*110*	*72*	*76*	*80*	*52*	*118*	*58*	*132*	*96*
A	*Speech*	*222*	*134*	*88*	*64*	*88*	*70*	*42*	*121*	*59*	*101*	*121*
B	*Voice*	*117*	*70*	*47*	*19*	*42*	*56*	*25*	*53*	*39*	*63*	*54*
C	*Emotional*	*531*	*257*	*274*	*119*	*199*	*213*	*105*	*257*	*169*	*275*	*256*
D	*Motor*	*506*	*276*	*230*	*156*	*208*	*142*	*109*	*229*	*168*	*255*	*251*
E	*Social*	*774*	*420*	*354*	*281*	*296*	*197*	*135*	*446*	*193*	*397*	*377*
	TOTAL	*2150*	*1157*	*993*	*639*	*833*	*678*	*416*	*1106*	*628*	*1091*	*1059*
	Mean	*9.43*	*9.81*	*9.03*	*8.88*	*10.96*	*8.48*	*8.00*	*9.37*	*10.83*	*8.27*	*11.03*
	SD	*2.11*	*1.85*	*2.01*	*1.77*	*2.58*	*2.18*	*2.14*	*3.41*	*3.13*	*2.11*	*3.04*

*X^2: 12.0; df: 4; p: 0.018;**X^2: 62.8; df: 8; p: 0.001; ***X^2: 24.8; df: 8; p: 0.002; ****X^2: 3.25; df: 4; p: 0.517;

In terms of gender variable, boys in this sample appear to show greater number of during test session behavior reactions (N: 118; Mean: 9.81; SD: 1.85) than girls (N: 110; Mean: 9.03; SD: 2.01). The boys have greater frequency and variety of such behaviors affecting their

'speech-voice', 'motor' and 'social' domains while the girls are found to be show more 'emotional' reactions during the test situations (X^2: 12.0; df: 4; p: 0.018) (Table 1). More specifically, the during test behavior reactions of girls include 'shyness', 'timidity', 'crying or teary eyes', 'nervous, fidgety and trembling'. Contrast this with boys, who show high count of behaviors like 'remaining silent, reluctant or refusing to talk/answer', 'stuttering', 'speaking in hoarse voice', 'showing angry outbursts', 'giggling, laughing or joking inappropriately', 'showing tics or startles', 'being restless or overactive', 'getting off seat frequently', 'being sullen', or 'not waiting to answer'.

In relation to **grades**, the children between classes 5-7 emerge as the most affected group (N: 76; Mean: 10.96; SD: 2.58) compared to those between classes 1-4 (N: 72; Mean: 8.88; SD: 1.77) and those between 8-10 classes (N: 80; Mean: 8.48; SD: 2.18)(X^2: 62.0; df: 8; p: 0.001). Likewise, students from 'aided' schools (N: 58; Mean: 10.83; SD: 3.13) show higher during test behavior reactions than those from 'private' (N: 118; Mean: 9.37; SD: 3.41) and 'government' (N: 52; Mean: 8.00; SD: 2.14) institutions (X^2: 24.8; df: 8; p: 0.002). By the same count, the type of curriculum studied by the tested children (either 'state' or 'central' syllabus) does not emerge as a significant variable to influence the frequency or type of their during test behaviors (X^2: 3.25; df: 4; p: 0.517) (Table 1).

More specifically, young children in their primary grades were observed to 'talk too much', 'show difficulty in following instructions', 'pose counter questions', 'frequently getting off seat or away from the task at hand', 'slouching or lying down on the chair', 'throwing things', 'clinging to their mother for approval or an answer', 'seeking repeated assurance from examiner', 'wanting to quit' and so on. The children in secondary grades were noticed to be 'hesitant, shy or withdrawn', 'stutter', 'speak in low voice', 'cry', 'tremble', 'sit on the edge of the seat', 'bite finger nails', 'chew writing instruments', 'make unnecessary mouth movements while writing', 'stare blankly', or 'avoid eye contact'. The high school students were found to be 'anxious or apprehensive about making mistakes', 'appeared confused or perplexed', showed a 'sullen, stubborn or irritable countenance', 'apathy or amotivation', 'were drumming on the table', or occasionally even 'boasting about oneself'.

An **item-wise analysis** (Table 2) across behavior domains show that 'fidgets or nervous' (C12; N: 88 out of 228; 38.40%) is the most common during test session behavior reaction among children with academic problems followed by 'shy or timid' (C13; N: 71 out of 228; 31.14%), 'trembles' (D23; N: 67 out of 228; 29.39%), 'fear of making mistakes' (C14; N: 64 out of 228; 28.07%), and 'cries or teary eyes' (C15; N: 63 out of 228; 27.63%). Even though less in frequency, reactions like falling 'silent' or 'being reluctant, refusing to talk/answer', showing 'verbal or body tics and startles', and/or 'seeking repeated assurance from examiner' are also recorded in about a fifth of the children in this study. Despite observations in literature on test behavior reactions in children, such as, their use of bizarre or unusual language, producing odd noises, attempting to cheat, telling lies, trying exaggerations or making up things, seeking to leave the room or swearing/using abusive language, these responses were only minimally observed in the present sample (Mantynen et al., 2001).

Table 2
Frequency distribution of during test session behaviors

Sno.	***Behavioral Reactions***	***N***
	N	***228***
A	***Speech***	
1	*Silent, reluctant or refuses to talk/answer*	*55*
2	*Stutter*	*49*
3	*Shows difficulty in following instructions*	*33*
4	*Talks too much*	*28*
5	*Uses verbal stereotypes: 'mmm…', 'don't know' or 'I know'*	*24*
6	*Posing counter questions*	*22*
7	*Uses bizarre unusual language*	*11*
	Sub-Total	***222***
B	***Voice***	
8	*Low voice*	*48*
9	*Hoarse voice*	*31*
10	*Croaky voice*	*27*
11	*Makes odd noises*	*11*
	Sub-Total	***117***
C	***Emotional***	
12	*Fidgets or nervous*	*88*
13	*Shy or timid*	*71*
14	*Shows fear of making mistakes*	*64*
15	*Cries or teary eyes*	*63*
16	*Shy, hesitant or withdrawn won't open mouth with outsiders*	*54*
17	*Appears confused or perplexed*	*44*
18	*Angry outbursts*	*38*
19	*Excess giggling or laughing*	*32*
20	*Stubborn, sullen and irritable*	*31*
21	*Repeated coughing as during recitation of tables*	*24*
22	*Jokes inappropriately or too much*	*22*
	Sub-Total	***531***
D	***Motor***	
23	*Trembles*	*67*
24	*Verbal or body tics or startles*	*51*
25	*Does not sit still, restless or overactive*	*49*
26	*Sits on edge of seat*	*47*
27	*Plays with own fingers, chair or table*	*46*
28	*Frequently off seat or off task*	*44*
29	*Bites fingernails*	*41*
30	*Chews or sucks things that are not edible*	*37*
31	*Uses eraser or crosses out frequently*	*34*
32	*Makes mouth movements while writing or drawing*	*26*
33	*Slouches or lies down in chair*	*24*
34	*Picks or scratches nose or other parts of body*	*22*
35	*Throws things*	*18*
	Sub-Total	***506***

E	***Social***	
36	*Seeks repeated assurance from examiner*	*51*
37	*Clings or turns to mother after each question to seek approval*	*47*
38	*Does not listen to instructions fully*	*47*
39	*Delayed RT or take time to answer*	*40*
40	*Does not wait to answer*	*41*
41	*Sullen, somber or serious countenance*	*41*
42	*Needs frequent coaxing to continue or complete work*	*38*
43	*Needs repetitive instructions*	*36*
44	*Poor or avoids eye contact*	*36*
45	*Stares blankly*	*28*
46	*Wants to quit*	*28*
47	*Complains feeling sleepy, headache, or dizziness during session*	*25*
48	*Acts young for age*	*25*
49	*Argues/blames task as not taught or out off syllabus*	*23*
50	*Shows apathy, unmotivated*	*22*
51	*Complains tasks are too difficult, boring or too easy*	*21*
52	*Avoid or evade entry into testing room*	*21*
53	*Does not pay attention to questions asked*	*18*
54	*Demanding feedback on performance*	*18*
55	*Yawns*	*18*
56	*Messy work*	*17*
57	*Appears impatient*	*17*
58	*Shows defiance, sarcasm or talks back*	*16*
59	*Contradicts own statements*	*15*
60	*Gets lost in thoughts or day dreams*	*14*
61	*Boasts, brags or pretends overconfident*	*14*
62	*Makes demands to be met immediately*	*11*
63	*Attempts to cheat*	*11*
64	*Attempts to leave room often under some pretext (toilet, water)*	*9*
65	*Exaggerates or makes up things*	*9*
66	*Tells lies*	*9*
67	*Swears or uses abusive language*	*8*
	Sub-Total	***774***
	TOTAL	***2150***

Inter Observer Reliability Estimates

The tripartite inter-observer agreement as measured by Fliess Kappa was 0.93. This is interpreted as 'almost perfect agreement' (Landis & Koch, 1977). Content validity or item total percentage agreement consistently measured a range of values between 90-93 percent between the identified list of major domain categories earmarked in this study, thereby indicating high internal consistency (p:<0.01). Face validity is also found to be high for the classification of major domains and its items covering test behaviors of the addressed children.

Discussion

Psychological testing is a new and growing phenomenon for school children in India (Shyam & Khan, 2009). With the promulgation of free and compulsory school education for children in recent times, many first generation learners as well as their teachers in particular and the education system in general is only now coming to grips with newer concepts like

psychological testing for their intelligence, interests, achievement or aptitude (Kapur, 2011). The emphasis on child rights, provision of equal opportunities and barrier free inclusions for students with special needs is on the upbeat (Antony, 2013). Parents are beginning to consult psychologists for academic problems in their children albeit reluctantly. The earlier stigma generally attached to such professional help seeking is decreasing. There are reasons to believe that they delay, deny or defend against such consultations for fear that their child will be nick-named. That the parents escorting their child for psychological consultation display a variety of proactive and negative behavioral reactions themselves while testing of their children is in progress is documented in a contiguous study (Venkatesan, 2013).

The value of during test session behavior reactions is historically acknowledged (Glutting, Oakland, & McDermott, 1989; Glutting & McDermott, 1988) as needed clinical skill to interpret the behaviors in children. However, reservations have been expressed on how much of such behaviors can be reliably generalized outside the test situations. Available research suggests that the practice needs to be engaged in with great caution (Brassard & Boehm, 2011). Bracken (1983), for example, calculated high impact of such behaviors on test scores (r=.48) although it was reported to be marginal (r=.17) in children between 6-14 years outside test session environment. Similar observations are made by other investigators (Konold, Maller, & Glutting, 1998; Glutting et al., 1996).

During test behavior observations of children can help determine the frequency, strength and spread of the targeted behavior. There are ethical issues involved in the observation process. Pains must be taken to ensure that behaviors of people are observed with their consent even as they are carried out without their knowledge at any given point of time. Reactivity is a major factor that can unwittingly influence the observer and observed (Hersen, 2006). Such overt observations maybe brief. Still, they can serve as poignant red flags on the covert, inner and intimate psychological status of a given child. Their poor motivation, oppositional stance, restless-over activity, disinterest, fear of failure, outward anxiety, diminished level of cooperation, passive-aggressive tendencies, task avoidance, reluctance to speak or engage in specific test tasks, withdrawal, poor goal directedness, frequent need for reassurance or redirection are all to be understood as an outward cry for help.

They are to be read as sign of malaise or a cue requiring deeper probe. Rather than viewing them as categorical 'present/absent' dichotomous behaviors, they must be viewed as falling along a situational continuum. A child can be, at the same time, pensive, defensive as well as offensive during a test situation. In mirroring the child's mental status, they may be silently indicating on the need for inquiring into home, school and/or parent-child relationships.

Summary

In sum, it may seem overly simplistic to assume that the child's test behavior is the 'cause' of his/her 'poor' or 'good' test performance. This may not be really the case. The test behaviors may be actually the 'consequence' of an undisclosed grim situation back at school, home, with peer group and/or their family. They may be also used as prognostic indicators for guessing or estimating the future status of a given child. Evidently, the brief observations obtained during testing needs to be supplemented by observation of the child's routine activities and/or interviews with parents and teachers. They cannot and need not be generalized directly to non-test situations. However, they can be definitely considered as hunches based on the examiner's perceptions which need to be followed by observations in the home, classroom, peer group or other naturalistic settings.

References

Antony, P. J. (2013). *Segregation hurts: voices of youth with disabilities and their families in India.* The Netherlands: Sense Publishers.

Aylward, G., & Mac Gruder, R. (1986). Test behavior checklist. Brandon, VT: Clinical Psychology Publishing.

Bracken, B. A. (1983). Observing the assessment behavior of preschool children. In K. D. Paget & B. A. Bracken (Eds.), *The psycho-educational assessment of preschool children* (pp. 63–79). New York: Grune and Stratton

Brassard, M. R., & Boehm, A.E. (2011). Preschool assessment: principles and practices. New York: The Guilford Press.

Caldwell, B. (1951). Test behavior observation guide. In R. Watson. (Ed.). *The clinical method in psychology.* (pp. 67-71). New York: Harper and Brothers.

Cohen, L. G., & Spencer, L. J. (2009). *Assessment of children and youth with special needs.* New Delhi: Dorling Kindersley.

Cohen, R. J., & Swerdik, M. E. (2002). Psychological testing and assessment: an introduction to tests and measurement. New York: Mc-Graw-Hill.

Daleiden, E., Drabman, R. S., & Benton, J. (2002). The guide to the assessment of test session behavior: validity in relation to cognitive testing and parent-reported behavior problems in a clinical sample. *Journal of Clinical Child and Adolescent Psychology,* 31(2): 263-271.

Fleege, P. O., Charlesworth, R., Burts, D. C., & Hart, C. H. (1992). Stress begins in kindergarten: a look at behavior during standardized testing. *Journal of Research in Childhood Education,* 7(1): 20-26.

Fleiss, J. L. (1981). *Statistical Methods for Rates and Proportions.* New York: John Wiley.

Fleiss, J. L., & Cohen, J. (1973). The equivalence of weighted kappa and the intra-class correlation coefficient as measures of reliability. *Educational and Psychological Measurement,* 33(1): 613–619.

Frisbie, D. A., & Andrews, K. (1990). Kindergarten pupil and teacher behavior during standardized achievement testing. *The Elementary School Journal.* 90(4): 435-448.

Gabel, S, Oster, G. D., & Butnik, S. M. (1986). Psychological testing in children: a guide for health professionals. New York: Plenum.

Glutting, J. J, & Mc Dermott, P. A. (1988). Generality of test session observations to kindergarteners' classroom behavior. *Journal of Abnormal Child Psychology*. 16(5): 527-537.

Glutting, J. J, & Oakland, T. (1993). Guide to the assessment of test session behavior. San Antonio, TX: The Psychological Corporation.

Glutting, J .J, Oakland, T., & McDermott, P. A. (1989). Observing child behavior during testing: constructs, validity, and situational generality. *Journal of School Psychology,* 27(2): 155-164.

Glutting, J. J., Youngstrom, E. A., Oakland, T., & Watkins, M.W. (1996). Situational specificity and generality of test behaviors for samples of normal and referred children. *School Psychology Review,* 25(1): 94-107.

Heinonen, J., Aro, T., Ahonen, T., & Poikkeus, A M. (2011). Test taking behaviors in a neuro-cognitive assessment: associations with school age outcomes in a Finnish longitudinal follow up. *Psychological Assessment*, 23(1): 184-192.

Hersen, M. (2006). *Clinician's handbook of child behavioral assessment*. California: Academic Press.

Jahoda, H., & Goldfarb, W. (1957). Use of a standard observation for the psychological evaluation of nonspeaking children. *American Journal of Orthopsychiatry*, 27: 745–753.

Jones, H. G. (1970). Principles of psychological assessment. In P. Mittler. (Ed.). *The Psychological Assessment of Mental and Physical Handicaps.* London: Tavistock.

Kapur, M. (2011). *Counseling children with psychological problems.* New Delhi: Dorling Kindersley.

Konold, T. R., Maller, S. J., & Glutting, J. J. (1998). Measurement and non measurement influences of test session behavior on individually administered measures of intelligence. *Journal of School Psychology,* 36(4): 417-432.

Landis, J. R., & Koch, G. G (1977). The measurement of observer agreement for categorical data. *Biometrics,* 33(1): 159–174.

Maller, S. J., Konold, T. R., & Glutting, J. J. (1998). WISC-III factor invariance across samples of children exhibiting appropriate and inappropriate test session behaviors. *Educational and Psychological Measurement. 58,* 3, 467-474.

Mantynen, H., Poikkeus, A. M., Ahonen, T., Aro, T., & Korkman, M. (2001). Clinical significance of test refusal among young children. *Child Neuropsychology*, 7: 241-250.

McConaughy, S. H., & Achenbach, T. M. (2004). Manual for the test observation form for ages 2-18. Burlington, VT: University of Vermont, Research Center for Children, Youth, & Families.

Oakland, T., Gulek, C., & Glutting, J. J. (1996). Children's test taking behaviors: a review of literature, case study, and research on Turkish children. *European Journal of Psychological Assessment*, 12(3): 240-246.

Ounsted, M., Cockburn, J., & Moar, V. A. (1983). Developmental assessment at four years: are there any differences between children who do, or do not, cooperate? *Archives of Disease in Childhood*, 58: 286-289.

Pellegrini, A. D., Symons, F. J., & Hoch, J. (2004). *Observing children in their natural worlds-a methodological primer*. New York: Routledge.

Perrone, V. (1991). On standardized testing. *Childhood Education*, 67(3): 131-142.

Roth, K., Eisenberg, N., & Sell, E. (1984). The relation of preterm and full-term infants temperament to test-taking behaviors and developmental status. *Infant Behavior and Development,* 7: 495-505.

Sarma, K. V. S. (2010). *Statistics made simple*. New Delhi: PHI Learning Private Limited.

Shyam, R., & Khan, A. (2009). *Clinical child psychology.* Delhi: Kalpaz.

Stapel-Wax, J. L. (2011). *Behavioral observation.* Encyclopedia of Child Behavior and Development. New York: Springer. Pp. 231-232.

Sundberg, N. D., & Tyler, L. E. (1962). *Clinical psychology: an introduction to research and practice*. New York: Appleton Century Crofts.

Venkatesan, S. (2009). *Ethical Guidelines for Bio Behavioral Research*. Mysore: All India Institute of Speech and Hearing.

Venkatesan, S. (2013). While testing behaviors and handling strategies in parents of children with academic problems. *Journal of Indian Academy of Applied Psychology*, 39(1): 26-34.

Wocaldo, C., & Rieger, I. (2000). Very preterm children who do not cooperate with assessments at three years of age: Skill differences at five years. *Journal of Developmental and Behavioral Pediatrics*, 21: 107-113.

Wren, D. G., & Benson, J. (2004). Measuring test anxiety in children: scale development and internal construct validation. *Anxiety, Stress and Coping,* 17(3): 227-240.

Zatz, S., & Chassin, l. (1985). Cognition of test-anxious children under naturalistic test-taking conditions. *Journal of Consulting and Clinical Psychology*, 53(3): 393-401.