Original Research Paper

Volume - 11 | Issue - 05 | May - 2021 | PRINT ISSN No. 2249 - 555X | DOI : 10.36106/ijar

Clinical Psychology

# SEGUIN FORM BOARD TEST: A PROPOSAL FOR APPROPRIATE NORMS

**Dr. Roopesh B Nagaraj***

M.Phil., Ph.D. Additional Professor Consultant: Child & Adolescent Mental Health Unit Dept of Clinical Psychology NIMHANS, Bengaluru-560029 *Corresponding Author

**ABSTRACT** Seguin Form Board Test (SFB) is one of the popular tests of intelligence for children in the Indian subcontinent. It is a much sought-after test due to its simplicity, ease of administration and takes less than 10 minutes to arrive at an IQ score. Further, it elicits interest in children, is easy to score and can be administered to even shy, speech impaired, autistic spectrum and/or attention deficit and hyperactive children. There are standard norms and to keep up with the Flynn effect, researchers are conducting studies to update normative data to keep up with the times. However, all the available norms in India with respect to SFB, follow incorrect practices, such as, using SFB for children older than 10 years; inappropriate processes, such as, using only the mean value and ignoring standard deviation; and erroneous way of converting interval scale score to ordinal scale score to arrive at the ratio IQ method. These errors have been existing for decades, despite the possibility of having better and appropriate norms that matches the deviation IQ. This article attempts to highlight the pitfalls of adopting the existing norms with appropriate examples and figures. Further, the article proposes a better and more appropriate norms to arrive at an IQ.

**KEYWORDS :** Seguin Form Board, Ratio IQ, Deviation IQ, Age-appropriate norms

## INTRODUCTION

Assessing intelligence has been one of the main roles of clinical and educational psychologists, especially with respect to determine mental sub-normality. To cater to this need, there are a variety of tests that use different ways and abilities to purportedly measure the same construct of intelligence. Barring few, earlier tests of intelligence rarely matched the theories of intelligence. However, the recent trend in intelligence assessment is to match a test to a particular theory of intelligence as well as to certain neuropsychological/brain functions e.g., Wechsler's four indices, instead of verbal and performance subscale (Wechsler, 2008), and Das's Pass theory of Intelligence and the corresponding intelligence test 'Cognitive Assessment System' (Nagleri and Das, 2005; Nagleri and Conway, 2009). However due to various reasons, many tests that do not have correspondence with established neuropsychological/brain functions, still find a place in the routine assessment of intelligence.

One such test is the Seguin Form Board Test (SFB; Seguin 1907) which is one of the oldest tests of intelligence that has withstood the test of time and changes in intelligence theory and assessment. SFB is popular among psychologists to assess intelligence. Some use it as an additional test along with other tests, while some others use it as a stand-alone test to arrive at an IQ. The reasons for the popularity of SFB are that it takes very less time; is easy to administer; elicits good interest/motivation from the child to participate; its ease of scoring; follows the popular method of interpreting ratio of mental age to chronological age; and it can be administered to all types of children including those who have speech impairment, where there are language incompatibility issues and across socioeconomic and educational backgrounds. Further, as this test takes very less time, the chance of children being uncooperative is less particularly in children with attention deficit hyperactivity disorder and autistic spectrum disorders. Because of all the above, SFB is considered to be one of the very few 'culture-fair tests' currently in circulation. However, like few other intelligence tests already in use in India it too has a few significant issues that need to be addressed and corrected. This article is an attempt in that direction.

**Issues and concerns**:

**1. What does SFB measure?**

Seguin Form Board is said to measure visuospatial perception, organization and discrimination; psychomotor dexterity, coordination and development; as well as motor and cognitive speed (Cattel, 1953; David and Virginia, 1972; Venkatesan, 2014). SFB is said to measure 'g' factor of intelligence, but only till the age of 8 - 10 years, after which it is said to become a test of manual dexterity instead of intelligence. Put it in other words, the correlation of SFB performance with IQ drops significantly after about 9 - 10 years. That is, after about 9 - 10 years of age it will mainly assess 'motor speed and ability' and not intelligence. In addition, there is a 'ceiling effect' for motor dexterity skills that is required for SFB does not develop much after about 9 – 10 years (David and Virginia, 1972; Spearman, 1927; Anastasi, 1961). Hence, using SFB after say, 9 to 10 years is not advisable (The current article's focus is mainly on the scoring norms, and hence detailed discussion about what SFB measures is not carried out here).

**2. Some popular norms and normative values**

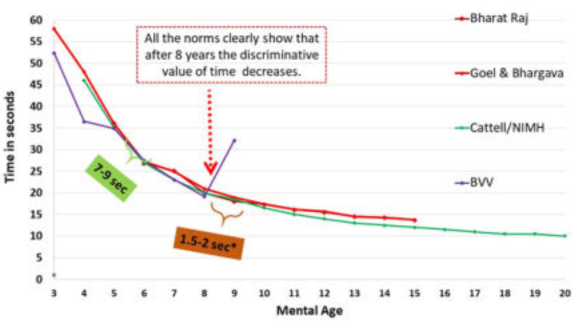
As SFB is one of the popular tests among psychologists, and due to the Flynn effect (Flynn, 1984), many researchers have attempted to keep the norms current. In this regard, the most available norms (in Mental Age) referred/used to in India are of Bharat Raj's (BR; Bharathraj, 1971) Goel and Bhargava (G & B; Goel and Bhargava, 1990), National Institute for the Mentally Handicapped, Secunderabad, India (NIMH-S; 1989), and Basavarajappa, et al., (BVV; Basavarajappa, Venkateshan, and Vidya, 2009). BR provides norms from 5 to 15, G & B provides norms from 3 to 15 years, NIMH-S from 3.5 to 20 years, whereas BVV does not provide norms per se, but provides the mean and SD values from 3 years to 10 years of Chronological Age (refer table 1).

**Table 1: Showing The Shortest Time (among The 3 Trials) Taken In Different Studies/norms**

| Age | Popularly circulated and used norms | | | BVV mean |
|---|---|---|---|---|
| | B R | G & B | Cattell/NIMH-S | |
| 3 | - | 58 | - | 52.37 |
| 4 | - | 48 | 46 | 36.51 |
| 5 | 36.1 | 36 | 35 | 34.93 |
| 6 | 26.8 | 27.2 | 27 | 27.44 |
| 7 | 25.1 | 25 | 23 | 23.04 |
| 8 | 20 | 21 | 20 | 19.15 |
| 9 | 18 | 18.9 | 18.5 | 32.18 |
| 10 | 17.3 | 17.4 | 16.5 | - |
| 11 | 16.1 | 16.2 | 15 | - |
| 12 | 15.5 | 15.7 | 14 | - |
| 13 | 14.5 | 14.5 | 13 | - |
| 14 | 14.3 | 14.2 | 12.5 | - |
| 15 | 13.7 | 13.8 | 12 | - |
| 16 | - | - | 11.5 | - |
| 17 | - | - | 11 | - |
| 18 | - | - | 10.5 | - |
| 19 | - | - | 10.5 | - |
| 20 | - | - | 10 | - |

Among the above, NIMH-S norms are exactly the same as Cattell's 1953 norms, i.e., it appears that NIMH-S (1989), has merely utilized Cattell's norms. Further, according to the available literature, Cattell restricted the norms only till 15 years (Cattell, 1953), however, NIMH-S norms goes up to 20 years. It is not sure, as to why and how this has happened and/or whether there is any basis for this. Children of every age in Cattell's study have taken lesser time, i.e., about 1 to 2 seconds, compared to BR and G & B normative values (for almost every age) (Goel and Bhargava, 1990). However, BVV results closely matches Cattell's norms, implying that the performance of the Indian children in 2009 matched the performance of the US children tested in 1953, reflecting the working of Flynn effect in Indian population (Cattell, 1953; Flynn, 1984; Basavarajappa et al., 2009). Apart from the Flynn effect gains observed in the BVV study, all other studies show almost the same *pattern* of results in terms of shortest time taken among the 3 trials (refer figure 1) indicating a universal development pattern with respect to SFB. Only the BVV study shows little difference in the time taken/pattern where the 9-10 year-old children have taken more time (Purple colored line going upwards in figure 1).

**Figure 1: Shortest time and their corresponding Mental Age in 4 different norms**




### 3. Decimal values

Almost all the above-mentioned norms (in table 1) show decimal values. This seems like the average/mean value of their respective normative group. However, when administered to an individual child, one rarely gets decimal values, and even if one gets it, there are high chances that it does not reflect the true performance. For example, not all will be adept at exactly starting the stopwatch when the child picks up the first block and stop the stopwatch exactly when the child correctly puts the last block down. Even a half second delay or a half second advancement has significant positive or negative implications for the child, whereas both may not reflect the child's ability and might actually be erroneous. To explain the above, let us assume that the examiner uses Cattell/NIMH-S norms, and that a child of 14 years takes the test and finishes it in 12.5 seconds. However, if the examiner starts the stopwatch just half-a-second late, then the child's performance would be shown as completed in 13 seconds, and the norms show the child's mental age as 13 years, which is one year less than her/his actual age. Hence, using decimal values might lead to erroneous judgement especially when the variation between two age levels is hardly one or two seconds.

### 4. Item discrimination / item results variability

To consider any test as a good test, especially an intelligence test, it must have a few characteristics, and the most important one is the fact that the scores should adequately discriminate between individuals who have

- average intelligence from individuals who have below/above average intelligence
- borderline intelligence from individuals who have mild mental retardation
- borderline intelligence from individuals who have average intelligence
- above average intelligence to superior level of intelligence, and so on

For e.g., to differentiate among the top 5 percentile points / high scores in Standard Progressive Matrices (Raven, 1938), Raven developed 'Advanced Progressive Matrices' (Raven, 1988; Raven 1991). However, this is not observed in any of the norms available on SFB in India. This can be better understood with an illustration. In the current article, figure 2 depicts the Cattell/NIMH-S norms in terms of differences in time taken (in seconds) between two adjacent age levels. As seen in figure 2, at younger ages the scores show much differentiation/variation (for e.g., between 4 to 5 years, it is approximately between 46 to 35 seconds, which is about 11 seconds difference/variation), compared to higher ages (for e.g., between 9 to 10 years, it is approximately between 18.5 to 16.5 seconds, which is about 2 seconds difference/variation).

As the variation is wider in the younger years, different psychologists adopt different scoring methods. For example, a 6 years old child's shortest time is 32 seconds. If we use Cattell/NIMH-S norms, this falls between 35 to 27 seconds, i.e., between an MA of 6 to 5 years. Here, one psychologist might consider that 31 is closer to 35 seconds, and hence the MA should be 5 years. This might reduce the child's IQ (see table 1). Another psychologist might divide the range and adjust it to the age. That is,

- Between 27 to 35 seconds, there are (range is) 8 seconds
- A year has 12 months, so 12 divided by 8 = 1.25 months
- 35 (normative value of 5 years) minus 32 (time taken by the child) = 3 seconds
- 3 x 1.25 = 3.75 months (after rounding off, we can consider it as 4 months)

So, subtracting 4 months from 6 years (or 72 months) will yield 5 years 8 months as the Mental Age of the child.

Even though the above adjustments reduce the erroneous scoring of SFB, it is still not the correct practice and can still lead to inaccurate reporting of IQ.

**Figure 2: Differences/variation in time (in seconds) - G & B norms**


Another example is, let us assume that a 11-year-old child gets the shortest time of 17 seconds. According to G & B norms, this falls between 10 to 11 years mental age (between 16.2 and 17.4). Here a genuine problem would be regarding which age to consider. Taking 10 years will reduce the child's mental age and subsequent under reporting of her/his IQ as 91. However, many psychologists overcome this issue by giving benefit of doubt to the child by considering it as 11 years so that the child will get an IQ of 100.

However, the issue is further complicated if the child takes just a few seconds extra. This can be understood with the help of a few scenarios where a delay between just one and three seconds can adversely impact the child. Cattell/NIMH-S norms are used for calculation to explain this scenario. As can be seen from table 2a, a mere one second delay in higher ages, can make one to consider an adolescent of 17 years to have dull normal IQ, and a 3 second delay in higher ages, can make one to consider the IQ to be in the lower borderline (close to mild mental retardation) range. No test or norms can/should reduce a person's IQ from normal range to borderline/mild intellectual disability range just because they took 3 seconds extra on a motor dexterity task.

In this regard, the BVV results substantiate the fact that SFB should not be administered to higher ages. That is, in the BVV study, children of 10 years took more time than the 9-year-old, despite the fact that the authors did not intend this test to be used with children aged above 9 years.

The above issue can be understood better if we compare the SFB with Binet Kamat Test (BKT; Kamat, 1934) and Vineland Social Maturity Scale (VSMS; Doll, 1953). BKT and VSMS though use 'age scale and ratio IQ', and they have intermediate values/scores between two adjacent age levels, and hence have better discriminatory power than SFB. For example, in BKT there are 6 items for each year till age 10 and VSMS has at least more than 3 items for each year till age 12. For higher ages, while there are fewer items, they still have more than one item for each year. Table 2b shows what happens when an individual fails in 1, 2 or 3 items of an age level (that is equivalent to her/his chronological age) on BKT. As compared to SFB, 'failing between 1 to 3 items', on BKT does not reduce the score and corresponding IQ of a child with average/normal intelligence (table 2b). For e.g., at 14 years, failing in 1 item (compared to their own chronological age level items) will result in an IQ of 98, failing in 2 items will result in an IQ of 95, and failing in 3 items will result in an IQ of 93. However, 98, 95 and 93 IQs are still in the normal IQ range (Kamat, 1934).

**Table 2: Shows How Mild Reductions From Age-appropriate Performance Affects The Iq On Two Different Tests Norms.**

| 2a: Seguin Form Board: Just 1 to 3 seconds delay (from age-appropriate performance) can significantly reduce the IQ (Calculation is based on Cattell/NIMH-S norms) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| CA | Norms | 1 sec delay | MA | IQ | 2 sec delay | MA | IQ | 3 sec delay | MA | IQ |
| 10 yrs | 16.5 sec | 18 sec | 9 yrs | 90 | 19 sec | 8 ½ yrs | **85** | 20 sec | 8 yrs | **80** |
| 12 yrs | 14 sec | 15 sec | 11 yrs | 92 | 16 sec | 10 yrs | **83** | 17 sec | 10 yrs | **83** |
| 14 yrs | 12.5 sec | 13 sec | 13 yrs | 93 | 14 sec | 12 yrs | **85** | 15 sec | 11 yrs | **79** |
| 17 yrs | 11 sec | 12 sec | 15 yrs | **88** | 13 sec | 13 yrs | **76** | 14 sec | 12 yrs | **71** |
| 20 yrs | 10 sec | 11 sec | 17 yrs | **85** | 12 sec | 15 yrs | **75** | 13 sec | 13 yrs | **65** |

| 2b: Binet Kamat Test: Failing in 1 to 3 items (from age-appropriate performance) **does not** significantly reduce the IQ | | | | | | |
|---|---|---|---|---|---|---|
| CA | Fail in 1 item* | | Fail in 2 items* | | Fail in 3 items* | |
| | MA | IQ | MA | IQ | MA | IQ |
| 8 yrs/96 months | 94 months | 98 | 92 months | 96 | 90 months | 94 |
| 10 yrs/120 months | 118 months | 98 | 116 months | 97 | 114 months | 95 |
| 12 yrs/144 months | 140 months | 97 | 136 months | 94 | 132 months | 92 |
| 14 yrs/168 months | 164 months | 98 | 160 months | 95 | 156 months | 93 |
| 19 yrs -> 16 yrs/192 mnths** | 222 months | 116 | 216 months | 113 | 210 months | 109 |
| 22 yrs -> 16 yrs/192 mnths** | 258 months | 134 | 252 months | 131 | 246 months | 128 |

* from their chronological age level

** failed 1/2/3 in item/s as compared to their actual age level of 19 yrs and 22 yrs and not the ceiling CA used in BKT (these two illustrations for 19 and 22 years is only for the comparison with the SFB purpose and not to be used BKT administration)

The above two types of values in table 2 (2a and 2b) clearly shows that current SFB norms significantly penalizes even for a mere one second delay. Here someone can think/say that "*why does a child show a delay in performance on SFB; every child will always perform at his/her optimum, and that the time taken on SFB represents the actual ability of the child…*". However, there can be extraneous factors that affect the child performance on any test/s. It has shown that motivation and interest alone can affect about 15 IQ points. In addition, there will always be children who perform at their own pace, despite being told to perform as fast as they can. This can be observed in children, who are shy, have social anxiety, autistic spectrum disorder, or cerebral palsy.

### 5. Point scale' to 'Ratio IQ'

Intelligence tests typically have two types of items or item scores. First is '*age scale*' scores, i.e., passing an item will result in the person getting credits in months/years. For e.g., in BKT, passing an item results in either 2 months, 4 months or 6 months credit depending on the age level of the item passed in the test. Second is '*point scale*' scores, i.e., passing an item yields a numerical raw score. For e.g., in Wechsler's tests, passing an item yields a score/s in point/s. Usually, IQ tests follow either,

- 'Age scale to Ratio IQ' format as followed in BKT (MA/CA x 100 = IQ) or VSMS (SA/CA x 100 = SQ); or
- 'Point scale to Deviation IQ' format as followed in the Wechsler's tests (Raw score to Standard score to IQ).

**'Ratio Iq' Falls Under The 'ordinal Scale Of Measurement'** And **'deviation Iq' Falls Under 'interval Scale Of Measurement'**. Generally, interval scale of measurement is considered better than an ordinal scale of measurement, and hence Wechsler's tests are considered to be better as compared to the BKT/VSMS.

The score obtained in SFB 'time taken to complete the trial', is a 'time score'. '*Time score*' matches with '*interval scale of measurement*' because time differences between any two values are equal across the scores. For e.g., the time gap between 10 to 20 minutes and the time gap between 90 to 100 minutes are the same, which is 10 minutes. Given this, rarely does one convert a point scale score to an age scale score, because point scale score is more desirable than the age scale score. However, **SFB commits this serious error. It uses a time score, which is in the point (interval measurement) scale and converts it to an age (ordinal measurement) scale, thereby obtaining mental age and ratio IQ**.

Another reason that substantiates the assertion of error is that, in BKT/VSMS one needs to add up the individual items passed. This score is then used to get the mental/social age. That is, if the child passes all the items until year 5, and then passes 5 items in year 6, and fails all items in year 7. Then the child score would be

- Passing all items till 5th year, so 5 years x 12 months = 60 months
- Passing 5 items in year 6, so 5 items x 2 months = 10 months
- 60 + 10 months = 70 months. Therefore 70 months is the Mental Age of the child.

Whereas in the available SFB norms, just one score, i.e., 'time taken' to complete a trial will directly yield a child's mental age (without adding any score).

The whole world is moving towards deviation IQ, as Ratio IQ can lead to significant errors (For a detailed description, please refer Roopesh, 2020). Psychologists in India who use BKT and VSMS follow mental/social age to ratio IQ approach and the very format of these tests makes it difficult to adapt them to get a deviation IQ / percentile score. However, with respect to SFB, one does not have to be limited, as the 'time taken to complete' score can be used to obtain deviation IQ (section 6 and section 12 discusses this issue in more detail).

### 6. Using Only 'average/mean' Values And Ignoring 'standard Deviation'.

The BR, G & B, and Cattell/NIMH-S norms consider only the 'average/mean' values to determine mental age (BVV study does not provide any norms but mentions both the mean and the SD). This is a serious mistake in intelligence/ability assessment. Every intelligence test, especially those that follow percentile point/deviation IQ follows the 'normal probability curve' distribution and consider both mean and the SD. There are relatively standard rules as to how much or what percentage of scores are considered as average IQ, and how much is considered as below or above average IQ scores and so on. For example, 68.26 % of the population usually get scores between ±1 SD, intellectual disability is considered only when the score obtained is lesser than 2 SD (less than 70 IQ on Wechsler's tests), and around ±0.66 SD (i.e., between 90 to 109 IQ in Wechsler's tests and/or between 25 to 75 percentile points in Raven's matrices) is considered as the 'normal/average ***IQ range***'. This ***range*** is determined by the mean and the SD. The same thing is observed in percentile points. Whereas, the available SFB norms in India completely ignores the SD. For a clearer understanding, table 3 shows two type of norms, one is the SFB (BR norms) and other is the sample of the Raven's Standard Progressive Matrices norms (Pune & Mumbai norms; Raven, Raven and Court, 2000).

**Table 3: Showing the different scoring systems.**

| SFB (BR norms) | | | | Raven's SPM (Pune and Mumbai norms) | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Pp | Chronological Age | | | | | | | | | | |
| MA | S Time | MA | S time | | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 |
| 5 | 36.1 | 11 | 16.1 | 95 | 39 | 44 | 46 | 49 | 52 | 53 | 54 | 55 | 56 | 56 | 55 |
| 6 | 26.8 | 12 | 15.5 | 90 | 36 | 41 | 43 | 46 | 49 | 51 | 52 | 53 | 54 | 54 | 53 |
| 7 | 25.1 | 13 | 14.5 | 75 | 31 | 34 | 37 | 41 | 45 | 47 | 48 | 49 | 50 | 50 | 49 |
| 8 | 20 | 14 | 14.3 | 50 | 19 | 21 | 28 | 33 | 39 | 41 | 43 | 44 | 45 | 45 | 44 |
| 9 | 18 | 15 | 13.8 | 25 | 13 | 13 | 17 | 22 | 30 | 33 | 36 | 38 | 39 | 39 | 37 |
| 10 | 17.3 | | | 10 | 11 | 11 | 12 | 14 | 18 | 23 | 27 | 29 | 31 | 31 | 30 |
| | | | | 5 | 10 | 10 | 11 | 12 | 14 | 17 | 20 | 24 | 23 | 26 | 25 |

S Time = Shortest time; Pp = Percentile point

In table 3, under SFB norms, it can be observed that the time a child takes to complete **'will only yield one mental age score', irrespective of the age of the child**. For example, if time taken to complete is 20 seconds, then irrespective of the child's age, a mental age is 8 years. In contrast, in the RPM norms, the score a child gets can overlap across all age ranges; and the child's score has value only after what s/he gets is checked/compared with the chronological age level norms. Further, within a particular chronological age, a child can get different scores. To explain further, in contrast to SFB, where one score will always yield one Mental Age, in the Raven's Tests, a raw score of 25 can be obtained by any child from 8 to 18 years; and hence has value only when one knows the child's chronological age to check the child's percentile point; and a child of 9 years can get any raw score from 0 to 60.

Further, in SFB in order for a child to get 'normal/average' IQ, the child's score must exactly/closely match the average score mentioned in the norms, and even a small delay of a second or two might bring down the child's IQ drastically (refer table 2). However, this is not the same in Raven's norms. Here, for a child (say 10 years old) to get a 'normal/average' intelligence, s/he should get a raw score anywhere between 17 to 37 (between 25th to 75th percentile points). Figure 3a clearly shows the differences in the distribution of 'z values' of normative scores of SFB and three other tests, i.e., Raven's SPM, Bhatia's Performance test of Intelligence (Bhatia), and WISC-IV (Raven, Court and Raven, 1988; Bhatia, 1955; Wechsler, 2003). It can be seen in figure 3a that the z score of SFB normative values does not show that the norms are normally distributed, whereas the z scores of Raven's SPM (figure 3b), Bhatia's PTI (figure 3c), and WISC-IV (figure 3d) normative values, show that they are normally distributed. That is, the '***inverted S***' pattern signifies that the distribution is normal.

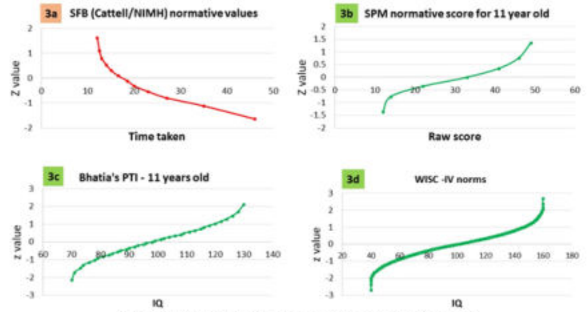



Figure 3: Showing the distribution of normative scores of different tests

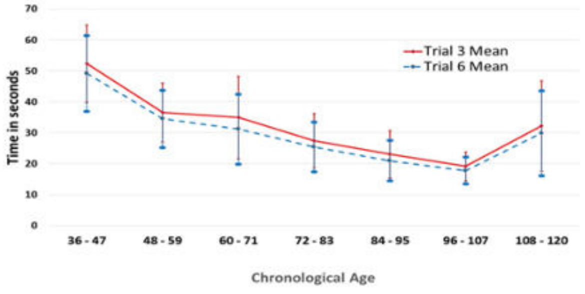
Another technical aspect one should remember is that, due to its very nature, age scale and ratio IQ values does not fit the 'inverted S' shape, when their z values are plotted. This is because they are not the values derived from standard deviation, deviation IQ and/or percentile points, which are based on normal/actual distribution of the scores. This is applicable to BKT and/or VSMS scores. However, one should recall what was mentioned above, that SFB mental age values are not actually the 'age scale' scores, but are the "actual time taken" scores and hence are 'point scale' scores. Technically therefore, point scale scores can be used to plot the z values, and if used correctly, point scale scores does yield '***inverted S***' shaped curve (refer 6d). However, it depends on how the point scale scores are used for the norms. All the above-mentioned normative values of Cattell/NIMH-S, BR and G & B norms have incorrectly used the time taken values (i.e., converted point scale scores to ratio IQs) and hence they do not adhere to the normal distribution (refer figure 6a, 6b, and 6c). Please refer to section 12 for a detailed information of how to derive appropriate norms based on 'time taken' values.

## 7. Three or more trials

In one of the studies (Basavarajappa, Venkatesan and Vidya, 2009), the authors have attempted to evaluate the optimum number of trials that takes to reach the shortest time taken, and the results showed that it is around the 6th trial. Given this, the authors suggested that the optimum number of trials to administer can be six trials. Even though this seem to have sound logic and face validity, on deeper evaluation, it clearly shows administering only 3 trials is sufficient. Administering only 3 trials saves time and saves confusion in terms of different people following different methods (number of trials).

The reason as to why only 3 trials are sufficient can be found among the principles of psychological testing of individual differences. IQ testing is a comparative method, where the IQ and the percentile point (i.e., performance of the child) are always understood in relation to where s/he stands among her/his age peers. IQ tests try to measure just that. The items in any IQ tests are arranged from most simple to complex items, thus matching the normal probability distribution. That is, few items are easy where most test takers will pass, some items are of average difficulty where about 68% of the people score within ± 1 SD, and few items are very difficult, which only a few will pass. In IQ testing, items which everybody passes and/or items which everybody fails are not included, because these items do not differentiate among people, and differentiating among people is the hallmark of IQ testing.

Figure 4: Difference between 3rd & 6th trial time (in seconds)




Given the above, the data obtained by the Basavarajappa et al., study, clearly shows that the distribution of the time taken (shortest) does not vary between 3rd trial or 6th trial, i.e., though the results shows that children have taken slightly shorter time in 6th trial compared to 3rd trial, they both show the same pattern of distribution with marginal difference in variation (refer figure 4). When the pattern of distribution of the scores are same, it does not matter whether it is the 3rd, 6th or 10th trial. Figure 4 shows that both the 3rd trial (red solid line) and 6th trial (blue dotted line) are almost identical in the pattern of distribution. The simple example given in table 4 clearly explains this phenomenon. It can be observed in the table that even though the time taken in column A is different from time taken in column D, and the time gap varies between the two (column B and E); the distribution pattern is the same. That is, if one divides each time gap values in column E by the number 2, all the values in column E matches with that of column B. This has been shown in column G, which is exactly the same as column B. Column C and F substantiate this in terms of similar/same ranks.

**Table 4: Showing The Two Time Taken Scores In Running And Their (same) Distribution Pattern**

| Pattern Pink | | | Pattern Purple | | | G |
|---|---|---|---|---|---|---|
| A | B | C | D | E | F | Column E divided by 2 |
| 100 mts run time | Time gap (in column A) | Rank (for column A) | 200 kms run time | Time gap (incolumn D) | Rank (for column D) | |
| 10 - 11 sec | 1 sec | 1 | 20 - 22 sec | 2 sec | 1 | 1 |
| 12 - 13 sec | 1 sec | 2 | 23 - 25 sec | 2 sec | 2 | 1 |
| 14 - 16 sec | 2 sec | 3 | 26 - 30 sec | 4 sec | 3 | 2 |
| 17 - 20 sec | 3 sec | 4 | 31 - 37 sec | 6 sec | 4 | 3 |
| 20 - 23 sec | 3 sec | 5 | 38 - 44 sec | 6 sec | 5 | 3 |
| 24 - 28 sec | 4 sec | 6 | 45 - 53 sec | 8 sec | 6 | 4 |

## 8. Separate Norms For Different Socioeconomic Status And Other Variables

Psychologists especially those who adapt/develop/standardize tests in India, opine and recommend that there should be separate norms for different socioeconomic class, residential (e.g., rural and urban) and educational (e.g., literate and illiterate) status (Bhatia, 1955). The reasons for this recommendation can be mainly attributed to at least four things (which are mutually related).

They are

- On many ability tests, children who are born to those with more education, high socioeconomic status, urban background, children from a developed country and/or white racial background perform better as compared to children of less educated parents, low socioeconomic status, rural background, who are from a developing country and/or non-white racial background.
- Due to the history of psychology which witnessed improper sampling procedures and/or standardization practices; and the associated criticisms of the intelligence tests not being culturally fair, especially with respect to certain race, certain population group, and lower socioeconomic strata.
- The sheer extent of differences that exists in India with respect to different sociodemographic, cultural, racial, regional, and/or economic differences.

- IQ assessment being the main 'ability' test that has a wide and long-lasting effect, which can determine/affect, either positively or negatively the child's circumstances, academics and/or future.

Being aware of the above, it is natural for any psychologist to feel that they have a moral or ethical responsibility to avoid any incorrect assessment and improper categorization/labeling of the IQ of a person. An associated thought might be that if it cannot be controlled/avoided before the assessment, then it needs to be corrected or adjusted during the scoring of the test at least. One of the ways, that the psychologist thinks of, to correct/adjust this 'perceived' injustice is by having separate norms.

Substantial number of psychologists think that the mistakes lie with the intelligence tests, be it in terms of the type of factors involved, the type of test items, appropriateness of sampling and/or the standardization process. However, psychologists should know that the IQ differences mentioned above are also observed in well designed and standardized IQ tests, such as Wechsler scales, which have significantly reduced test/item bias and have shown cross-cultural validity, at least among the Western population (Wechsler, 2003; Georgas Weiss, Van de Vijver and Saklofske, 2003). Then, if the bias is not in the intelligence tests or not in the sampling/standardization process, then how can the significant differences in IQ among different groups (as mentioned above) be explained?

It is easier to explain if one understands that IQ as an ability is not determined solely by genes/biology and that, it is not a fixed entity but one, that reflects a person's current level of functioning, with or without receiving/utilizing adequate stimulation to achieve optimal/full cognitive potential (Prifitera, Saklofske and Weiss, 2005). Wechsler explains these differences in IQ not as an indication of lower intelligence among certain groups, but as an indication of differences observed in our society, and how these variations in economic, social, medical, and political opportunities have an impact on the intellect. Based on the IQ differences observed among different groups, he suggested to make improvements in socioeconomic conditions that leads to these differences in intelligence rather than condemn IQ tests. He further states that, "the cause is elsewhere and the remedy not in denigrating or banishing the IQ but in attacking and removing the social causes that impair it" (Wechsler, 1971).

Apart from the above explanation, a psychologist must remember that, several factors significantly affect IQ, such as whether the child is reared by a single parent or both the parents; amount of time parents spend with their children; amount of money spent on the child; nutrition consumed; appropriate medical interventions; and exposure to toxins (Prifitera, Saklofske and Weiss, 2005). This does not mean that there can be separate norms for all these different factors. To understand this further, one can look at the common phenomenon that happens all around us, which is, the way in which various ability as well as achievement tests (for e.g., annual final school exams, entrance exams for specialty/professional courses) are conducted. None of these tests have separate scoring criteria for different groups/variables/factors, be it gender, socioeconomic groups, caste, religion, or race. However, considering the differences obtained on such tests (i.e., after scoring the performance) by some groups/population, various governmental/other organizations provide special consideration as a social justice. However, one should remember that this special consideration happens after the obtained results of ability/achievement tests and not before or not during scoring. The same is followed with respect to intelligence tests, where people who score less than the normal/average IQ range get different benefits (for e.g., extra time during exam, and disability benefits).

**9. Diagnosing mental retardation/intellectual disability with SFB**

After going through all the issues above, a pertinent question to ask would be whether the SFB is suitable to diagnose mental retardation / intellectual disability. The usual practice observed among the students and professionals is that when other tests are administered, the performance on the SFB will be taken as one of the corroborative findings. However, when other core intelligence tests (e.g., BKT, Raven's CPM, MISIC, or WISC) could not be administered due to various reasons (such as child having speech delay, poor motivation, being on the autistic spectrum, or in a community camp assessment setting) some students/professionals do consider only SFB results to diagnose intellectual disability. Doing so is not a good practice, due to the reasons listed below.

a. The important reason is that SFB mainly measures visuospatial perception, organization and discrimination; psychomotor dexterity, coordination and development; as well as motor and cognitive speed. However, though it is said to measure the 'g' factor till about 8 – 9 years, it does not measure the main aspects of intellect, which is abstract thinking, reasoning, and so on. Further, to diagnose intellectual disability, one has to consider other important factors, such as social and adaptive functions. This is one of the main reasons as to why Raven did not try to attempt to diagnose intellectual disability with his progressive matrices (Raven, 1941). Cole, Burkheimer, Steinberg et al. (1968), studied the scores on SFB with standard scores on Wechsler Scales of Intelligence for a sample of 172 children **with intellectual disabilities** between 6-15 years; to find that only Digit Span and Comprehension subtests showed a statistically significant correlation. Due to this they concluded that the SFB is not a valid test for **such** children.

b. As far as correlation with other tests was concerned, a study (Goel and Bhargava, 1990; Goel and Sen, 1984) was carried out to assess the suitability and applicability of SFB to intellectually disabled children (Goel and Sen, 1984). They administered Draw-a-Man Test (DAP; Pathak, 1956); Seguin Form Board, Colored Progressive Matrices (CPM, Raven 1956), Peabody Picture Vocabulary Test (PPVT; Dunn, 1959), Stanford-Binet Test (S-B; Terman and Merrill, 1937) and VSMS, on 60 children who were earlier diagnosed as having mental retardation. The results of the study showed moderate correlation within the range of 0.31 to 0.50 between SFB and the other tests. All correlations were significant, and due to this the authors claimed that the Indian SFB norms can be used to assess mental development. However, there are several errors in planning as well as interpreting the above study results. They are (i) the sample of 60 children were between the ages of 10 and 18 years 10 months, with a mean age of 14.9 years. As discussed above, higher the age, the efficacy of SFB to measure intelligence decreases; (ii) the above-mentioned various tests measure different things (i.e., drawing ability, vocabulary, social and adaptive functions, reasoning and problem solving, motor dexterity), each yielding different scores (i.e., Social Age, Percentile point, Mental Age) (iii) SFB norms used were again limited in many aspects (for e.g., using only mean/average score and not using SD); and (iv) though the correlations were significant, the magnitude of correlations (0.31 to 0.50 between SFB and other tests) and the regression results were very low, indicating that they measure different things and hence, that SFB cannot be substituted for other tests.

To explain it further, any test claiming to measure intelligence, should correlate quite highly with established intelligence tests. For example, the correlation of VSMS with Stanford Binet test is quite high around 0.8 to 0.9 (Doll, 1953). Hence, as much as possible, SFB should not be used as a sole test to diagnose intellectual disability.

**10. Few wrong practices**

**a. Arbitrarily extending the norms till old age**: Another disturbing trend has been observed among few students/professionals that they are following an age-correction norms conversion without any rationale or basis. That is, they extend the Cattell/NIMH-S norms till 74 years for e.g., considering time score that matches 15 years to match for 25 to 29 years; time score of 13 years matches to 50 – 54 years; time score of 11 years matches to 60-74 years, and so on. This is absolutely a wrong practice and should not be carried out under any circumstances.

**b. Random stacking of blocks**: Almost all psychologists follow the three-column stacking order in the placement of blocks. However, some of them randomly place the blocks within each column. The manual clearly says that the star and the cross should be at the bottom of the pile. There is a reason for the recommendation of placement of the blocks, which is, certain blocks such as the star, inherently take more time to correctly place in the right slot. Therefore, if the star is kept on the top, it might increase the time taken to complete all the blocks and if the first block itself takes more time, the child might lose motivation to carry on with the task, which might affect the assessment and/or erroneously lead to a miscalculation of the IQ of the child.

**C. Using Only Sfb To Determine Intelligence:** Though performance on SFB correlates significantly with intelligence till about 8 – 9 years, it is not advisable to depend on only this test to determine the intelligence of the child. Intelligence is a broader concept and/or involves multiple aspects/abilities. As mentioned in section 1, SFB mainly depends on motor dexterity of the child and hence it is not advisable to use only SFB as the sole test to determine the intelligence of the child. It can be one of the tests to corroborate other intelligence test findings.

## 11. Probable reasons for continuation of inappropriate practices

After going through the issues and errors listed above, one can wonder then as to why so many people and for so many decades kept on following this erroneous scoring practice. One cannot pin point the exact reason, but it can be hypothesized. The probable reason for this can be traced back to the scoring and reporting of IQs in early 20th century. As the deviation IQ method was not practiced or was not popular in the mid-20th century (before, Wechsler introduced his tests with Deviation Iqs), researchers in India might have followed their Western counterparts in using the mental age/ratio IQ method. Further, during that time, the Stanford Binet Test was more popular and it followed the mental age and ratio IQ method. Similarly, other tests, such as Draw-A-Man test and VSMS also followed the ratio method of scoring. In addition, Cattell's norms on SFB also followed the mental age and ratio method. Given this, it appears that it might have become sort of a 'mental set' or 'functional fixedness' in people to automatically consider the 'mental age and ratio IQ' scoring method for SFB; and all new researchers who tried to develop the recent / latest norms (citing Flynn effect) continued with the same 'mental set' or had functional fixedness in terms of the type of the method of IQ calculation. That is, the new researchers concentrated mainly on finding the “time taken” among the existing population, rather than concentrating on the type of scoring itself.

Other reasons might be few, such as, (i) blind belief that any published material has passed all the necessary requirements, and is hence valid; (ii) when several senior professionals/teachers follow those methods, then it might be the correct one; (iii) when the test is originally from a Western/developed country, then it would be correct; and/or (iv) though a few critically minded professionals have some concerns, they might not having enough time and/or may not know how/where to find the mistake and/or how to correct it.

## 12. Proposing new norms for SFB

Pointing out errors in scoring practices alone is not enough, and one should try to find ways to rectify the errors. In this regard, this author, has tried to derive norms to calculate IQ based on whatever limited data is available. For this, this author has used the mean and SD results mentioned in the Basavarajappa et al., article. It should be remembered that the norms arrived by this author here, is 'derived' based on only mean and SD values; and is not 'developed' based on the entire sample data from each individual subject. Hence, the norms given here can be considered to be a relatively 'crude norms', and hence, can/ should be used until more appropriate norms are developed.

The norms given here are the 'age wise norms' (refer table 5). As the Basavarajappa et al. (2009), study showed the ceiling to be at age 9 years, the normative values also stop at 8 years 11 months. Here one has to find the child's chronological age (from 3 years to 8 years 11 months) on the left first column. Then check the shortest time taken by that child. After this, the column where the child's score indicates the IQ range (i.e., Superior, Above Average, Average, Below Average, Borderline or probably Intellectually disabled).

The norms are derived based on the SD values of normal probability distribution for each IQ category (refer figure 5). For example, the “time taken” range which falls between ±0.66 SD is considered as the 'average' category. That is, if a child of 6 years 4 months takes 29 seconds (shortest time of the 3 trials), then the child is said to have 'average' intelligence.

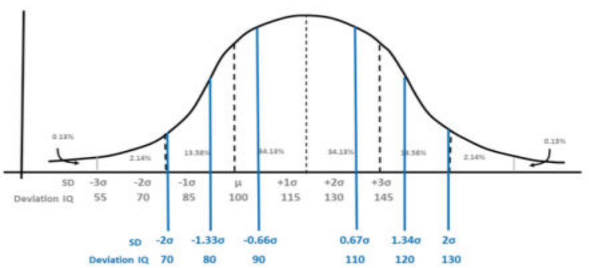



**Figure 5: NPC characteristics and the range for IQ classification, based on WHO/Wechsler criteria of m=100 and SD = 15**

**Table 5: Showing The 'derived Normative' Values To Score/ interpret The Sfb's Shortest Time By The Current Author**

| Range in SD* | 1.33 to 2 | 0.67 to 1.32 | 0.66 to -0.66 | -0.67 to -1.33 | -1.34 to -2 | < - 2** |
|---|---|---|---|---|---|---|
| Age (in months) | IQ range (time taken in seconds) | | | | | |
| | Superior | Above Avg | Average | Below Avg | Borderline | Probable IDD |
| 36 - 47 | 27 - 35 | 36 - 43 | 44 - 61 | 62 - 69 | 70 - 77 | > 77 sec |
| 48 - 59 | 17 - 23 | 24 - 29 | 30 - 43 | 44 - 49 | 50 - 56 | > 56 sec |
| 60 - 71 | 8 - 16 | 17 - 25 | 26 - 44 | 45 - 53 | 54 - 62 | > 62 sec |
| 72 - 83 | 10 - 15 | 16 - 21 | 22 - 33 | 34 - 39 | 40 - 45 | > 45 sec |
| 84 - 95 | 8 - 12 | 13 - 17 | 18 - 28 | 29 - 33 | 34 - 38 | > 38 sec |
| 96 - 107 | 10 - 12 | 13 - 15 | 16 - 22 | 23 - 25 | 26 - 28 | > 28 sec |

* The range is given for description/better understanding of the theory used to divide the scores. Overall SD ranged between -1.98 to 1.98 (as exact 2 SD was difficult to arrive at). So, this range/row can be ignored when scoring

**Pls. note: Higher time taken means poor performance (vice versa); Values are rounded off

To further highlight that this author's proposed 'derived' normative values given in table 5 adheres to the normal distribution of the norms, a curve was plotted (refer figure 6d) based on the values/norms for 4 year-old given in table 5. The plot shown in 'figure 6d' has the '***inverted S***' shape and hence can be said to adhere to the normal distribution. Further, one can note the figure 6d's similarities with the other three norms of Raven's, Bhatia's and WISC – IV norms (figure 3b, 3c and 3d respectively). This newly proposed derived appropriate norms for SFB shown in 6d, can be contrasted with the other three norms of Cattell/NIMH-S, BR and G & B norms (figure 6a, 6b and 6c respectively). It can be observed that **none** of these latter normative values follow 'inverted S' shape and hence it can be said that they do not follow the normal distribution.

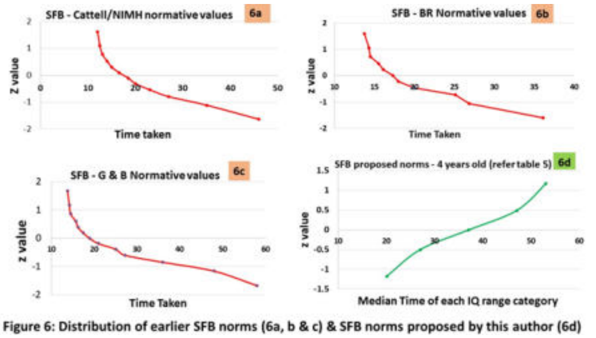



**Figure 6: Distribution of earlier SFB norms (6a, b & c) & SFB norms proposed by this author (6d)**

The figure 6 clearly shows that the new proposed norms by this author (figure 6d) is similar to the norms of Raven's, Bhatia's and WISC-IV test norms (figure 3b, c, and d); and hence more appropriate compared to the existing SFB norms (figure 6a, b and c).

## 13. Conclusion:

Changing the views, attitudes and behaviors that one has held for a long time is not easy. It is natural for anyone to question anything that is new, especially if the new thing is different from the earlier/familiar ones. However, when one encounters such a scenario, one has to be open minded to at least evaluate the pros and cons of both the earlier practices as well as the new suggestions. Science develops only by correcting earlier mistakes. Further, the main role-definition of the clinical psychologist is that of a 'scientist-practitioner'. Given this it is better to evaluate what is correct and what is not. The current article has pointed out the erroneous practices that were long used in terms of Seguin Form Board's scoring methods, and provides a better option. This has been done using simple examples, tables and graphs. Further, one has to remember that the proposed-appropriate norms given by this author, is derived based on only the mean and standard deviation values, and hence can be considered as a relatively crude method. Therefore, this can be considered only as a temporary solution until appropriate norms are developed similarly, based on the data from the entire sample of each individual subject.

## REFERENCES:

1. Anastasi, A. (1961). *Psychological Testing*. Macmillan.
2. Basavarajappa, Venkateshan, D., & Vidya, M. (2009). Normative data on Seguin Form Board test. *Indian Journal of Clinical Psychology, 35(2)*, 93-97.
3. Bharathraj, J. (1971). AIISH norms on Seguin Form Board with Indian children. *Journal of All India Institute of Speech Hearing, 2*, 117-127.

4. Bhatia, C. M. (1955). *Performance Tests of Intelligence Under Indian conditions*. Oxford University Press.
5. Cattell, R. B. (1953). A Guide to Mental Testing. In S.K. Goel & M. Bhargava. *Handbook for Seguin Form Board*. National Psychological Corporation.
6. Cole, S., Burkheimer, G. J., & Steinberg, J. (1968). Validity of Seguin Formboard with Retarded Children, *Psychological Reports, Suppl 22(3)*, 1143-1144.
7. David, R. L. & Virginia, C. S (1972). *Seguin Form Board; Technical Report 20. Disadvantaged Children and their first School Experiences*. EST Head Start longitudinal study. Technical Report Series ED0-81833.
8. Doll, E.A. (1953). *The Measurement of Social Competence: A Manual for the Vineland Social Maturity Scale*. Educational Publishers Inc.
9. Dunn, L.M. (1959). *Peabody Picture Vocabulary Test*. American Guidance Services Inc.
10. Flynn, J. R. (1984). The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*, 95, 29–51.
11. Georgas, J., Weiss, L. G., van de Vijver, F., & Saklofske, D. H. (2003). *Culture and Children's Intelligence: A Cross-cultural Analysis of the WISC-III*. Elsevier Sciences.
12. Goel, S.K. & Bhargava, M. (1990). *Handbook for Seguin Form Board*. National Psychological Corporation.
13. Goel, S. K. & Sen, A. K. (1984). *Mental Retardation and Learning*. National Psychological Corporation.
14. Kamat, V. V. (1934). A revision of the Binet scale for Indian children. *British Journal of Educational Psychology, 4(3)*: 296-309.
15. Kulshreshtha, S. K. (1971). *Hindi Adaptation of Stanford Binet Intelligence (1960) Scale*. Manas Sewa Sansthan.
16. Malin, A. J. (1965). *Vineland Social Maturity Scale: Test Items and Manual of Instructions*. Nagpur Child Guidance Centre.
17. Naglieri, J. A. & Conway, C. (2009). *The cognitive assessment system.* In J. A. Naglieri & S. Goldstein (Eds.), *Practitioner's Guide to Assessing Intelligence and Achievement (pp. 27-59).* John Wiley & Sons Inc.
18. Naglieri, J. A. & Das, J. P (2005). *Planning, attention, simultaneous, successive (PASS) theory: A revision of the concept of intelligence.* In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues*. The Guilford Press.
19. National Institute for the Mentally Handicapped. Mental Retardation (1989). *A Manual for Psychologists*. NIMH.
20. Neisser, U., Boodoo, G., Bouchard, T. J. Jr., Boykin, A. W., Brody, N., Ceci, S.J, et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51(2)*, 77–101. https://doi.org/10.1037/ 0003-066X.51.2.77
21. Phatak, P (1956). *Directions for Using Draw-A-Man Test for Indian Children*. M.S. Baroda University.
22. Prifitera, A., Saklofske, D.H., & Weiss, L.G. (2005). *WISC-IV Clinical Use and Interpretation: Scientist-Practitioner Perspectives.* Elsevier Academic Press.
23. Raven, J., Raven, J. C. & Court, J. H. (2000). *Manual for Raven's Progressive Matrices and Vocabulary Scales (Section 3)*. Pearson.
24. Raven, J. C., Court, J. H. & Raven, J. (1988). *Manual for Raven's Progressive Matrices and Vocabulary Scales (Section 4)*. H. K. Lewis.
25. Raven, J. C., Raven, J. & Court, J. H. (1991). *Manual for Raven's Progressive Matrices and Vocabulary Scales (Section 1)*. Oxford Psychologists Press.
26. Raven, J. C. (1956). *Coloured Progressive Matrices: Sets A, AB, B*. Lewis.
27. Raven, J. C. (1938). Standardization of progressive matrices. *British Journal of Medical Psychology, 19,* 137–150. https://doi.org/10.1111/j.2044-8341.1941.tb00316.x
28. Roopesh, B. N. (2020). Binet Kamat Test of Intelligence: Administration, Scoring and Interpretation –An In-Depth Appraisal. *Indian Journal of Mental Health, 7(3)*, 180-201.
29. Seguin, E. (1907). *Idiocy: Its Treatment by the Physiological Method*. Columbia University.
30. Spearman, C (1927): *The abilities of man*. Macmillan.
31. Terman, L. M. & Merrill M. A (1937). Measuring Intelligence: A guide to the Administration of the New Revised Stanford-Binet Tests of Intelligence. Houghton Mifflin.
32. Venkatesan, S. (2014). Celebrating a century on form boards with special reference to Seguin Form Board as measure of intelligence in children. *Global Journal of Interdisciplinary Social Sciences, 3(6)*, 43–51.
33. Wechsler, D. (1971). Intelligence, Definition, Theory and the IQ. In R. Cancro (Ed.), *Intelligence: Genetic and environmental influences*. Gruene and Stratton.
34. Wechsler, D. (2003). *Wechsler Intelligence Scale for Children – IV. Technical and Interpretive Manual. San Antonio*. The Psychological Corporation.
35. Wechsler, D. (2008). *Wechsler Adult Intelligence Scale -IV*. *The Psychological Corporation.*