Reliability of Physical Fitness Field-based Tests in Down Syndrome

Ting Cao*
Physical Education department, Huanggang Normal University, Huanggang Hubei China
*Corresponding author e-mail: caoting@hgnu.edu.cn

Abstract. Physical fitness field-based test is an important means to understand the physical health status of the Down Syndrome. This systematic review was designed to identify the reliability of the field-based physical fitness tests study in Down Syndrome. The SCOPUS, SPORTS DISCUS, MEDLINE were screened for papers published from 1970 to 2018. Each study classified as high, middle and low according the results of the sample number, sample description, test methods, time interval and statistical methods. There are 30 high quality papers included in the present review. The reliability of assessing body composition, musculoskeletal fitness, motor fitness and cardiorespiratory fitness. Although some test methods for further studies are needed, but the reliability of the evaluation for mental disabled people physical field test provides a level of evidence based recommendations.

Keywords: Down Syndrome, Physical Fitness, Field-Based Test, Reliability, Systematic Evaluation

1. Introduction
Physical fitness test can be divided into laboratory test and field test. Laboratory tests can objectively and accurately measure physical fitness levels, but this method is costly, complex and difficult to implement on a large scale. Field testing provides an alternative means. However, the current physical fitness field test methods are all developed for healthy people. Because the mentally disabled group is in the task of understanding the task and using the disabled group qualitative test method, the reliability and validity are the two most basic indicators to be considered. Among them, reliability refers to the degree to which the results of repeated tests on the same sample using the same method are consistent [1]. It can be considered that reliability is the primary consideration indicator. Because reliability is the basis of validity, without reliability, validity cannot be discussed. Specifically, reliability mainly includes three types: (1) Internal Consistence refers to the consistency of test content; (2) Test-retest Reliability refers to the measurement results at different times Stability; (3) Inter-rater Reliability refers to the degree of consistency of the results of different testers testing the same object. This research mainly focuses on the systematic evaluation of the reliability of the mass field test of the mentally disabled, in order to provide a reference for the further theoretical research and practical application of the mass qualitative test of the mentally disabled [2].
2. Research Objects and Methods

2.1 Research Objects
This research is based on publicly published academic literature on the reliability of field tests of Down Syndrome. From the perspective of constituent elements, physical fitness research at home and abroad mainly includes five constituent elements: cardiopulmonary endurance, body composition, muscle ability, athletic ability and flexibility. Based on the literature on the reliability research of the five components of physical test, extract sample information (number, age, gender, degree of disability), test methods, reliability research types, statistical methods, research results and other information.

2.2 Research Methods
The two authors searched for relevant documents in SPORTS DISCUS, SCOPUS, MEDLINE, PUMBMED, Web of Science and Google Scholar database systems. The retrieved phrases include words for Down Syndrome, characteristic words for reliability test (Reliability/Reproducibility/Repeatability), and words for physical fitness test (Physical Fitness; Body Composition/Skinfold/Body Fat/Circumference) /Weight; Motor; Molecular Strength/Endurance/Force/Hand Grip/Dynamometer; Cardiorespiratory/Aerobic Capacity; Flexility). Search English literature published from January 1970 to December 2015.

The two authors screened the retrieved literature separately, and the criteria were: (1) the research object was aimed at the mentally disabled; (2) the literature on body composition, muscular capacity (muscle strength and muscular endurance), cardiorespiratory endurance, and exercise capacity ( The reliability of testing methods such as flexibility, balance, coordination, speed, and flexibility were studied; (3) the test method studied was field testing; (4) full English literature and full text available. Literature screening was performed by two authors respectively screening and extracting the title and abstract information of the literature. If the information meets the literature screening criteria, the full text content is tested. At the same time, consult the references to find the missed documents. Finally, the two authors checked the included literature and related information, and negotiated and resolved the inconsistencies.

3. Research Results
Based on literature search and preliminary screening, 58 literatures on the reliability of the qualitative test of Down Syndrome were determined. According to the literature screening method, 12 articles were deleted from the included literature due to factors such as non-English language, laboratory testing, and failure to obtain the full text. According to the literature quality assessment method, 30 articles were included in high-quality articles (67.4%), 16 articles were medium-quality articles (32.6%), and there were no low-quality articles. The degree of agreement between the two researchers was 87% (Kappa=0.795), and the inconsistencies were resolved through mutual negotiation.

In terms of the characteristics of high-quality literature, the number of samples ranges from 6 to 81, and the age distribution ranges from 5 to 89 years old. Among them, school-age teenagers and children are the main group, and the Down Syndrome groups under 10 and over 20 years old Less research. According to the American Psychiatric Association's classification of the degree of Down syndrome, namely, borderline Down Syndrome (IQ: 70-85), Down Syndrome (IQ: 55-70), moderate Down Syndrome (IQ: 35-55), and severe Down Syndrome (IQ: 20-35) and deep Down Syndrome (IQ<20). The degree of Down Syndrome of the sample is mainly based on the mild to moderate intelligence group, of which 7 studies are at the mild or critical line level, 11 documents are at the mild to moderate level, and 5 documents are at the moderate to severe level. There are also 7 documents that do not clearly indicate the level of Down Syndrome. There are also 8 articles on the research object specifically for Down syndrome patients in the type of Down Syndrome. In addition, there are multiple disability samples in some high-quality literature, such as the concurrent Down syndrome disability and visual disability, the concurrent Down syndrome and hearing disability, and the
concurrent Down syndrome and dyskinesia.

In terms of test items, most of the test items are derived from items developed for healthy human fitness testing, but some items have been adjusted. For example, the grip strength test is adjusted to a sitting position instead of a standing position; standing long jumps can be tried multiple times; the sit-ups are adjusted to curl-ups, and the tester does not have to sit up completely; the sit-ups are set from 30 The second is extended to 60 seconds. In the timed test, the testees were continuously encouraged; in the round-trip running test, each group was reduced to 1-2 people, shortening the round-trip running track to 16 meters; and the study provided assistants in the round-trip running test.

In terms of cardiorespiratory ability, there are a variety of methods involving cardiorespiratory ability tests, including fixed time tests (6 minutes walking, 9 minutes running), fixed distance tests (300 yards run, 600 yards run, 880 yards run, 0.5 mile run, 1 mile run) Walk), round-trip running test (16 meters, 20 meters) and step test. Specifically, there are a total of 17 high-quality documents related to the reliability research of cardiorespiratory ability. The most commonly used test methods are 6-minute walking (4 items), 20-meter round trip (4 items), 880 feet (3 items), and step test. (2 items), 1/2 mile run/walk (1 item), 600-yard run (2 items), 300-foot run (2 items), 16-meter round-trip run (1 item), PACER (Progressive Aerobic Cardiovascular Endurance Run, 1 item) and 1 mile walk (1 item).

In terms of muscular ability, there are 11 high-quality documents related to the reliability of muscle strength and muscular endurance. Among them, the most commonly used test methods are grip strength (4 items), sit-ups (3 items), and handheld dynamometer (Hand-held Dynamometer, 2 items), standing long jump (2 items), arm suits items), soft ball throw (1 item), isometric push-ups (1 item), bench press (1 item) and 30-second sitting Station test (1 item).

In terms of body composition, there are a total of 5 high-quality documents related to related research. The most commonly used test methods are skinfold test (4 items), body mass index (3 items), waist circumference (2 items), and bioelectrical impedance (1 item) ) And tibia length (1 item).

Table 1. High-quality literature on the reliability of group qualitative testing of Down Syndrome

| Dyer                  | C  | 0 | 2 | 2 | 2 | 1 | 2 | 9 |
|-----------------------|----|---|---|---|---|---|---|---|
| Pitetti.K.H.eta       | C  | 2 | 2 | 2 | 1 | 1 | 2 | 2 | 10 |
| Montgomery D L        | C  | 2 | 2 | 1 | 2 | 1 | 1 | 1 | 9  |
| Casey etal            | C  | 2 | 2 | 1 | 0 | 2 | 2 | 0 | 9  |
| Nasuti G              | C  | 1 | 2 | 2 | 1 | 1 | 0 | 2 | 9  |
| Elmahgoub.S.S         | C  | 2 | 2 | 2 | 2 | 2 | 1 | 2 | 12 |
| Vis etal              | C  | 2 | 2 | 2 | 1 | 2 | 2 | 2 | 11 |
| Fernhall etal         | C  | 2 | 2 | 2 | 1 | 1 | 0 | 2 | 9  |
| Fernhall etal         | C  | 2 | 2 | 1 | 2 | 1 | 1 | 0 | 9  |
| Teo-kohS.M.et al      | C  | 2 | 2 | 2 | 2 | 1 | 1 | 2 | 10 |
| Beets.M.Wet al        | C  | 2 | 1 | 2 | 2 | 2 | 1 | 1 | 10 |
| Gillespie             | C  | 2 | 2 | 1 | 1 | 2 | 2 | 9  |
| Baumgartner T         | C  | 2 | 2 | 1 | 2 | 2 | 2 | 2 | 11 |
| Waninge A et al       | B  | 2 | 2 | 2 | 2 | 1 | 2 | 2 | 11 |
| Havingatop A M        | B  | 2 | 2 | 1 | 1 | 1 | 0 | 2 | 9  |
| Winnick et al         | M1 | 2 | 2 | 1 | 1 | 1 | 2 | 1 | 9  |
| Aufsesser.P.M         | M1 | 1 | 1 | 2 | 2 | 2 | 1 | 1 | 9  |
| George                | M1 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 10 |
| Mercer V et al        | M1 | 2 | 2 | 2 | 2 | 1 | 1 | 1 | 10 |
| Wuang.Y.P.et al       | M1 | 2 | 1 | 2 | 1 | 1 | 2 | 1 | 9  |
| Gupta.et al           | M2 | 2 | 1 | 1 | 2 | 2 | 1 | 2 | 9  |
| Shield et al          | M2 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 10 |
| Nicolimietal          | M2 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 10 |
| Verma et al           | M2 | 2 | 2 | 1 | 1 | 2 | 2 | 1 | 11 |
| Villamonte            | C+B+M1+F | 2 | 2 | 2 | 2 | 2 | 1 | 1 | 11 |
4. Analysis and Discussion

4.1 Cardiorespiratory Capacity

Cardiopulmonary ability is the most tested content in high-quality literature. Among them, the 6-minute walk is the most commonly used test method, and there are 4 high-quality documents [3]. The 6-minute walk test requires the subject to walk as far as possible within the specified time without running. The ICC for 6 minutes walking between 11 to 26 years old (IQ: 40-70) is between 0.76 and 0.97. Some studies believe that ICC from 0.9-0.99 indicates high reliability, 0.80-0.89 indicates good reliability, 0.7-0.79 indicates general reliability, and <0.69 indicates poor reliability. There are also studies that believe that ICC can be classified as follows to judge the consistency of repeated tests: 0-0.2 has poor consistency, 0.3-0.4 has fair consistency, 0.5-0.6 has medium consistency, 0.7-0.8 has good consistency, and >0.8 consistency. The most ideal. It can be seen that ICC in different tests should be treated with caution. In the 4 sets of tests on 55 (27 males) people with Down Syndrome aged 11 to 26 in 6 minutes, ICC=0.84 for test 1 and 4, ICC=0.95 for test 2 and 4, and ICC=0.97 for test 3 and 4. The four groups of tests subdivided into three categories of mild, moderate and severe disability also showed an increasing phenomenon. Therefore, it is recommended to warm up twice before the formal test to obtain the best test results [4]. In a 39 obese or overweight adolescents with Down Syndrome (16.8±2.6 years old, IQ: 40-70), the ICC of the 6-minute walk test was 0.82, and there was no significant difference in repeated tests (P>0.05)[5]. Nasuti G and others have good reliability in the test of 13 adults (7 males, mild to moderate) aged 18 to 44 years old who walk for 6 minutes. The Cronbach's alpha of the two tests is 0.98, and the system error is -0.52 meters. The consistency limits are -54 and 43.6 meters. In the test of 29 Down Syndrome adults (39±11 years old, 19 males) who walked in 6 minutes, it was revealed that there was no significant difference in the 6-minute walk distance of Down Syndrome adults with or without severe heart disease (310±88 meters, 317±85 meters, P>0.05), CV=11%.

Limited distance running is one of the commonly used 880 yards run is equal to the 0.5 mile run, we discuss them together. There are 3 high-quality documents on 880 yards and 0.5 mile runs, with ICC ranging from 0.49 to 0.94. Pizarro is testing two groups of 81 adolescents with moderate Down Syndrome from 12 to 15 years old (educable group IQ: 55 to 70, 44 people, 25 males; trainable group IQ: 55 to 40, 37 people, 20 males) 880 yards in the running test, r=0.8 for the educable group and r=0.9 for the trainable group. In the test of 3 groups of 880 yards, the first group included 26 adolescents with moderate Down syndrome (16 males) aged 10 to 14 with ICC=0.75, and there was no significant difference between repeated tests. The second group included 12 adolescents with moderate Down syndrome (6 men) aged 15 to 18 with an ICC=0.49, and there was a significant difference between repeated tests; the third group included 15 adolescents with moderate Down syndrome (15 men) aged 12 to 14 The ICC=0.94, there is no significant difference between repeated tests. In the test of 23 mild to moderate 0.5-mile runs (13 males and 10 females) from 10 to 17 years old (13 males and 10 females), the repeated test correlation coefficient r=0.96.

Other tests involving aerobic endurance reliability test are 600 yards run, 300 yards run, 16-meter round-trip run and 1 mile walking test. Fernhall et al. performed the ICC of the 600-yard run test on 34 adolescents with mild to moderate Down syndrome (22 males, 14±2.3 years old), respectively, 0.97 and 0.96. The ICC of the 16-meter round-trip test was 0.96. In 40 adolescents with mild to moderate Down Syndrome (14.1±1.3 years, IQ: 50.3±10.6) from 10 to 17 years old, the correlation
coefficient of the 1-mile walk repeated test was 0.97. There was no significant difference in the 5 repetitions of Aufsesser PM for 36 (26 males) adolescents with Down Syndrome aged 12 to 20 years of 300 yards, $r=0.79-0.93$.

4.2 Body Composition

There are 5 reliability tests for calculating body fat percentage from skinfold thickness. 16 adolescents with mild to moderate Down syndrome (11 males, 15.4±2 years old, IQ: 35-55) in the triceps skinfold test, ICC=0.85, subscapular skinfold ICC=0.64, there was no significant difference in repeated tests. Pizarro was divided into two groups to divide 81 adolescents with moderate Down syndrome from 12 to 15 years old (IQ for education group: 55 to 70, 44, 25 males; IQ for trainable group: 55-40, 37, 20 males) In the skinfold test, ICC=0.98 in the educable group and ICC=0.99 in the trainable group [6]. 20 adolescents with mild Down syndrome (15±1.2, IQ: 63±11.5) between 14 and 17 years old, biceps (Midbiceps), triceps (Midtriceps), scapular muscles (Subscapular), and anterior psosas muscles (Suprailiac) ICC ranges from 0.94-0.99, the sum of the four skinfolds ICC=0.99, and the percentage error ranges from 5% to 47.5% [7]. In 45 severely disabled adults (28 males and 17 females, males 38 ± 11 years old, females 40 ± 10 years old) skinfolds were tested on the biceps, triceps, scapula, and anterior psosas muscles. There is no significant difference in the sum of the folds. The correlation coefficient of the sum of the four skinfolds is ICC =0.91, 95%CI=2.4mm, and the percentage error is 6.7%. There is no significant difference between repeated tests.

There are 3 studies on the reliability of the Body Mass Index. In the study of body mass index, in the test of body mass index of 17 adolescents with mild to moderate Down syndrome (15.4±2 years old, IQ: 35-55, 12 males), the body mass index of ICC=0.95 (95%CI=0.88-0.98), there is no significant difference between repeated tests, the first test SEM=22.34±4.03kg•m⁻², repeated test SEM=22.18±3.84kg•m⁻². For 45 severely disabled adults (28 males, males 38 ± 11 years old, females 40 ± 10 years old) body mass index test, repeated test ICC=0.98 (95%CI=0.98-0.99), the average percentage of the limit of agreement It was 5.8%, and there was no significant difference between repeated tests (Wilcoxon rank sum test $p=0.554$). In Donncha C's test of height and weight for 63 people with mild Down syndrome (IQ=63±11.5), the ICC of height and weight was 0.99, the 95% CI=8.4mm for height test, and the average percentage error was 0.5%. The 95% CI of the weight test is 1.6 kg, and the average percentage error is 2.8%.

4.3 Sports Ability

Balance ability includes static balance and dynamic balance. The dynamic balance ability test includes getting up and walking (Times Up and Go Test, referred to as TUG, 1 piece), four square step test (Four Square Step Test 1 piece), and inclined balance board. ICC=0.82 in the test of getting up and walking on 40 Down Syndrome (16 males, 10.6±4.4 years old). ICC=0.70 (95%CI=0.64-0.90) in the Four Square Step Test (FSST) repeated test on 13 teenagers (8 males, 12.8±2.9 years old), but there are significant differences in the repeated test. The homogeneity reliability test ICC=0.78 (95%CI=0.62-0.91), but there are significant differences in repeated tests. Villamonte tested the balance ability of 21 mentally disabled people (10 males) aged 5 to 31 (the dynamic balance test included 10 items such as walking on the balance beam, sitting and standing, etc., static balance such as standing on hard and soft floors with eyes closed) 6 items), only 3 items have good reliability for adolescents, 5 items have good reliability for adult women, and 9 items have good reliability for adult men. The test results questioned the reliability of some common test items for men and women and all ages, and pointed out that the balance test results of different age groups should be treated with caution.

The reliability research of the needle and the combination of sensitivity and speed is 4*100 meters. The test method comes from the ALPHA system. In the repeated test of 17 adolescents (12 males, 15.4±2 years old) 4*100 round-trip running, ICC=0.92 (95%CI=0.78-0.97), and there is no significant difference in repeated tests. Hilgenkamp TI tested the 10-meter incremental round-trip test. The
10-meter incremental round-trip test was repeated on the same day ICC=0.90 (95%CI=0.77-0.96), and ICC=0.76 (95%) every two weeks CI=0.48-0.90), the mean difference is 3.5±1.5m.

In the treatment of 100 adolescents with Down Syndrome aged 4 to 12 (64 mild IQ: 55-70, 36 moderate IQ: 25-54) Bruininks-ozerskys Test of Motor Proficiency (abbreviated as BOT-2, including sensitivity, Coordination, balance and speed). In the homogeneity reliability test, the total Cronbach’s alpha=0.92, the Cronbach’s alpha of each element=0.81-0.88, and the sub-test method Cronbach’s alpha=0.81-0.88. In repeated tests, the ICC of each component and sub-test method was 0.88-0.99[8].

4.4 Flexibilit

In the flexibility test, sitting forward bending is a classic test method. A total of 3 studies have explored the reliability of sitting forward bending. Pizarro was divided into two groups to seat 81 adolescents with mild to moderate Down syndrome aged 12 to 15 (IQ for education group: 55 to 70, 44, 25 men; IQ for trainable group: 55 to 40, 37, 20 men) In the test of flexion, $r=0.90$ in the educable group and $r=0.97$ in the training group. For 20 adolescents with mild Down syndrome (15±1.2, IQ: 63±11.5) between 14 and 17 years old, the correlation coefficient of sitting forward bending was ICC=0.97, 95%CI=4.9cm, and the average percentage error was 47.6% [9]. In another group of 36 elderly people (12 males, 50-89 years old) with different degrees of Down syndrome who were sitting forward with one leg, the left leg ICC=0.96 (95%CI=0.91-0.98) was repeated on the same day. Right leg ICC=0.95 (95%CI=0.89-0.98), repeat the test two weeks apart, left leg ICC=0.63 (95%CI=0.29-0.83), right leg ICC=0.71 (95%CI=0.43-0.87).

5. Conclusion

This study conducted a systematic evaluation of the reliability of the field test of Down Syndrome groups. The results show that, for a long time, the academic circle has conducted a lot of research on the reliability of the mass field test methods of Down Syndrome. In the high-quality literature, the research objects are mainly focused on adolescents and children with Down syndrome; the degree of Down syndrome is mainly mild and moderate Down syndrome. It should be pointed out that due to the differences in the research design and conclusion reporting of the high-quality literature of the mass-site tests for the disabled, although the reliability of some test methods is reliable, the limited number of test samples leads to limited reliability. Moreover, the validity of some test methods needs further study, therefore, the conclusions of this article need to be treated with caution in clinical practice [10]. In spite of certain limitations, this article can still form a framework for the reliability of the mass field test method for Down Syndrome based on the level of evidence (see Table 2).

| Evidence | Cardiopulmonary ability | Muscle capacity | Body composition | Athletic ability | Flexibility |
|----------|-------------------------|----------------|------------------|-----------------|------------|
| Stronger | 6 minutes walk, 880 yard run/0.5 mile run, 20-meter round trip/PACER | Grip, Sit-ups | Skinfold thickness test, Body mass index, Waist measurement | | |
| Medium evidence | Step test | Standing long jump, Hand-held force measurement, Arm support | | | Sitting forward bending |
| Not | 600 yards run, | Softball | Bioelectrical | Get up and walk, | |
enough evidence

| 300 yards run, 16-meter round trip, 1 mile walk |
|------------------------------------------------|
| throwing, 30 seconds sitting and standing test, isometric push-ups, Bench press |
| impedance, Tibia length |
| Square step test, Tilt balance board, 4*100m running, Go back and forth in 10m increments, BOT-2 |
| sitting position |

References

[1] Ortega F B, Ruiz J R, Castillo M J, et al. Physical fitness in childhood and adolescence: a powerful marker of health. International Journal of Obesity, 2007, 32(1):1-11.

[2] Winnick J P, Short F X. The Physical Fitness of Sensory and Orthopedically Impaired Youth: Project UNIQUE. Final Report. 1982:318.

[3] López-Girona A, Bachs O, Agell N. Prevalence of Chronic Health Conditions in Children With Down syndrome: a Systematic Literature Review. Intellectual & Developmental Disabilities, 2011, 49(2):59-85.

[4] Š Golubović, J Maksimović, B Golubović, N Glumbić, et al. Effects of exercise on physical fitness in children with Down syndrome. Research in Developmental Disabilities, 2012, 33(2):608-614.

[5] Patrick C, Sami E, Dirk C. Physical and Metabolic Fitness of Children and Adolescents with Intellectual Disability - How to Rehabilitate? Latest Findings in Intellectual and Developmental Disabilities Research. InTech, 2012.

[6] Westendorp M, Houwen S, Hartman E, et al. Are gross motor skills and sports participation related in children with Down Syndrome? Research in Developmental Disabilities, 2011, 32(3):1147-1153.

[7] Carmeli E, Ariav C, Bar-Yossef T, et al. Movement skills of younger versus older adults with and without Down syndrome. Research in Developmental Disabilities, 2012, 33(1):165.

[8] Oppewal A, Hilgenkamp T I M, Wijck R V, et al. Physical fitness is predictive for a decline in daily functioning in older adults with Down Syndrome: Results of the HA-ID study. Research in Developmental Disabilities, 2014, 35(10):2299-2315.

[9] Currell K, Jeukendrup A E. Validity, reliability and sensitivity of measures of sporting performance. Sports Medicine, 2008, 38(4):297-316.

[10] Essendrop M, Maul I, Läubli T, et al. Measures of low back function: A review of reproducibility studies. Clinical Biomechanics, 2002, 17(4):235-249.