

# Code of fair testing practices in education

[Environment](#), [Air](#)



## **Abstract**

The report analyses the code of fair testing practices in education and its relevance to the chosen topic in the previous topic that is unemployment and emotional or psychological disturbance caused by it to unemployed people. The code of fair testing has nine elements and all these elements are analysed. It is found that the code of fair testing provides good insights regarding ethical following during the test and various measures consideration for quality of test such as competencies and knowledge of test developers and test users, prevention of obscene language and other parameters. The test is found useful to education industry and educational people, but it is not found feasible for the employment topic. Therefore, the recommendation is to discard the test for the topic chosen.

### Code of fair testing practices in education

The code of fair testing in the education offers guidelines to professionals to fulfil their obligations towards using tests those are feasible and fair to all people irrespective of their age, ethnicity, gender, disability, linguistic background, religion and other personal characteristics. Purpose of fairness is determined by the code of fair testing. Fairness term implies that participants are informed about the general nature and purpose of test. This code is broadly applied to the education sector such as admissions, educational diagnosis and educational assessment as well as student placement. The code is not designed to cover certification testing, performance testing and computer based tests.

The code is primarily addressing the test developers and test users roles. Test developers are those people and companies that construct the tests. Test users of people as well as agencies select and administer test, commission the test development services or draw decision on test scores basis. The code offers guidance in four critical areas such as development and selection of appropriate tests, administration and scoring of tests, reporting and interpreting of the test results and communicating to the test takers.

Test developers have these responsibilities mentioned below. They are required to provide the evidence that the test measures and recommended uses of the test, as well as intended test takers. Strengths and limitation of the test should also be mentioned and inclusion of level of the precision in test should be described. The test developers are required to provide information regarding selection of content and skills to be tested and to communicate requisite information regarding the characteristics of a test in a streamline fashion that is understandable to the test users.

The test developer is required to provide evidences regarding the technical quality by including validity and reliability in the tests findings or the tests meet the predetermined purposes. A test developer is required to offer qualified test users representative samples of the test questions, directions to use the tests, answer sheets, score reports if required by participants and manuals. The test developer must not use the offensive content or offensive language while developing the tests questions and other related materials. He or she is responsible to modify the tests forms or administration

procedures for those people who have disability and need special arrangement to participate in the tests. The test developer is also required to obtain and offer evidences regarding the performance of test takers from diverse subgroups and to make significant efforts to create a feasible sample size for subgroup analysis. The evidences should be evaluated to know the differences in performance related to skills assessment.

A test users needs to consider these aspects. The test users must define the testing purpose, content of tests and skills being tested as well as test takers characteristics. He or she is required to use most feasible test on the basis of review of available information. The materials should be reviewed by the person supplied by test developer to select accurate, clear and complete tests. The test user is required to select test by a process including those people having correct skills, knowledge and training. The test user is required to evaluate the tests questions sample, directions of use, manuals, answer sheets, ad score reports. He or she is required to evaluate materials and processes that are used and must make sure to avoid the offensive content. The materials and procedures used by test developers must be evaluated by them along with the resulting test. The tests should be selected in appropriate modified forms for the test takers and disabilities must be considered. The test users should evaluate the available evidences regarding the performance of test takers by considering the diverse subgroups. The test user is also required to consider the feasibility of performance differences that are caused due to those exogenous factors.

Now considering the requirements of test developers and test users, the selected tests are not analysed on the base of application of code of fair testing practices in education.

The tests used in the previous reports are activity vector analysis test that consists of questionnaire and uses four factors such as emotional control, aggressiveness, sociability and social adaptability.

### **Element one: purpose**

As the purpose is required by the code of fair testing practices in education, choosing the activity vector analysis personality test to know the emotional and psychological problems of people who are emotionally impaired. The report mentions in analysis of week two regarding the past evidences of what the test measures. It is used to test the character of a person or broadly about humanity (Mount, 2015). The test is specifically developed to workplace issues solutions and therefore, the purpose of choosing those people who are jobless and emotional and psychological harms to them is amply assessed by this method (Activity Vector Analysis, 2018).

Another tool is BEM sex role that determines the gender biases issues. The Bem sex role inventory shows significance in terms of analysing various hormones associated with one gender class. The purpose of the assignment is to judge the factors of gender biases in test. The validity of this test to serve the purpose is established by other studies (Hoffman, & Borders, 2001). Another test chosen is of California psychological inventory that analyses the traditional values and its validity is also established by other studies findings (Stewart, 2008).

## **Element 2: appropriateness**

The chosen tests are found appropriate to conduct study of chosen topic.

The topic is studying the emotional and psychological impact of unemployment over them and as the first test activity vector analysis analyses the psychological and emotional part, its content has become appropriate for study. The second test BEM is used to understand whether the recruiters adopt bias approach to select males and females to different positions because as it is cited before that women are not considered for laborious and hard work, this kind of biases can be effectively dealt by BEM test.

The third test, California psychological inventory considers the traditional values and as it is found that males and females people working as recruiter may be caught with traditional thinking biases (Beattie, G., & Johnson, 2012). Therefore, the content of the California psychological inventory is found appropriate.

## **Element three: material**

The materials collected for such study consists of 434 questions for tests.

The tests consist of true and false options. There are 18 scale measurements and three in them are validity scales. The result of CPI scales reflects findings related to empathy, self-control, confidence, responsibility, independence and other findings. The activity vector analysis results are determined by an interview and stability and instability of emotional state of participants. BEM sex role personality test measures the behavioural tendencies. But analysis of week two assignment also reveals that the

methods used for the study and materials or content used in them are not specifically clear and understandable as the streamlined language has not been used.

### **Element 4: training**

The test users require specific knowledge, skills and training but the assignment matters suggest a shallow knowledge of the tools used for the study. It is mentioned that a psychologist having appropriate knowledge and skills will conduct the study to determine actual relationship between two genders. The reason of using the median test over the t-test is not elaborated. For the activity vector analysis, BEM and CPI, various mechanisms are used such as attribution questionnaire, general behaviour inventory, interpersonal circumflex, Stanford sleepiness scale and other mechanisms but the rationality of choosing these mechanisms to conduct the test is not elaborated as code of fair testing practice in education requires.

### **Element 5: technical quality**

The synthesis of reliability evidence is presented. To assess technical quality of the tests, a similar test of personality assessment inventory technical quality is analysed that reflects a sampling and relevant data to conduct the study to know the personality assessment inventory role to address the study objectives and it is found that the validity is established (Reidy, Sorensen, & Davidson, 2016).

Similarly the method used in another study reveals a sample being chosen for data collection purposes to know the personality attributes. The study

reveals a validity of chosen method of study that is item response theory related to personality traits measurements (Oswald, F. L., Shaw, A., & Farmer, 2015).

Another study to establish technical quality of measurement of personality traits is found to establish validity. For this purpose, statistical control is used to determine validity. The validity is established by this study as statistical controls help to reduce about 90 percent of faking effects (Sjöberg, 2015).

Next study also establishes validity while taking into account non-cognitive traits among the participation. The study focuses towards predictive validity of test and finds negative performance associated with aloofness and empathy (MacKenzie, Dowell, Ayansina, & Cleland, 2017).

Another study also establishes validity. The study investigates the predictive validity with curriculum-sampling tests that establish high predictive validity. But this method may not be employed to assess the emotional and psychological harms to jobless people. Therefore, its validity may be questioned.

Next study uses a temperament and personality questionnaire tool to assess personality styles in depression. There is a sample drawn to investigate validity of tool. Structural equation model is used to determine external validity. The resultant effect found is temperament and personality questionnaire is an effective tool to assess personality in depressive patients (Spanemberg, Salum, Caldieraro, Vares, Tiecher, da Rocha, & Fleck, 2014).



### **Element 6: test item and format**

Selection of test item was those people who had depressive thoughts due to unemployment. The participants participated in a questionnaire interview. But the number of participants was only 20 that do not produce generalization and more validity to the findings. The number of questions contained only 25 that were also not sufficient. The participants were allowed to allocate scores by themselves and they recorded their response in written format. This adds flexibility and requirements of code of fair testing practices in education.

### **Element 7: test procedure and materials**

The test procedures adopted by the chosen application vector test does not violate the privacy and emotional sensitivity. The questions do not comprise that situation that compelled participant to answer uncomfortable question. This also suggests that the application of code of fair testing practices in education is applied because no obscene question is being asked. The questions did not use offensive language regarding caste, color or creed.

Element 8: Modifications and Accommodations: the chosen test for study does not reveal any kind of consideration towards disable people to participate. This reveals that application of code of fair testing is not applied.

### **Element 9: group difference**

There is no evidence regarding the test taker' performance of diverse subgroups and creation of adequate sample sizes to subgroup analysis. Application of code of fair testing practices in education for tests is clearly not followed in this case.

## Summary

Application of fair testing practices in education has certain strengths and weaknesses regarding the topic I have chosen. These strengths and weaknesses are as mentioned. One of the strength of this test is its ability to provide ethical guidelines to conduct tests that are fair to all types of test takers irrespective of their age, disability, gender, race, national origin, ethnicity, linguistic background, sexual orientation or other kinds of personal traits. Strengths of application of code of fair testing practices are standardization of test. This test offers test takers to demonstrate their own knowledge their contribution to performance. Third strength of the test is offering an opportunity to test takers being informed about the content and nature of the test.

For development of test, this test considers taking into account evidences measurement, intended test takers, and strengths and limitations of the test to improve the precision of test scores. The guidelines to develop test and use the test by the application of code of fair testing practices are its strengths. The test demands to know the skills and content testing for development of tests, and communication of characteristics of test to the intended test users. The technical quality evidence of test is required.

Guidelines such as review of test, review of evaluate representative samples, technical quality, materials, appropriateness of test content, procedures and materials, selection of tests with appropriate modified forms, and review of available evidence on the test takers performance of diverse subgroups.

These requirements of the test are its strengths and therefore, the test has become appropriate to judge the personality attributes. The test also protects the rights of participants as it does not permit obscene sentences in test (Code of fair testing practices in education, 2018).

But there are certain limitations too of this test. The code is specifically meant to be applied to the education sector for the purpose of admissions, educational diagnosis, educational assessment and student placement. The test is not appropriate to cover testing of employment or the certification or licensing testing or other kinds of testing that are not the part of education sector. Another weakness is that the code does not allow adding any new principle other than the standards it follows that reduces its flexibility.

## **Recommendations**

The test is found feasible for the educational industry and educational tests. But the test has one inherent weakness that it cannot be used for employment testing and therefore, I recommend that the test is not appropriate for the topic I have chosen.

## **References**

- *Activity vector analysis*, (2018). Retrieved from <https://www.ava-assessment.com/>
- *Code of fair testing practices in education*, (2018). Retrieved from <http://aac.ncat.edu/Resources/documents/Code%20Final%20Edit%209-02revFINAL12Wall.pdf>
- *Code of fair testing practices in education*, (2018). Retrieved from <https://www.apa.org/science/programs/testing/fair-testing.pdf>

<https://assignbuster.com/code-of-fair-testing-practices-in-education/>

- Hoffman, R. M., & Borders, L. D. (2001). Twenty-five years after the Bem Sex-Role Inventory: A reassessment and new issues regarding classification variability. *Measurement and Evaluation in Counselling and Development*, 34, 39-55.
- Stewart, C. O. (2008). The Validity of the California Psychological Inventory in the prediction of police officer applicants' suitability for employment.
- Beattie, G., & Johnson, P. (2012). Possible unconscious bias in recruitment and promotion and the need to promote equality. *Perspectives: Policy and Practice in Higher Education*, 16 (1), 7-13.
- MacKenzie, R. K., Dowell, J., Ayansina, D., & Cleland, J. A. (2017). Do personality traits assessed on medical school admission predict exit performance? A UK-wide longitudinal cohort study. *Advances in Health Sciences Education*, 22 (2), 365-385.
- Oswald, F. L., Shaw, A., & Farmer, W. L. (2015). Comparing simple scoring with IRT scoring of personality measures: The Navy Computer Adaptive Personality Scales. *Applied psychological measurement*, 39 (2), 144-154.
- Reidy, T. J., Sorensen, J. R., & Davidson, M. (2016). Testing the predictive validity of the Personality Assessment Inventory (PAI) in relation to inmate misconduct and violence. *Psychological assessment*, 28 (8), 871.
- Sjöberg, L. (2015). Correction for faking in self-report personality tests. *Scandinavian journal of psychology*, 56 (5), 582-591.

- Spanemberg, L., Salum, G. A., Caldieraro, M. A., Vares, E. A., Tiecher, R. D., da Rocha, N. S., ... & Fleck, M. P. (2014). Personality styles in depression: testing reliability and validity of hierarchically organized constructs. *Personality and Individual Differences* , 70 , 72-79.