INNOVATIVE ASSESSMENT APPROACHES OF CREATIVITY

Ananda Kumar Palaniappan, Ph.D
Teh Ying Wah, Ph.D
Selva Ranee d/o Subramaniam, Ph.D
University of Malaya
anandak@um.edu.my

Assessment of any cognitive dimension has always been a challenge. More so if the construct that is being assessed has eluded accurate definition since gaining prominence. One such cognitive construct is creativity. Assessing creativity has always been a challenge since there is no one definition that has been universally accepted. One definition that comes close is Torrance's (1974) definition which has led to the development of the world famous creativity tests – Torrance Tests of Creative Thinking (TTCT). However, TTCT have been known to be a battery of tests that require would-be scorers to undergo training which is only available at the University of Georgia's Torrance Center for Creative Studies. This makes it difficult for users of this test to be trained before they are certified to score these tests accurately and reliably. Hence, many users of the TTCT have resorted to scoring the tests based on their readings from manuals and other sources. This paper will discuss one innovative approach that has been designed to not only assess creativity using an adapted version of TTCT quickly but also easily by any users anywhere in the world through the web-based creativity assessment system or WBCAS. Research findings will be presented to show that this system can be used to assess one's creativity quickly and accurately. This paper will also discuss some of the other advantages of using WBCAS in assessing creativity.

Keywords: Innovative assessment, creativity, web-based creativity assessment system

Creativity has been a much cherished and sought after cognitive ability ever since it gained prominence at the turn of the 19th century. It is an important construct much sought after in both the education and work environment. Since the 1980s, there has been an ongoing effort to enhance creativity in the teaching and learning process in schools in Malaysia. The Education Ministry has consistently emphasised the importance of thinking skills, especially creativity in the curriculum. This is also reflected in the philosophy of education which guides all initiatives in education since its formulation in 1988.

Rationale of the Study

Creativity assessment has always posed a big challenge. Firstly there is no one single definition of creativity that has been accepted by all creativity theorists and researchers (Treffinger, 1996). Creativity tends to manifest itself in different domains and dimensions; hence, there has never been a comprehensive assessment that is capable of incorporating all these dimensions.

Secondly, there is a vast array of creativity tests that assess creativity from many different perspectives. The type of assessment used depends on the definition of creativity employed (Treffinger, Young, Selby, & Shepardson, 2002). Among the most common ones are those that assess creativity as a process. Again there are a number of ways for assessing creativity as a process.

Thirdly, there is the problem of reliability and validity. To what extent do these tests or questionnaires assess creativity with reliability and validity? While some tests have a large research evidence to back them up on these issues, others are readily available but lacking in evidence of validity and reliability.

Fourthly, the ease of scoring and interpretation of the scores. Some instruments are well established but are very difficult to score unless one has a considerable amount of practice. The Torrance Tests of Creative Thinking (TTCT) is one that falls into this category. These instruments

have high reliability and validity but can be very tedious to score. Hence, there is a need to design a valid and reliable scoring technique that is easy to use without compromising on reliability and validity.

Most valid and reliable creativity tests, especially the TTCT, both the Verbal and Figural versions which are being used worldwide, have been very difficult to score. The laborious scoring procedures have driven researchers and educators to use other measures which are easier to score although they may have to compromise on the validity or reliability.

To overcome this problem, a computer program called the Web-based Creativity Assessment System (WBCAS) was written to help researchers and educators as well as employers and recruiters to score and interpret responses of their students or employees easily.

This computer software was designed to assess creativity based on the responses given by the respondents. This can be done either on a stand-alone CD version or on-line via a website. This way it is possible to get immediate results and interpretation of the respondents' scores as well as get feedback and pre-set recommendations on how to enhance the level of one's creativity.

Objective of the Study

This paper presents results relating to the validity, reliability and efficiency tests of a computer system called Web-based Intelligent Creativity Assessment System (WBCAS). It also discusses the approach used to assess the four dimensions of creativity proposed by Guilford (1956, 1967) and Torrance (1963).

Review of Literature

Creativity has been assessed in many different ways. These assessments were mainly based on the researchers' definition of what creativity is. Early measures of creativity used inkblots. As early as 1898, imaginative responses of intellectually superior students to a series of inkblots were studied (Dearbon, 1898). Kirkpatrick (1900) also used inkblots to study the level of creativity of preschool children. Later tests of creativity were measures of mental capacity and also social and personality aspects. There were also measures based on biographical information. Each of these measures has its strengths and weaknesses.

The inkblots test, for example, was much maligned for its inaccuracies and unreliability. These tests called Rorschach inkblots tests, named after the Swiss psychiatrist Hermann Rorschach (1884–1922), were said to be able to tap into the subconscious minds of the test-takers and reveal their inclinations towards divergent thinking among others.

Tests based on words or semantics sometimes called verbal tests, were also developed. For example, words association tests and remote association tests tend to tap into the ability of the person to be able to associate meanings to other objects that are uniquely related.

Some tests assess the person's ability to generate possible consequences. For example, test takers are asked to indicate what would happen if there were no sunlight or gravity. Again credit is given to unusual responses.

Other tests require the person to give possible titles to short stories or complete short essays. These require the person to be able to come up with clever titles and also be able to complete the story with unusual endings. Again, they are assessed based on uniqueness.

Some tests tap into the ability to think of unusual objects or situations based on different stimuli, for example, Thinking Creatively with Sounds and Words (TCSW) (Khatena & Torrance, 1998) which comprises two sub-tests, namely, Sounds & Images (S & I) and Onomatopoeia & Images (O & I).

More recent tests require test-takers to complete incomplete figures and to give possible titles for the completed figures. One world-renowned test is the Torrance Tests of Creative Thinking, Verbal and Figural Forms A and B (Torrance & Ball, 1984). The Figural Form comprises three activities. Respondents are given ten minutes to complete each activity. In the first activity an oval shape object is presented and respondents are required to draw lines in or around the object to make a picture that no one else will think of. The test-takers are required to give a clever and appropriate title

for the picture. The second activity comprises incomplete figures and require test-takers to complete these figures by drawing pictures that no one else will think of and give a suitable title for each. The third activity consists of 30 parallel lines and respondents are required to draw lines on and around the parallel lines to make a picture that tells a story. Again, the respondents are required to give a unique title for each picture.

The Verbal Form comprises six activities and respondents are given 45 minutes to complete them. The first three activities require respondents to look at the picture given and ask-and-guess what is taking place, guess causes and what the consequences will result from the situation depicted in the picture. The fourth activity deals with product improvement, while the fifth requires respondents to generate as many unusual uses for a cardboard box as possible. The sixth activity requires test-takers to think of what would happen if there were strings coming down from the clouds – a hypothetical situation.

Most of these tests appear to tap on different aspects of creativity. Hence, akin to the blind man trying to describe an elephant, each of these tests may be in fact looking at different dimensions of the same construct which may be presenting itself in different ways.

Scoring the tests presents a different challenge altogether. Some are easy to score; others are very time-consuming and require special training. One such test is the TTCT which comprises the Figural and Verbal forms. It may take up to 20 minutes to score each TTCT Figural Test and even more for the Verbal Tests. Hence, it is of important to think of ways of scoring these tests in the easiest manner possible and at the same time maintaining the accuracy of the scoring.

To help alleviate this time-consuming scoring problem, the fast growing software programming technology is used to design a software. The software written for this purpose is called Web-based Creativity Assessment System (WBCAS). It is designed to not only score the responses easily but also to maintain accuracy as well as provide instant scores and reports for anyone around the world. This paper describes this software and also discusses some preliminary results on its accuracy.

Definition of Terms

There are several definitions of terms that are used in this paper that relate to the various divergent thinking skills assessed by WBCAS.

The definition of creativity used is based on Torrance's (1974) definition of creativity:

> *a process of becoming sensitive to problems, deficiencies, gaps in knowledge, missing elements, disharmonies, and so on; identifying the difficult, searching for solutions, making guesses or formulating hypotheses about the deficiencies; testing and retesting these hypotheses and possibly modifying and retesting them; and finally communicating the results.* (p. 8)

The four main divergent thinking skills assessed are:

- Originality: the ability to generate responses that are unique and statistically infrequent
- Fluency: the ability to generate many relevant and meaningful responses
- Flexibility: the ability to generate relevant and meaning responses of different categories
- Elaboration: the ability to add details to an idea.

Methodology

Three activities were chosen based on Torrance Tests of Creative Thinking (TTCT) Figural Form A and B (Activity 3) which is the activity where respondents are required to complete as many parallel lines / circles as possible by drawing original pictures and naming them. These activities are then programmed into the software and programs were written to score the responses given.

The originality dimension was scored based on the principle of statistical infrequency. The scoring system is dynamic unlike the static scoring procedures employed in the manual. Moreover the scoring was based on local norms rather than the norms proposed in the manuals.

In phase one of the test, responses based on over 500 Malaysian responses were scored and entered into the database. These formed the initial database from which the new responses are scored.

As each respondent typed their responses into the software, the system will enter them into the database located in the secure server at the university. The system will then calculate the infrequency of these responses and allot points based on the following procedure:

Responses	Originality Points
< 1 %	3 points
1- 3 %	2 points
3 - 5 %	1 point
> 5 %	0 point

For every response entered into the computer, the system will add these into the database and recalculate the percentage of that response. If that response is less than 1% of the total responses given by the respondents who have taken the test so far, that particular response is given 3 points for originality. If it is between 1 to 3 percent, 2 points are given, and if it is between 3% and 5% then 1 point is given. However, if the response is a common response where more than 5% of the respondents have given it as the answer, no credit is given to that response.

Since the database continuously changes as more and more respondents take the test, the database is dynamic and ultimately will be able to give a more accurate assessment of a person's originality score based on the local norms. Hence, this computer system has the potential to give a valid assessment of a person's originality.

Individualized Scoring

Another improvement is being added into the second version of the software. Since it is felt that it may not be accurate to subject the scoring of responses of people of all ages and vocation to the same database, the second version will assess the person's originality scores based on different databases. Based on the demographic data the respondents type in during the registration, the WBCAS will choose the most age and vocation appropriate database to use during the scoring procedure.

The scoring of Fluency is relatively easy compared to that of originality since it is the score of the number of parallel lines attempted by the respondents.

For scoring Flexibility, the system will check each response typed in with those already in the database. If the response given fits the response in the system, it will check the category in which the response is placed and assign the category. This is done for each response. Then the system will assess the number of different categories the responses fall into and assign the total number of categories as a value for flexibility. As for elaboration, the system currently used has not been programmed to assess the number of details added to the basic idea. However, this is being programmed for the next version of the WBCAS.

Results

The following are the results of validity tests that were undertaken to ascertain the validity of WBCAS. This is undertaken by assessing each respondent's responses. Using the administrator option in the first screen of the software, administrators will be able to use a designated password to access each individual respondent's responses using the scoring procedures recommended in the TTCT scoring manual. These responses are then hand-scored for fluency, flexibility and originality. Correlation analyses were then used to ascertain to what extent these hand-scored values are related to those calculated by WBCAS. The following are the results of three pilot studies undertaken to test the validity of WBCAS in accurately scoring the three components of creativity.

First Validity Study

WBCAS was tested for validity using 36 University undergraduates. The mean age of the sample was 21.5 (SD = 1.81) years and there were 16 males and 20 females in the group. The responses were also hand-scored based on the TTCT scoring manual to obtain the actual Fluency, Flexibility and Originality scores. Both scores were analyzed using Pearson Product Moment Correlation for each component to obtain the validity indices.

Table 1

Pearson Product Moment Correlations between WBCAS and Hand-scored Scores

	FluHand	FlxHand	OriHand
FluCom	1.00**	.66**	.75**
FlxCom	.66**	.93**	.53**
OriCom	.73**	.60**	99**

* - Sig. at $p < .05$.   ** - Sig. at $p < .01$

FluHand = Hand-scored Fluency score   FluCom = WBCAS Fluency scores
FlzHand = Hand-scored Flexibility score   FlxCom = WBCAS Flexibility scores
OriHand = Hand-scored Originality score   OriCom = WBCAS Originality scores

There is a perfect correlation between the Fluency scores generated by WBCAS and the hand-scored Fluency scores (r = 1.00, $p < .01$). This is expected to be so because programming for the scoring for Fluency is quite straightforward. As for Flexibility, the correlations was also very strongly correlated (r = .93, $p < .01$). There is an even stronger correlation for Originality, the programming for which was quite difficult (r = .99, $p < .01$).

Second Validity Study

In this study the validity of WBCAS was tested using 116 university undergraduates (Mean age = 24.30 years; SD = 2.56) There were 52 males and 64 females in the sample. Responses were also hand-scored based on the manual to obtain the actual Fluency, Flexibility and Originality scores. Both sets of scores were correlated for each component to obtain the validity indices. The results are shown in Table 2.

192

Table 2
Pearson Product Moment Correlations between WBCAS and Hand-scored Scores

	FluHand	FlxHand	OriHand
FluCom	1.00**	.64**	.94**
FlxCom	.90**	.77**	.82**
OriCom	.91**	.71**	.94**

* - Sig. at $p < .05$.   ** - Sig. at $p < .01$.

FluHand = Hand-scored Fluency score   FluCom = WBCAS Fluency scores
FlzHand = Hand-scored Flexibility score   FlxCom = WBCAS Flexibility scores
OriHand = Hand-scored Originality score   OriCom = WBCAS Originality scores

Again there is a perfect correlation for the scoring of Fluency. As for Flexibility the correlation is somewhat lower but significant ($r = .77$, $p < .01$). For originality, the two scores correlated very strongly and significantly ($r = .94$, $p < .01$). This shows that WBCAS is a valid software to accurately assess these three components of creativity effectively and quickly.

Third Validity Study

In this third study involving 60 respondents, 18% of the respondents were from outside Malaysia. The majority of these were from the USA and China. Again, similar results were obtained (Table 3): Fluency ($r = 1.00$), Flexibility ($r = .89$) and Originality ($r = .94$). They were all significant at $p < .01$.

Table 3
Pearson Product Moment Correlations between WBCAS and Hand-scored Scores

	FluHand	FlxHand	OriHand
FluCom	1.00**	.87**	.90**
FlxCom	.89**	.89**	.80**
OriCom	.90**	.84**	.94**

* - Sig. at $p < .05$.   ** - Sig. at $p < .01$.

FluHand = Hand-scored Fluency score   FluCom = WBCAS Fluency scores
FlzHand = Hand-scored Flexibility score   FlxCom = WBCAS Flexibility scores
OriHand = Hand-scored Originality score   OriCom = WBCAS Originality scores

Discussion

The need for a quick and reliable measure of one's creative potential cannot be overemphasized. The need to have a model that is valid and at the same time comprehensive in assessing creativity is also equally important. The beta version of the computer software, WBCAS, designed to assess creativity accurately and quickly appears to be able to assess creativity effectively. The three validity studies undertaken have shown that the software is able to assess Fluency, Flexibility and Originality very well. The fact that this can be done on-line makes it even more useful as anyone anywhere will be able to access it.

Since WBCAS is able to generate the results immediately, it enables the respondent to instantly obtain a measure of his or her creativity in these three dimensions. The system is also able to generate recommendations based on his or her scores on all three components. Although the feedback is auto-generated and quite standardized, it does help respondents to look at the suggestions given to improve their level of creativity.

Apart from being able to provide hard copies of their scores and recommendations, WBCAS is also able to send these to the respondents via email. Test-takers will be able to take these tests anywhere there is Internet access and print out their scores. Employers may be able to ask prospective employees to take the test on-line at the applicants' own convenience and have the scores fed back to their database where they can retrieve these scores when needed.

Hence, WBCAS has immense potential in assessing creativity in terms of Fluency, Flexibility and Originality accurately. This next version of WBCAS will incorporate the individualized database based on the respondents' age and vocation to provide an even more accurate creativity assessment.

Implications

WBCAS has great potential in both the academic as well as the work environment. Both educators and employers will be able to use it obtain immediate feedback on their students' or employees' level of creativity. As for educators, this will help them to obtain a valid assessment of their students' level of creativity. The results can be used to plan and organize teaching and learning experiences aimed at improving students' creativity. Teachers will also be able to monitor the effectiveness of their teaching in enhancing creativity using WBCAS.

Employers will gain from WBCAS in two areas, recruitment and selection. Employers can choose new creative employees easily using this software. New prospective employees may be asked to take the test online and these scores could be fed back to the employers' database for use in recruitment. As more and more capabilities are built into the program, it appears that there is a huge potential for this software in creativity assessment for all purposes.

References

Dearbon, G. V. (1898). A study of imagination. *American Journal of Psychology, 5*, 185-190.

Guilford, J. P. (1956). Structure of intellect. *Psychological Bulletin, 53*, 267-293.

Guilford, J. P. (1967). *The nature of human intelligence.* New York: McGraw-Hill.

Khatena, J., & Torrance, E. P. (1998). *Thinking creatively with sounds and words: Norms-technical manual.* Bensenville, IL: Scholastic Testing Service (Originally published by Personnel Press, 1973).

Kirkpatrick, E. A. (1900). Individual tests of school children. *Psychological Review, 5*(1), 274.

Kirschenbaum, R. J. (1998). The creativity classification system: An assessment theory. *Roeper Review, 21*(1), 20-26.

Torrance, E. P. (1963). *Education and the creative potential*. Minneapolis: University of Minnesota Press.

Torrance, E. P. (1974). *Torrance Tests of Creative Thinking: Norms-technical Manual*. Bensenville, IL: Scholastic Testing Service. (Originally published by Personnel Press, 1966)

Torrance, E. P., & Ball, O. E. (1984). *Torrance Tests of Creative Thinking: Streamlined (revised) manual including norms and directions for administering and scoring Figural A and B*. Bensenville, IL: Scholastic Testing Service.

Treffinger, D. J. (1996). *Creativity, creative thinking, and critical thinking: In search of definitions*. Sarasota, FL: Center for Creative Learning.

Treffinger, D. J., Young, G. C., Selby, E. C., & Shepardson, C. (2002). *Assessing creativity: A guide for educators*, The National Research Center for the Gifted and Talented, University of Connecticut.