TEST AS TOOL FOR EFFECTIVE RESEARCH IN THE HUMANITIES

Stephen Billy OLAJIDE

University of Ilorin, NIGERIA

Abstract. Research, the painstaking gathering and verification of evidence that builds into ideas that should lead to the resolution of given problems, is one scholarly endeavour that has impacted greatly on the quality of man's life. As a result, considerable resources are devoted to research in different disciplines and at various levels of education across the world years in years out. In spite of the importance of research, it would seem that man has not applied it rigorously enough for problem solving, especially in the humanities, which include the arts, social sciences, law and education. It is possible for research in these academic areas to be strengthened with more purposeful use of tests, which are carefully designed instruments for gathering information that could lead to balanced judgements regarding knowledge dissemination and application. Thus, this paper is an attempt to advocate more appropriate use of test in the conduct of research in the humanities. In so doing, the paper explores the meaning, nature and scope of test and suggests ways of more effectively suiting it to research.

Keywords: test, research, humanities

Introduction

Test is an instrument designed, produced and implemented to elicit information about an individual respondent in respect of his knowledge, attitude, skills, assumed values and preference. In the humanities (which include the arts, education, social sciences, and law - at times the social sciences are made to stand alone), test is used to evaluate teaching and learning, and obtain data for research.

Test usually takes place before, during, and after an instruction in the humanities. The teacher may wish to determine what previous knowledge his group of learners have before taking his instruction. And as the instruction progresses, he may test in order to find out if the learners have been sharing additional experiences. At the end of the lesson, the teacher would conduct a test to find out the general intellectual standing of the learners based on the classroom instruction he has just given. While the test before a lesson is "diagnostic", the one holding as the class progresses in the "formative test "; that which takes place after the lesson is the "summative test". Being goal-oriented, test is purposive in nature; it is hinged on implemented educational activities. This may be why Lawal (2009) opines that in the context of schooling, test (like evaluation) is objective driven. Something cannot be built on nothing.

However, the utilitarian value of test makes it a tool that may be used beyond the confines of education. It aids research and researching in almost all disciplines. The systematic nature of test makes it a powerful tool of research, which is why the present paper is an attempt to show that research in the humanities can be made quite meaningful with the application of test.

Research has to do with probing into an existing situation or phenomenon relying on available evidences with a view to achieving a new and improved understanding of the situation. Denscombe (2001) describes research as an activity that is "significant in relation to current issues in society". Ac-

cording to Denscombe (2001), characteristics of good research include relevance, feasibility, coverage, accuracy, objectivity and ethics. Valid research is not haphazard. For research in the humanities to be of value, it needs to: (i) be based on incontrovertible data; (ii) address genuine problems; (iii) avoid biases; (iv) reflect that the researcher is aware of the weakness of the methods/approaches adopted in the study; (v) concentrate on the cardinal issues involved in the problem or phenomenon under investigation; (vi) contain appropriate and detailed answers based on the data obtained; (vii) cover all the issues pertaining to the topic under investigation; (viii) be properly funded; (ix) enjoy good timing – without sufficient time, the research may not be comprehensive and reliable enough; and (x) be hinged on relevant theories.

As already explained, the humanities include disciplines in Arts, Education, Law, and at times the Social Sciences, even though there are scholars who believe that the Social sciences have their distinct characteristics and may be allowed to stand apart from humanities (Frankfort-Nachmias & Nachmias, 2000; Denscombe, 2001).

While the task here is not to discuss the view as to whether the Social Sciences should be part of humanities, it may be pointed out that both groups of disciplines are as distinct as they are interrelated. One great thing they have in common is research; and another is that both share approaches to the conduct of research. Consequently, they employ test as an instrument for gathering data for research. And they are both amenable to the quantitative and qualitative strategies of analysing such data (Bazemer & Jewitt, 2001). In spite of the tendency for the social sciences to stand apart from the humanities in the grouping of academic disciplines, the present writer believes that both may better go together as the humanities.

Never the less, applying test in the methodology of research in the humanities can be demanding. Apart from the general characteristics of test already highlighted, it has the specific features of validity and reliability - two

attributes that make it a tool that can be used with confidence. Also, there are different types of test that may be used in the humanities with specific procedures for constructing and administering them.

Some types of test

Being a tool of evaluation (Ojerinde, 1986), test aids researcher judgement based on correct interpretation of data, but for the tool to achieve its usefulness, it has to be carefully designed, so as to fit into the nature, scope and rigour of the humanities. It also must be constructed by the researcher bearing in mind the attributes of the would-be respondents (the subjects) for the investigation. This is one consideration that makes test look like examination, although, as observed by Lawal (2009), both are two different things.

As there are types of examination, there are types of test, even when both may be administered under the same conditions. The researchers in humanities may conduct an examination using the platform of testing, and official examination scores could be turned to data of immense valuable.

Discussing the meaning, nature and scope of evaluation, Lawal (2009) and Ojerinde (1986) are agreed that there are generally two types of test - the free response and objective type tests.

The free response test

This is the essay type of test and is common in the humanities. It usually allows more time for respondent's responses, and it permits the respondent free use of words in the expression of thought.

Grading responses in the test, here, can be somewhat demanding. For example, a testee's power of expression may lure the tester into awarding a mark that could be hard to link with the quantity and quality of testee's responses.

The researcher in the humanities that opts for the free response type of test as a means of gathering data needs to be aware of its nature, weaknesses and strengths.

In the free response test family are others like matching/pairing, cloze, problem set, concept mapping, and free response/constructed items types.

Lawal (2009) views them all as test types "requiring short sentences or paragraphs.

Matching/pairing

This is a kind of test where the researcher requests respondents to connect one set of suggested responses to the other correctly with a view to ensuring that the respondents really have the ability to link ideas correctly. Items may be listed on the left-hand side of the page and respondents are directed to link them to their actual meaning on the right-hand by drawing lines.

Weaknesses:

- The free response type, particularly the essay, is not easy to score.

- The respondents involved in the study may be tempted to guess question ahead, thereby indirectly confounding the findings from the study.

- Different scorers may award different grades to the responses, even when the researcher-tester has provided a strong marking guide. And the researcher-tester, trying to ensure greater reliability of data by remarking the same set of responses, is most likely not going to give the same marks as those originally awarded to specific responses.

- Here, too, emphasis is more on the mastery of the cognitive aspect of the humanities content, whereas memorization may not yield data that would be indicative enough of the desired respondent behaviour.

- The free response test questions may not induce responses relevant to the critical areas of the study.

Strengths:

- It is quite easy to prepare free response questions.

- Also, it does not allow the respondent to cheat. The seemingly clumsy corpus in the front of each respondent cannot be easy to decipher by a dishonest other respondent.

- Then the free response test helps to indicate the t respondent's cognitive endowments. It reveals his knowledge, flow of thoughts and originality of ideas almost readily.

- Its enables the researcher to determine the level of the respondent's dexterity at deploying supportive ideas to arrive at the revealed response.

The items may also be represented with a space after each, for the respondents to provide in the correct letter; description or definition. The matching test may be prepared in a variety of other ways.

One weakness of the matching test however is that is draws mainly from the teestee's power to memorize facts; another is that it hardly prevents the respondent from recalling unimportant information. Then it does not prevent the respondent from guessing.

Cloze test

This is a test where the researcher makes use of graded passages from which words are systematically deleted. Many aspects of the humanities content may be tested at the same time with the cloze procedure.

The researcher in the humanities wishing to use the cloze method should decide on the ratio of words to delete from a passage, for example one word after every five (5) to seven (7) other words throughout the passage. The words that count here are the nouns, verbs, adjectives, pronoun, adverbs, and prepositions as they are the lexemes that have meaning and can call up specific images in the minds of respondents. It should also be noted that the further the

deleted words are apart, the easier the test becomes for the respondent, because the grater the contextual aid they would receive. Also, the researcher should begin deleting words as from the second or third sentence into the passage. The rubrics (instructions) in the test must be clear, and tasks involved must be in the light of the time available for the test. The cloze test has its weaknesses and strengths:

Weaknesses:

- There can be the temptation for the researcher to opt for a passage which readability may not have been determined suitably enough.

- It may be hard to determine the cognitive thrust that the context of the passage and mode of deletion place on the respondents, which may have implication for their performance, especially as they may be tempted to guess answers.

- It takes time to prepare a good cloze test.

Strengths:

- The cloze test is easy to prepare.

- It is also easy to grade.

- It confers contextual support on the respondents, as the grater the spaces in deletion, the higher the number of words that give semantic clues to the respondents.

- The cloze procedure (as it is sometimes called) can give the researcher, especially in language, data that can be useful for determining proficiency level.

- It closes task the memory of the respondents very well.

- It is also useful for revealing their mastery of specific aspects of the topic under investigation.

Problem set

In this type of test, the researcher prepares two or more short-answer items related to specific illustrations, maps, graphs, charts, pictures, or passages. The questions require the respondents to utilize their knowledge of the media presented in, say, labelling, explaining, describing, identifying and differentiating them.

Concept mapping

Concept maps are visual aids for thought and discussion, and can help the researcher to determine how much understanding the respondents have of basic concepts or terms. The researcher may, for example, wish to determine a population's level of knowledge of basic concept in History or some other discipline in the humanities. He then takes an appropriate sample from the population and selects his subjects to whom he would present a package of cards relevant to the concepts or terms of concern. He subsequently instructs the subjects (respondents) to: (1) sort the cards to known and unknown concepts; (2) arrange the cards on a surface to show how the subjects think of the words in the cards and/or the relationships among them; (3) draw and label the links among the concepts and terms.

The objective test

To Ojerinde (1986), objective test is that in which "... there is no possibility of difference of opinions among scorers as to whether responses are to be scored right or wrong." (p.228). He describes a test as a behaviour that may not be easily displayed by other means, which implies that precision is the strength of the objective test. Researchers that are interested in ascertaining and describing human behaviour have a reliable tool in the objective test.

Across levels of education, and in different types of research in the humanities, it is used. However, popular as it is, the objective test has its weaknesses.

Weaknesses:

- Objective test is difficult to prepare.

- It can be costly to produce as it usually involves many questions.

- It somehow permits respondents to cheat.

- Also, objective test ignores problem-solving skills.

- It hides respondent thought processes.

Strengths:

- The objective test allows the coverage of content as may be appropriate to the topics and intent of research.

- Reveals the respondent's mastery of facts, principles and tends of the ideas being investigated.

- It prevents respondents from guessing likely questions, especially in a study that involves teaching specific topics to respondents at the treatment stage in a semi-experiment to, say, determine the effectiveness of a method of teaching.

- It very well gives the result of the respondent's hidden thoughts - the result is concretised as the actual response made by the respondents or the score eventually indicated by the tester.

Attributes of tests, appropriated to research in the humanities

Good tests have attributes, the most important two of which are validity and reliability.

Validity

Validity is a research procedure by which the tool used is sure to measure what it has been prepared to measure. An instrument that performs unpredictably is not going to give reliable results Validity also means the judgment made by a respondent or the researcher of the internal structure and content of the instrument or tool in use. Researchers have different statistical ways of determining validity (Rasinger, 2010).

Any interpretation of a test depends on the outcome of the test. In other words, the accuracy and relevance of test (its true value) depends on the veracity of all claims arising from the correct use of information provided about the development and purpose of the test; and as observed by Lawal (2009) identifies six types of validity: (i) Face validity, which makes a test look like what it should be; (ii) Content or curriculum validity, if the test sufficiently covers the stipulated content; (iii) Empirical validity, whether the respondent will score as much as, more than, or less than he has scored on the test if he takes the same or an equivalent test sooner or later-sooner; (iv) Construct validity, where the tester establishes the theoretical basis of the test; (v) Criterion-related validity, that shows that a respondent's performance on a set of objectives in a test can be repeated by him on a more standard test that has similar objectives; (6) Convergent validity, by which the respondent performs the same way when the skills or attitudes tested in one test are tested again in equivalent items in different but related contexts.

Reliability

A test has reliability if it measures the knowledge, skills, attitudes and values it is supposed to measure consistently. Scholars are agreed that reliability has influence on validity (Chapelle & Brindley, 2010; Rasinger, 2010; Lawal, 2009; Ojerinde, 1986; Chastain, 1972).

Accuracy of measurement

Test to be used in the humanities should also give true mental capacity of the respondents selected for the research, by which the tests would be deemed to possess the attribute of accuracy. The test that lacks accuracy would present strong respondents as being weak; the converse is also true. Inaccurate test is inappropriate because it is not discriminatory and data from it would not lead to valid findings that may be generalized.

Representativeness

The good test does not deviate from the objectives of an investigation. If a test is meant to elicit data on attitude, it must not emphasize knowledge or skills, otherwise findings based on the results of such test would not be worth the researcher effort. The implication is that the findings would not extend the frontiers of knowledge in any way (Olajide, 2011; Abimbola, 1995).

Fairness

The test designed for research is said to be fair if it gives the respondents equal chance to demonstrate their knowledge, attitudes and skills. Thus, the tester should pay attention to the content from where the items of the test are motivated as well as reflect the goals of the research.

Systematics

The researcher must ensure that his test ranges through existing syllabus, and the questions posed should not be ambiguous and vague. If the test items are haphazardly prepared and presented, it may be difficult to use the test on the respondents effectively. The haphazard test would not attract respondents. Even if it attracts them, their responses would not be adequate, and the data motivated would be inappropriate.

Variety

One type of test may not be enough to avail the data required for a study. A combination of different types of test could be what would yield the envisaged data. Hence, the humanities researcher desiring to make use of test for gathering his data should develop objectives in the light of research topic, and decide on the number of instruments to be employed for reliable results.

Procedure in the construction and administration of tests

Researchers wanting to use tests as their tool for gathering data should take certain steps in order to achieve success: (1) They should define the objectives of their tests in line with the problems and purpose of their researches, as hinted elsewhere in this chapter. Such objectives must also be behavioural, concise, and realistic. They should reflect the attributes of the respondents; (2) The researcher needs to define the topic and make clear the aspects of the content to be covered in the test. All items in the test will then derive from the target content; (3) There should be Blue-Print (or Table of Specification) for the test, which would stipulate how many questions to be drawn on each aspect of the content. This would ensure that the questions are not skewed and that the content is properly covered. If the questions go beyond or are fewer than the required to cover the content, the data that would emerge would not be useful.

The question should not just be targeted at the content. The researcher should watch the process of aiming at the content, letting questions induce the respondents to demonstrate knowledge, comprehension and analytical prowess, power of synthesis, and capacity to apply the knowledge gained in arriving at reasoned judgement: (4) The questions will have to be directed at the foregoing skills proportionately and judiciously, so that none is neglected, and

in order that the respondent behaviour that is being investigated may be described comprehensively enough. This means that the researcher should weigh the content to be investigated in the light of the significance of topic and related objectives before constructing test items. It should be noted too, that unless the respondents have been sufficiently exposed to the content being targeted for investigation, they may not be able to respond adequately. One other way the researcher could safeguard the allocating of questions to content is to use the percentage to indicate how many questions would test each of knowledge, comprehension, analysis, synthesis, application and evaluation, which are key concerns in the humanities and would make the research have relevance to the society; (5) Then the researcher needs to factor the time that will be available to the respondents into loading the questions. Whatever the type of test, the time given for it to be taken must be appropriate, otherwise, respondents may encounter difficulties and the data to be obtained from the test may be grossly affected.

Using tests in humanities research

Much as test is one instrument the researcher may use in the humanities, the test must not be poorly constructed. Which means the researcher should have a proper understanding of the topic to be researched (Abimbola, 1995). Unless test is amenable to the nature and thrust of research topic, it would be futile to use it. Based on his knowledge of the topic, the researcher delves well into the literature to background for the topic (Olajide, 2011). He ensures that all the dimensions of the topic are captured in the literature reviewed, before indicating the problem being focused in the research. Olajide (2011) also stresses that the goals (or objectives) of the study should be in line with the problem. All of the foregoing steps have implication for any test the researcher in the humanities will use in gathering the desired data: (1) Findings from the data gathered through test, if properly interpreted, may have far-

reaching implication for the growth and development of the humanities. This compels the researchers concerned to make extra efforts at developing strong tests. Researchers and others interested in critical issues in the humanities could rely on data from test in making balanced judgment or decisions because the good test is a pragmatic tool of research; (2) Test data could indicate where problems lie in the humanities being investigated. Since the topics being investigated cannot pretend to be new, the researchers handling them must have, through appropriate review of available literature, ascertained the current states of the art regarding the topics. New things revealed from the data gathered through researcher made or adopted/adapted test could amount to some sort of solutions to the existing problems, and may throw up fresh insights in the disciplines. For example, test of knowledge, attitudes and skills among learners of English may provide data that will reveal new information about learner strategies. Findings from the data could also show how teachers of the language may handle materials to make learning better: (3) That data emanating from test could be used in drawing inference about individuals within and outside of groups has been elucidated by Lawal (2009) and Frankfort-Nachmias & Nachmias (2009). Their view is shared by Bezemer & Jewitt (2010) and Rasinger (2010). Rasinger additionally emphasizes that such comparison may be strengthened by the deployment of the right analytical (mathematical) models.

Concluding remarks

The humanities are rich disciplines. They have been able not only to preserve themselves individually and collectively but influence other areas of intellectual engagement through especially research. It may also be correct to observe that over time, the humanities have beneficiated greatly from practices in other disciplines point to one thing-all intellectual activities results in knowledge, which itself understands.

Researchers in the humanities make use of test for amassing the data they require, but the nature, scope, process and outcomes of test make it a tasking matter. There are different types of test, each with its own weakness and strength. The process of constructing test is quite demanding.

The researcher in the humanities opting for test as instrument of data collection has technical and ethical issues to resolve, for a successful outing. He needs to understand the demands of the topic to be investigated. He also needs to appreciate that without correct methodology, it would be difficult to contribute to knowledge by researching into the topic. Appropriate methodology means the researcher is not only familiar with what has been done on the topic and the procedure involved, but also armed with the right knowledge of the content relevant to the topic and the population (locale) to be involved in the investigation, especially the empirical investigation. If he elects to use the test for data collection, his knowledge would then guide him in constructing and administering the test and actually using the data generated. Only the appropriate use of test could lead to generalizable and enduring contributions in the humanities.

REFERENCES

Abimbola, I.O (1995). The origin, nature and functi ons of educational research (pp. 1-12). In: Jimoh, S.A. (Ed.). *Research methodology in education: an inter-disciplinary approach.* Ilorin: University of Ilorin.

Bezemer, J. & Jewitt, C. (2010). Multimodal analysis: key issues (pp. 180-197). In: Litosseniti, L. (Ed.). *Research methods in linguistics.* London: Continuum.

Chapelle, C.A. & Brindley, G. (2010). Assessment (pp. 247-267). In: Schmitt, N. (Ed.). *An introduction to applied linguistics*. London: Routledge.

Chastain, K. (1972). *Developing second-language skills: theory to practice.* Chicago: Rand McNally College Publishing.

Denscombe, M. (2001). *The good research guide for small-scale social research projects*. Buckingham: Open University Press.

Frankfort-Nachmias, C. & Nachmias, D. (2000). *Research methods in the social sciences*. New York: Wadsworth.

Lawal, A. (2009). Evaluation of students learning outcome I: types and uses of tests (pp. 344-364). In: Abimbola, I.D. & Abolade, AO. (Eds.). *Fundamental principles and practices of instruction.* Ilorin: University of Ilorin.

Ojerinde, D. (1986). *Educational tests and measurements.* Ibadan: Codat Audio-visual services.

Olajide, S.B. (2011). Ensuring adequate review of literature in language research: the role of supervisor (pp. 270-278). In: Kuupole, D.D., Bariki, I. & Yennah, R. (Eds.). *Cross currents in language, literature and translation: festschrift for J.P.A. Ukoyen*. Porto-Novo: Editions Sonou d'Afrique.

Rasinger, S.M. (2010). *Quantitative research in linguistics: an introduction.* New York: Continuum.

✉ Dr Stephen Billy Olajide
Department of Arts Education
University of Ilorin
Ilorin, Nigeria
E-Mail: billyolajide@unilorin.edu.ng