ISSUES AND CHALLENGES OF MEASUREMENT IN SOCIAL GERONOTOLOGICAL RESEARCH IN NIGERIA

Nnachi Amos Imo
Ebonyi State University, Abakaliki

Abstract

Measurement in the behavioural sciences is quite different from measurement in the natural sciences. The difference is more obvious in gerontology research where the aged are fraught with physiological problems, especially in the areas of sight and hearing. The Objective of this paper is to study the issues and challenges of measurement in carrying out research among the aged. Survey and documentary designs were the designs for the investigation. Observation and documents were the major instruments used in collecting data. The results show that the social researcher encounters many challenges in the areas precision, prediction, value orientation, varying social patterns, and behaviour instability, in measuring variables in social gerontology research. It is, therefore, recommended that gerontology researches be skilled in observation and interviewing, and there should be replication and objectivity in research in order to conduct valid and reliable research.

Key words: Issues, Challenges, Measurement, Social Gerontology, Research

Introduction

Gerontological social work, or social work with the aged or older adults, is a rapidly expanding field of practice, globally and in Nigeria. Of course, the population of the elderly rose from 4.6 million in 1991 to 7million in 2006. The Nigeria government is currently indicating interest in social and medical gerontology. To effectively rise to the occasion, social workers in Nigeria, who work with older adults need to have specialized knowledge about social conditions that confront older adults in Nigeria, including their problems, formal and informal care networks, health care issues poverty employment, housing and mental health. They need to have adequate knowledge about the normal and successful aging processes, and changes that are functions of aging process. Since the population of category referred to as the elderly is rapidly increasing, social workers in Nigeria also need to possess adequate knowledge of the demographic characteristic of older adults.

Social work research enables practitioners to possess the necessary knowledge itemized above. It is essential, if gerontology social work is to be both trustworthy, robust, sensitive and responsive to the needs of the elderly, that it bases its activities on evidence which has been generated in a systematic, rigorous and disciplined way. So research in gerontology social research is concerned with understanding the problems experienced by elderly individuals within societies and the impact of social policies and professional

interventions on them. Research in gerontology social work seeks to provide answers and evidence which can contribute to the improvement of policy and practice, to reduce age related problems and distress, and to promote elderly well-being.

To contribute to sustainable development in Nigeria, the evidence produced in gerontology social research must have been gathered in an organized way, using methods which are appropriate to the questions being asked and generating information and conclusions which are capable of being tested, verified or refuted. Gerontology social work research is therefore, "a systematic investigation concerning the elder population, which is conducted using the most appropriate designs and verifiable methods and analysis". It seeks to find answers to questions relevant to the category of the population referred to as the elderly or older adults.

Scientific inquiry of this nature works in a definite way. A basic element of science is the concept, a mental construct that represents some part of the world in a simplified form (Macaronis, 2010). Gerontologist use concepts to label aspects of aging and older adult including, young old, middle, old, oldest old, family, success, and to categorize the elderly in terms of their gender , social class or well-being. These are variables. A variable is a concept whose value or attributes are capable of changing in different circumstances, for example, age, marital status, income. The use of variables depends on measurement. Some variables are easy to measure, as when you place a baby on a scale to see how much it weighs. But measuring gerontological variables can be far more difficult.

The focus of this article is to examine issues of measurement in gerontology research. We will look at challenges of measurement in social gerontology research and find ways to carefully plan to minimize errors in measurement and examine steps to check on the adequacy of our measures.

Epistemological Issues

In order to test research hypotheses, concepts need to be translated into variable that can be measured. Measurement is the process of assigning values, such as numbers, to variables (Lindsey & Beach, 2004). Measurement moves research from the abstract, conceptual level to the concrete, empirical level. Macionis (2010) defines measurement as a procedure for determining the value of a variable in a specific case. According to Babbie (2010), Measurement is a careful, deliberate observation of the real world for the purpose of describing objects and events in terms of the attributes composing a variable. Gabor and Grinnell define measurement as the process of systematically assigning labels to observations. Measurement is most applicable in positivism – the study of society based on systematic observation of social behaviour. In using measurement, the job of the gerontology social researcher is to discover reality by gathering empirical evidence.

Measurement is always somewhat arbitrary because the value of any variable, in part, depends on how it is defined. In addition, it is easy to see that there is more than one way to measure abstract concepts, such as ageing, frailty, intelligence, success, satisfaction and fulfilment. Good research, therefore, requires that social researchers operationalize a variable. Operationalization means specifying what is to be measured before assigning value to a variable. This means that you have to decide exactly what you are going to measure. The process through which we specify what we mean when we use particular terms in research is called conceptualization. It is the process whereby fuzzy and imprecise notions

(concepts) are made more specific and precise. For example, when you say older adults, you have to decide exactly what you are going to measure, say, those 60 years and above or those 65 years and above, or whether it is young older, middle old, or oldest old.

Statistical Measurement

Gerontologists use three different descriptive statistics to report averages. The simplest statistic is the mode, the value that occurs most often in a series of numbers. In research study of older adulthood gerontologist face the problem of dealing with huge numbers of people. For example, how does the gerontological researcher report the income of millions of aged Nigerians, the aged in both rural and urban areas? Listing streams of numbers will carry no meaning and will tell us nothing about the people as a whole. To solve this problem, sociologists use descriptive statistical measurement to state what is average for a large number of older adults. Although it is easy to identify, gerontologists rarely use the mode because it reflects only some of the numbers and is therefore a crude measure of the average.

A more common statistic, the mean, refers to the arithmetic average of a series of numbers calculated by adding all the values together and dividing by the number of cases. Because the mean is pulled up or down by an especially high or low value, it can give a distorted picture of data that include one or more extreme scores. Another statistic, the median, is the middle case, the value that occurs midway in a series of numbers arranged from lowest to highest. The median, unlike the mean, is not affected by any extreme scores. In such cases the median gives a better picture of what is average than the mean. The gerontologist should, therefore, know when to use either the mode, or the mean, or the median in his gerontological measurement.

Levels of Measurement

In research studies on aging and older adulthood data are collected to measure the variables of interest. Attributes that compose variables may be related in different ways. Because of these additional relationships among their attributes, different variables may represent different levels of measurement. Here, we examine four levels of measurement: nominal, ordinal, interval, and ratio.

Nominal Measures Nominal measures merely offer names or labels for characteristics. Variables whose attributes have only the characteristics of exhaustiveness and mutual exclusiveness are nominal measures. Examples of these include religion, gender, ethnicity, political party affiliation, birth place and age.

Imagine a group of elderly people being characterized in terms of such variable as age. Imagine asking them to stand according to their old age category (their attributes), all those young old, those middle old, and those oldest old. The variable would be age; the attributes would be young old, middle old, oldest old. All the people standing in a given group would share at least one thing in common, and would differ from the people in another group in the same regard. To facilitate the collection and processing of data, we assign different code numbers to the different categories, or attributes of nominal variables. Thus we record "1" to designate young old "2" to designate middle old, and "3" to designate oldest old. These code numbers have no quantitative meaning. They are only convenient devices to record qualitative differences.

The word nominal comes from the same Latin root used in words like nominate and nomenclature-words that have something to do with naming. No matter what code number we assign to them, no matter how high or low that number may be, the code refers only to a name, not an amount. Thus, in coding ethnicity, if we assign a "1" to Hausa, a "2" to Igbo, a "3" to Yoruba, a "4" to Efik, and a "5" to Ijaw, we are not implying that someone with a higher code number has more of something than someone with lower code number. Consequently, when we statistically analyze nominal data, we cannot calculate a mean or a median. Our analysis would be restricted to calculating how many people were in the various categories, such as when we say that 40 percent of the caseload is Igbo, 30 percent Hausa, 25 percent Yoruba, and so on.

Ordinal Measures Variables whose attributes may be logically rank-ordered are ordinal measures. The different attributes of ordinal variables represent more or less of the variables. Examples of such variables are social class, racism prejudice, conservatism, sexism, client satisfaction and the like.

An example of ordinal measurement could be if we ask older adults how satisfied they are with the services they receive from agencies, or to rate the quality of those services. We might ask them whether they are very satisfied, satisfied, dissatisfied, or very dissatisfied. This would tell us the rank order of their level of satisfaction, but it would not provide a quantity that allow us to say that client at one level of satisfaction were exactly twice as satisfied or three times more satisfied than clients at another level of satisfaction. Similarly, if one client rated the quality of service as excellent and a second client rated them as good, then we could say that the first client gave a higher rating to the services but not precisely how much higher the rating was. For example, we could not say that excellent is one third better than good, or two times better than good.

If elderly clients are rating service quality on an ordinal scale with the categories excellent, good, fair, or poor, the gerontological researcher might as with nominal measurement, assign code numbers to represent the ratings. For example, excellent "4", good "3", fair "2", and poor "1". Unlike nominal measures, these code numbers would have some quantitative meaning. That is the code 4 would have a higher rating than the code 3, and so on. But the quantitative meaning would be imprecise; it would not mean the same thing as having four cars as oppose to having three, two or no cars. Whereas we can say that an elderly with four cars has four times as many cars as an elderly either with one car, we cannot say that a client who felt the services were excellent (code 4) found them four times better than the client who felt they were poor (code 1). The word ordinal is thus connected to the word order, and means that we know only the order of the categories, not their precise quantities or the precise differences between them (Allen & Babbie 2005).

Interval Measures This is a level of measurement describing variable whose attributes are rank-ordered and have equal distance between adjacent attributes. Here, the actual distance separating the attributes that compose the variables does have meaning. The logical distance between attribute can be expressed in meaningful standard intervals. A physical science example is the Fahrenheit or Celsius temperature scale. The difference or distance between 80 degrees and 90 degrees is the same as that between 40 degrees and 50 degrees.

The interval measures commonly used in gerontological and most social scientific research are constructed measures, such as standardized intelligence tests that have been

more or less accepted. The interval separating IQ scores of 80 and 90 may be regarded as the same interval separating scores of 115 and 125 by virtue of the distribution of observed scores obtained by many thousands of people who have taken the test over the years. It must be noted that a person who received score of zero on a standard IQ test could not be regarded, strictly speaking, as having no intelligence, although we might feel he was unsuited to be a university lecturer or even a university students.

Ratio Measures This is a level of measurement describing a variable with attributes that have all the qualities of nominal, ordinal and interval measures, and in addition are based on a true zero point. In the physical sciences the Kelvin temperature scale is an example of ratio measure. It is based on absolute zero, which does mean a complete lack of heat. Examples from social work research would include age, length of residence in a long term care facility, number of children, number of days spent hospitalized and so on.

In a study on aging we might ask a gathering of people to group themselves by age. The fact that all members of a single group share the same age and that each different group has a different age satisfies the minimum requirements for a nominal measures. Arranging the several groups in a line from youngest to oldest meets the additional requirement of an ordinal measure and enables us to determine if one person is older, younger or the same age as another. Finally, because one of the attributes included in age represents a true zero (babies carried by women about to give birth), it also meets the requirement for ratio measure because it permits us to say one person is twice as old as another. When we compare two people in terms of ratio variables, it allows us to conclude (1) whether they are different (or the same) (2) whether one is more than the other, (3) how much they differ, and (4) what the ratio is one to another (Babbie, 2010).

Common Sources of Measurement Error

In studies on aging and older adulthood, gerontological researchers need to be mindful of the extreme vulnerability of the measurement process to sources of measurement error. Measurement errors occur when we obtain data that do not accurately portray the concept we are attempting to measure. Common sources of measurement error come in two types:

Systematic Error Systematic error occurs when the information we collect consistently reflects a false picture of the concept we seek to measure, either because of the way we collect the data or because of the dynamics of those who are providing the data. In researches on aging and older adulthood, the most common way our measures systematically measure something other than what we think they do is when biases are involved in the data collection. According to Allen and Babbie (2005) biases can come in various forms. We may ask questions in a way that predisposes individuals to answer the way we want them to, or we may smile excessively or nod our heads in agreement when we get the answers that support our hypothesis. Or we may distort their true views or behaviours. The former bias, agreeing or disagreeing with most of all statements regardless of their content, is called *the acquiescent response set*. The later bias, the tendency of people to say or do things that will make them or their reference group look good, is called *the social desirability bias*. The social gerontologist doing research on older adulthood should be wary of the social desirability bias as an important source of measurement error. This is because in many traditional societies, the elderly would be prone to give responses that

portray their family and kindred as good. Guidance to the gerontological research in relation to this problem is to suggest that you imagine how you would feel in giving each of the answers you offered to the respondents.

Another common source of bias which leads to measurement error is *cultural bias*. Cultural bias stem from cultural disparities. The often cited cultural bias in measurement is intelligence tests which are often cited as biased against certain ethnic minority groups. The potential for cultural bias in measurement is not controversial. So, gerontological researchers should be able to develop questionnaire items that have measurement equivalence in their study of older adults from diverse ethnic backgrounds.

Random Error In measuring variables in gerontological research, random errors have no consistent pattern of effects. They do not bias our measures. They make them inconsistent from one measurement to the next. However, researchers in aging and older adulthood should not assume that whenever data changes over time that we have random error. Sometimes things really do change. When thing change our measures should detect such change. What random error means is that if the things we are measuring do not change over time but our measures keep coming up with different results, then we have inconsistencies in measurement.

In aging and older adulthood studies, random error can take various forms. If our measurement procedures are cumbersome, boring, complex or fatiguing then our respondents or subjects may say or do things at random just to get measurement over with as quickly as possible. For example, halfway through a lengthy questionnaire full of complicated questions, elderly respondent may stop giving much though to what the questions really mean or how they truly feel about them. Here the cumbersome and complicated natures of the questionnaire produce inconsistencies in the subjects' responses.

Lack of familiarity with social service jargons can also produce random errors. For instances, elderly clients who are not familiar with social service jargons are asked whether they have received brokerage, advocacy, or linkage services. Not understanding what they were being asked, because of their ignorance of the service jargons, they might answer differently the next time they are asked, even though the situation remains the same. But if the elderly clients, though having no idea of the service jargons, decide to answer affirmatively to every question in order not to appear negative or get the practitioner into trouble, that would represent systematic error arising from social desirability bias or acquiesce response set.

Random errors can be very serious problem in measurement. These errors can make a highly effective intervention appear ineffective. On the other hand, they can make services which produce not effect to appear effective. They distort the results produced by measuring instruments.

Errors in Alternate Forms of Measurement

Four alternate options that are commonly used to measure variables in social gerontological research are written self-reports, interviews, direct behavioral observation and examining available records. Let us look at how each of these is vulnerable to measurement error.

Written Self-report Questionnaires or scales are the most commonly used measurement options, perhaps because they are relatively in-expensive and expedient way to collect data. Written self-report can be used to gather background information about people (their age, gender, ethnicity, sex, and so on) or to measure their knowledge, attitudes, skills, or behavior. Errors may arise from how we word our items which may create difficulty in understanding; or in the length of our instrument which may create fatigue; or in the complexity of the structure which may be confusing. The way we have worded our item may also produce bias, or result in socially desirable response set. Another error might arise from your instrument not measuring what it is intended to measure – that is not valid.

Interviews: Interviews might be costly and more time – consuming than written self-reports, but they have the advantage of providing the opportunity for the researcher to clarify questions the respondent does not understand, and that questions are not skipped, and for interviewers to observed things about the respondent and probe further. However, interviews are also vulnerable to social desirability biases, especially when interviewers introduce their own biases such as by smiling or nodding when respondents answer in ways that support a study's hypothesis. This biased response can be worse when the elderly client is required in a face-to-face interview to assess a workers therapy. It will be hard to say it has not helped.

The characteristics of different interviewers might lead to measurement errors or affect how respondents answer questions. For example, a Niger- Delta interviewer might get different responses from an Hausa interviewer when interviewing people about their stand on oil derivation policy. Again, a female interviewer might get different responses from a male interviewer when asking people how they feel about equal rights for women.

Direct Behavioural Observation This alternate form of measurement enables the gerontological researcher to observe the elderly directly and not rely on written elf-report or interviews. It produces first hand information observed by the researcher himself. Although it is costly and can be more time consuming, it has the advantage of seeing behavior for ourselves and not having to wonder whether the way people answer questions really reflect how they actually behavior. Nevertheless, direct observation, too, can be highly vulnerable to systematic error, such as social desirability biases. People are quite different from unthinking objects. People who know they are being observed may act in a much more socially desirable manner than when they are not being observed or when they do not know they are being observed. In addition, the observer might be biased to perceive behaviours that support their study's hypothesis. Differences in observation skills may also affect how observer observe and record the event being observed.

Examining Available Records Examination of available records is perhaps the least time-consuming and costly measurement option. Using this option, a gerontological social work researcher might want to examine the process notes in the record of those who provide services techniques or provide different services. Here errors might arise from practitioners tendency to exaggerate their records regarding the amount of time they spend on certain activities in the belief that someone might use those records to evaluate their performance. That would be systematic error. Sometimes, practitioners may resent all the record keeping

that is expected of them and thus are not careful in documentary their tasks. Here random errors are made.

Criteria for Measurement Quality

In research studies on aging and older adulthood, we need to be aware of the above-discussed potential errors. So we must take steps to deal with them. We must therefore make sure, before we implement the study, that the measurement procedures we use have acceptable levels of reliability and validity. Regardless of the specific variables that are of interest to a particular gerontological study, or the manner in which there are measured, it is essential that the measurement instruments used are both reliable and valid.

Reliability

In social gerontological research, reliability is said to be achieved when a particular instrument applied repeatedly to the subject yield the same results each time. Reliability has to do with the dependability, or consistency, of the instruments used to measure variables. Reliability has to do with the amount of random error in a measurement. The more reliable a measure is measure, the less random error in it.

It must be noted that reliability does not ensure accuracy. For example, if a large school football team player set his bathroom scale to shave seven kilogram's off his weight just to make him feel better. Although the scale would "reliably" report the same weight for him each time, the report will always be wrong though consistent. It would be wrong due to systematic error – that is biased scale.

In gerontological research, reliability problems crop up in many forms. One form is that in survey research, different interviewers get different answers from respondents as a result of their own attitudes and demeanours. Different coders also tend to code the same content differently. We also have reliability problem when we ask questions which the respondents do not know the answers: How many times have you been to the village square? Or when we ask people about things that are totally irrelevant to them: Are you happy with South African's current relationship with Costa Rica? Or when respondents do not understand what our questions mean, such as when we ask elderly rural women how good the android phone is. Or when we ask questions that are so complicated that even those who have clear opinion on the matter might give different answers at different times.

Types of Reliability: In studies on ageing and older adulthood, the type of measurement reliability that is most relevant to a particular study varies according to the study's purpose and design. ***Inter observer reliability or interrater reliability*** is the term used for the degree of agreement or consistency between or among observers or raters. Here reliability is achieved when there is agreement in the ratings or observation. The second type of reliability is ***Test-Retest Reliability***. This is a term for assessing a measures' stability over time. In studies on older adulthood it is important to use a stable measure - that is a scale that provides consistency in measurement over time. A third type of reliability is ***the internal* Consistency Reliability**. Internal consistency reliability is an assessment of whether the various items that make up the measure are internally consistent.

In gerontological research, it is very necessary that we create reliable measures. Below are some techniques we can use to create reliable measures:

(i) In asking the elderly for information, if your research design calls for that, be careful to ask only things the respondents are likely to be able to answer.
(ii) Ask about things relevant to than and be clear in what you are asking.
(iii) Another way is to use measures that have proven their reliability in previous research.
(iv) To guard against interviewer unreliability, supervisor in a survey can call a subsample of the respondents on the telephone and verify selected pieces of information.
(v) Replication by other researchers may help to produce reliable measures.
(Vi) Finally, clarity, specificity, training, and practice will avoid a great deal of unreliability and problem (Rubin & Babbie, 2005).

Validity

We stand a good chance of drawing appropriate conclusions from our data only if we are measuring what we think we are measuring. Validity is a term describing a measure that accurately reflects the concept it is intended to measure (Bostwick & Kyte, 1993). The concept of validity applies both to specific measurement instruments and to the findings of research studies. There are several different types of validity, but all have to do with whether we are measuring what we think we are measuring: In gerontological studies on the elderly, **Internal Validity** of a research study refers to the accurate identification and interpretation of the factors, or effects, responsible for an observation. On the other hand, the ***external validity*** of a study refers to whether findings obtained from the sample of study participants can be generalized to the population of interest (Erber, 2005).

Another type of validity is ***ecological validity***. Ecological validity refers to whether the results with a particular test instrument accurately reflect real-world functioning or real-world behavior. For example, scores on an intelligence test might be reliable for young and old adults. But do these scores inform us about the level of competence young and older adults are likely to demonstrate when they deal with real-world situations. Another type of validity is ***heterotypic continuity.*** Heterotypic continuity is a type of validity which has to do with whether a measure used to assess some underlying quality, or characteristic, has the same degree of internal validity for different age groups in a cross-sectional study, or for the same people as they are followed over time in a longitudinal study. ***Content Validity*** is another type of validity. The term refers to the degree to which a measures covers the range of meanings included within the concept. For example, a test of mathematical ability cannot be limited to addition alone but would also need to cover subtraction, division, multiplication, and so forth (Carmines & Zeller, 1979).

Face Validity is the quality of an indicator that makes it seem a reasonable measure of some variable. For example, that the frequency of an elderly man's visit to hospital is an indication of that man's health status seem to make sense without a lot of explanations. It has face validity. ***Factorial Validity*** refers to whether the number of constructs and the items that make up these constructs on a measurement scale are what the researcher intends. It refers to how many different constructs a scale measures and whether the number of constructs and the items that make up those constructs are what the researcher intends. **Criterion-Related Validity** refers to the degree to which a measure relates with an external criterion. When we assess the criterion validity of an instrument we select an external criterion we believe is another indicator or measure of the same variable that our instrument intends to measure. For example, the validity of a written driver's test is determined by the

relationship between the scores people get on the test and how well they drive. In this example, driving ability is the criterion. Finally, ***Construct validity*** refers to the degree to which a measure relates to other variables as expected within a system of theoretical relationships and as reflected by the degree of convergent validity and discriminate validity. **Convergent Validity** refers to the degree to which scores on a measure corresponds to other measures of the construct, while ***discriminate validity*** refers to the degree to which scores on an instrument correspond more highly to measure of the same construct than they do with scores on measures of other construct.

Sensitivity
Changes in elderly clients' problem level are often small. It is therefore important than a measuring instrument be able to detect small changes. What is needed is an instrument that is reliable or stable enough to ignore irrelevant changes and sensitive enough to detect small changes in the level of real problem.

Non-Reactivity
The very act of measurement sometimes affects the behavior, feeling, or knowledge level objective that is being measured. Nonreactive refers to measurement instrument that does not affect the behavior, feeling, or knowledge objective being measured. For example, an elderly cigarette smoker who begins to count the number of cigarettes he smokes may smoke fewer cigarettes simply as a result of the counting, not as a result of any intervention. A synonym for the concept nonreactive is unobtrusive. A gerontologist's aim is to record a measurement as unobstructively as possible.

Utility
Utility means usefulness. Utility implies instrument being useful and practical in a particular client. For example, a practitioner may discover a perfect instrument for measuring depression in order adults. It is valid, reliable, nonreactive and sensitive, but it is also five pages long. In addition, it takes a long time to score, and numerical score, once obtained is difficult to translate into a meaningful assessment of the client's depression. This instrument, though perfect in every other respect, is useless in practice because it takes too long to complete and too long to score and interpret. Instruments that have utility are acceptable to the client, they are easy and quick to administer and score, and they give results that reveal the client's current state (Gabor & Grinnell, 1998).

Some Challenges to Measurement in Gerontological Research
Scientific research is an important way of generating knowledge. It is an important way of knowing. Measurement plays a vital role in this knowledge generation. In real life application, measurement in studies of older adults has several challenges. This limitations apply to many other social scientific enquiries. Some of the limitations are examined below.

1. Human Behaviour is too Complex to Measure Individual Actions Precisely. Scientists in the physical sciences can measure objects with precision. For example, Astronomers calculate the movement of objects such as moon, comets, and planets in the sky with remarkable precision. However, the moon, planets, comets are unthinking objects. They do

not react to being observed. The elderly, by contrast, have minds of their own. They assess circumstances and react according to their evaluation. Each person has a mind of his own, so no two people react to any event (whether it be service provisioning or age related problems) in exactly the same way. Here measurements are prone to systematic and random errors. This limitation is not a failing, so to say, of gerontological enquiry. It is a function of the fact that the elderly we study are thinking, creative and spontaneous beings.

2. **People Respond to their Surroundings, the Presence of a Researcher may affect the Behavior of the Elderly Being Studied.** Most people, and more so, the wily elderly react to being observed. This does not apply to physical research. The astronomers gaze has no effect on the distant moon, comet or planet. A basic challenge of social research is that being observed affects low people behave. Researchers can never be certain precisely how this will occur. When the elderly is being studied, some may change their disposition to Ok or bad when it is really otherwise. Some elderly may not want their family to be seen in a bad light and so put up every pretence to distort reality when being observed. Furthermore, people being watched may become anxious, angry or defensive; others may be specially friendly or helpful. The act of studying people can cause their behavior to change and this is a serious challenge to measurement.

3. **Because Social Scientists are Part of the Social World they Study, they can never be 100 Percent Value-free when Conducting Social Research.** The "test tube" of social scientists is the society they live in. Therefore, social scientists may find it difficult to control-or even to recognize- personal values that may distort their work. Many researchers are grossly affected by racism, value-system, culture and social background that they intentionally or unintentionally distort reality. All these are not so in physical research. Chemists are rarely personally affected by what goes on in their test tubes.

4. **Social Patterns Vary, what is true in one Time or Place may not hold in another.** As people are physically different so do their behaviours differ from one another. Human behaviours are so variable that there are no universal sociological laws. But the laws of physics will apply tomorrow as also yesterday, and they hold true in Australia as in Africa. Cultural differences also promote variations in behaviours.

Recommendations

How do we, gerontologists, reduce these challenges and ensure reliable and valid measurement? One way to limit distortion caused by personal values is replication, that is, repetition of research by other investigators. In spite of our personal values, if other researchers repeat a study using the same procedures and obtain the same results, we can gain confidence that their results are accurate - both reliable and valid.

Remember, science assumes that reality is "out there" scientists need to study this reality without changing it in any way. Objectivity should be the watch-word of every researcher whether gerontological or physical. Objectivity means personal neutrality in conducting research. Objectivity implies that researchers carefully hold to scientific

procedures and do not let their own attitudes and beliefs influence the results. The scientific objectivity of course is an ideal rather than a reality. The ideal of objectivity is to keep a professional distance or sense of detachment from the results, however they turn out. With this in mind, you should do your best when conducting research to see that conscious or unconscious biases do not distort your findings. One extra precaution is that researchers openly state their personal values in their research report so that readers can interpret the conclusion with those considerations in mind.

References

Allen, R. & Babbie, E. R. (2005). *Research methods for social work* Belmont:/Cole.

Babbie, E. (2010). *The practice of social research.* Belmont: Wadsworth.

Bostwick, G. J., Jr. & Kyte, N. S. (1993). Validity and reliability. In R. M. Grinnell, Jr. (Ed.) *Social work research and evaluation* (4th ed.) 174 – 197, Itasca,/L: F. E. Peacock Publishers.

Carmines, E. G. & Zeller, R. A. (1979). *Reliability and validity assessment*. Beverly Hill, CA: Sage.

Erber, J. T. (2005). *Aging and older adulthood*. Belmont: Thomson Wadsworth.

Gabor, P. A. & Grinnell, R. M., Jr. (1998). *Evaluation and Quality improvement in the human services. Boston*: Allen and Bacon.

Lindsey, L. L. & Beach, S. (2004). *Sociology*. New Saddle River: Pearson Education Inc.

Macionis, J. J. (2010). *Sociology*. Upper Saddle River: Pearson Education Inc.