International Journal of Business and Management Review

Vol.7, No.3, pp.27-43, April 2019

Published by European Centre for Research Training and Development UK (www.eajournals.org)

AN ASSESSMENT OF THE RELIABILITY OF SECONDARY DATA IN MANAGEMENT SCIENCE RESEARCH

OLABODE, Segun Oluwaseun[1*]; OLATEJU Olawale Ibrahim[2]; BAKARE Akeem Abayomi[3]

1. Department of Management Technology, Faculty of Management Sciences, Lagos State university, Lagos, Nigeria.
2. Department of Management Technology, Faculty of Management Sciences, Lagos State university, Lagos, Nigeria.
3. Department of Management Technology, Faculty of Management Sciences, Lagos State university, Lagos, Nigeria.

ABSTRACT: *Secondary data (SD) provides major advantage in the use of existing data sources, with large amounts of information, at relatively cheaper cost and easily available for research purposes. Even some researchers argue that millions of person-years of experience in the database will be available through SD, which would be impossible to collect in prospective studies. But an unreliable data could impede on the quality of research results and conclusions. The study critical examination of literature has identified tools that can aid the assessment of SD reliability. The study believes that the use of the adjusted inter-raters/observer as proposed by the study will add value to the method of assessing the reliability of SD, because of it use of statistical tools to directly estimate the available data. The study also believes that this will serve as a base for other researchers to improve on the study of assessing the reliability of secondary data.*

KEYWORDS: *Inter-raters/observer, Reliability, Secondary data, Validity*

INTRODUCTION

The quality of data (primary or secondary) utilised in any research determines the outcome of the research and its importance for further research work and relevance to business or statistical institutes. Thus, the quality of the enormous data collected daily by relevant organisations and/or individuals (e.g. government agencies, universities, private organisations, non-profits, think tanks, public opinion polls, and students) in recent years should be of importance to any system/institution especially the academic environment. Most times, vast amounts of primary data are collected and archived by relevant institutions or researchers at points in time all over the word. These results have made more prevalent the possibility of utilizing exiting data tor research at a later point in time-i.e. use of secondary data (Andrews, Higgins, Andrews, Lalor, 2012; Johnston, 2014; Smith, 2008).

Depending on the researcher's perception, the term “secondar datal” (SD) are data or information that was either gathered by someone else (researchers, recognized organisations acceptable to a system, etc.) for records or other purpose than the one currently under consideration, or often a combination of the two (Cnossen, 1997; McCaston, 2005) and is thus sometime referred to as “second-hand" data. For the first researcher they are primary data, but for the second researcher,

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

they are secondary data (Peter & Piet, 2012). To Weijun (2008), SD include both raw data and published summaries.

SD sometimes save the researcher the time that would have been spent on the field collecting data and, accessing the area under study. It can provide relatively large database of good quality, that may not be feasible for any individual researcher to collect. However, some researchers in business and management studies especially indicants as proxies for constructs, perhaps due to concerns over the possibility of been outdated, inaccurate or validity issues (Houston, 2004; Houston & Johnson, 2000; Schutt, 2006).

Obtaining SD today could be a relatively routine and easy process depending on the environment (i.e. how often such environment updates its records and what kind of records are available). SD may be quite expensive, however, the upfront costs s c as registration fees, have dropped with the emergence of the World Wide Web and the increase in the numbers of digitally published research sites. (Routledge, 2004).

Depending on the environment in which the SD is collected and the purpose of collection, SD can be beneficial especially to management sciences, for example, SD collected from academic publications may have a high degree of background work needed for the present research in the literature reviewed. Its use in such publication could already have promoted the data in media and management academics environment. Hence it could make its pre-established degree of validity and reliability need not be re-examined by the researcher or environment who is re-using such data. It could also in some cases, be a baseline for comparison with collected primary data results to determine the originality of the present data (Management Study Guide, 2016).

However, SD has its own shortcomings as identified earlier. The SD may be outdated, inaccurate or have validity issues. It may not be relevant to the population under examination, or detail enough. For example, an administrative data, transactional data or data from the Internet, which is not originally collected for research, may not be available in the usual 'research formats' or may be difficult to get access to. This exposes researchers to possible errors that can affect the quality (reliability and validity) of the data and invariably affects the viability of the research.
This study therefore, believed a critical examination of the concept, and assessment tools in the reliability of secondary data is essential to aid management research.

Statement of the Problem

As identified earlier, SD may be advantageous especially in term of cost as a result of the large database it can provides for management research innovation, productivity, and drawing conclusions in academics' research. However, when utilizing SD to help draw important conclusions in academic research, failing to check the reliability of that data could lead to inaccurate analyses and inappropriate research findings and conclusion. This may be due to some of the following. First, with today's accessibility to data via the internet, anyone can publish anything from anywhere (Stewart, 2014). Secondly, some organisations fraudulently manipulate information to give investors and client an impression that may not reflect their true state, some organisations don't post or give out detail information/data needed for comprehensive business and management research (especially details of their working capital and other financial variables

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

that can aid a comprehensive research) (Bankole, 2003; James & Oyeniyi, 2017; Shabnam, Zakiah, & Mohd, 2016; Vlad, Tulvinschi & Chirita, 2011). Thirdly, as identified by Babbie (2010); Cowton (1998); Flintermann (2014); using SD means possibility of inappropriateness of data in research, little or no control over how research data was generated and collated, possible modification of data by a researcher, a potential poor documentation, that could make the data neither valid or reliable. Hence, the use of SD based on face value without checking for potential errors and bias before it is used (Flintermann, 2014) or determining the reliability cannot be trusted for business and management research. This corroborates the theory identified by Priezkalns (2016) that some researchers believed only primary source of data can be trusted. This situation is one of the several reasons some researchers in the academic field of business and management avoid SD sources in their research.

Though several researchers have tried to examined how to improve researchers confidence in the use of SD, by developing tools/methods of assessing SD. There have been drawback as a result of limited literature. Thus, Flintermann (2014) is of the opinion that, available literatures have not been able to identify a suitable tool/method for the assessment of the reliability and validity of SD. Despite the various literature identified by few researchers in this area, there is a dearth knowledge of how secondary information is correlated with the primary data in business and management research and the solution to it. This also corroborated the belief of Andrews et al., (2012); Johnston, (2014); Smith, (2008) that there remains a dearth of literature that specifically addresses the process and challenges of conducting a reliable secondary data analysis research. Thus taking a critical examination of the above, this study provides an exploratory of the available -tools available in management sciences to determine the reliability of SD.

Objectives of the Study

i) To examine the concept of secondary data, validity and 'reliability' of secondary data.
ii) To assess the available tools for determining the reliability of secondary data in management/business research.
iii) To identify which criteria can be used to assess the reliability of secondary data in management/business research.

LITERATURE REVIEW

The study examined various work from literature because the conceptual frame work needed to address the first objective of the study.

Conceptual Framework

According to Johnston (2014), the concept 'secondary data analysis' was first identified by Glaser's in the discussion of re-analyzing data; i.e. data which were originally collected for other purposes. Weijun (2008) is of the opinion that SD include both raw data and published summaries. To Weijun, most organisations collect and store a variety of data to support their operations. These data are available only in the format the organisation that produce want it, thus most likely required negotiation for it to be accessed. Researchers like Bankole (2003); Oyeniyi, Obamiro, Abiodun, Moses, & Osibanjo (2016); believe that SD is an existing information whose main source is from primary sources. To Boslaugh (2007) the difference between SD and primary

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

data depends on the relationship between the individual/research team who collected a dataset and the researcher who is analyzing. Boslaugh (2007) concept is an important one because the same data set could be primary data in one analysis and secondary data in another depending on the time interval, purpose and environment. For example, three researchers A, B & C examined the relationship between two research variables. While A used system D as a case study, B used system E as a study area. While A & B collected data on the field, researcher C analysed data collected from A & B for comparison of the relationship in the two environment within the same time frame. Since data collected by A & B from the field (primary data) were for different purpose, the same data given to C will be seen as a secondary data.

Hakim (as cited in Johnston, 2014) believed secondary data analysis is any further analysis of an existing dataset which presents interpretations, conclusions or knowledge additional to, or different from, those presented in the first report on the inquiry as a whole and its main results. Irrespective how researchers or professional alike conjure the definition or concept of SD, the time interval that differentiate between the original purpose of the data collection and later purpose differentiate it from primary data. This, is in convergence with the view of Watson (2013) that see SD as analytical works that comment on and interpret other works from primary sources and are thus "second hand, published accounts, because they are created after primary sources and they often use or talk about primary sources.

Data Reliability (DR) is a concept every researcher, especially in business, management, social sciences and basic sciences, are aware of (Shuttleworth, 2009). To Shuttleworth, (2009) it could be a way of maximizing the inherent repeatability or consistency in collated data. For maintaining reliability internally, a researcher will use as many repeat sample groups as possible, to reduce the chance of an abnormal sample group skewing the results. For example, if three replicate samples for each analysis, and one generates completely different results from the others, then there may be something wrong with the data collated.

To Golafshani (2003), it is the extent to which sampled research results are consistent over time and the accuracy of representation of the total population under study'. If there results of a study can be reproduced under a similar methodology, then the research instrument is considered to be reliable. This concept is not different from the works of most researchers like Bankole (2003); Oyeniyi, et al (2016); Phelan & Wren (2006); Roberta & Alison (2015). They described reliability as the consistency between independent measurements of the same phenomenon or consistency of a measuring instrument to produce the same result repeatedly when applied to the same object. This concept shows that, reliability is a measure of the level of consistency of the research instrument and not the data. Though, the initial data generated at different period with the instrument is used to assess the reliability of the instrument. This concept is in convergence with the concepts of researchers like Babbie (2010); Flintermann (2014); Pierce (2009); Tasic & Feruh (2012) etc. To these researchers, reliability is the degree to which a research instrument or process consistently yields the same results under the same conditions, regardless of how many time the process is repeated, or degree to which a researcher can rely on the source of the data and therefore on the data itself. Thus to Flintermann, researchers can improve the reliability of their research instrument by repeatability and increasing its internal consistency. He further identifies

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

the following as cited in the work of Golafshani, (2003) that reliability can be estimated by using the following tests especially in quantitative research:

- Inter - Rater/Observer reliability: Degree to which different raters/observers are giving the same answers or estimates
- Test-Retest Reliability: Consistency of a measure over time
- Parallel - Forms reliability: The reliability of two tests constructed the same way, from the same content
- Internal consistency reliability: Consistency of results across items, often measured as Cronbach's Alpha

In science, the definition is the same, but needs a much. narrower and unequivocal definition. Thus, Shuttleworth (2009) believed that just as in sciences reliability is extremely important externally. This is because in science, the theory is that another researcher should be able to perform exactly the same experiment, with similar equipment, under similar conditions, and achieve exactly the same results least the design is unreliable. For example, the cold fusion case, of 1989 where Fleischmann and Pons announced to the world that they had managed to generate heat at normal temperatures, instead of the huge and expensive tori used in most research into nuclear fusion. These findings shook the world, but other researchers that attempted to replicate the experiment, experience no success. Thus the conclusion is that, it is either the researchers lied, or genuinely made a mistake. Neither of the conclusion is unclear, but their results were clearly unreliable.

Just as Shuttleworth (2009) identified the similarities in concept, this study summarized the concept of data reliability as stated by Adefioye (2016) as the consistency, ability and repeatability of results i.e. the result of a researcher is considered reliable if consistent result have situations but different circumstances. It can also be overall consistency, accuracy and completeness of a measure of repeatability of findings from processed data, given the uses they are intended for. In this context, reliability means that data are reasonably complete and accurate, meet the intended purposes, and are not subject to inappropriate alteration.

- Completeness refers to the extent that relevant records are present and the fields in each record are populated appropriately.
- Accuracy refers to the extent that recorded data reflect the actual underlying information.
- Consistency, a subcategory of accuracy, refers to the need to obtain and use data that are clear and well defined enough to yield similar results in similar analyses (Adefioye, 2016).

It should be of note that, while researchers like Adefioye (2016) used the construct 'accuracy' in term of the actual underlying information within the stem understudy, others like Oyeniyi, et al (2016) used it in term of expected underlying information within the system understudy. Thus, Oyeniyi, et al (2016) believed that accuracy cannot be used to conceptualize reliability. For example, a critical examination of the faulty wrist watch example stated by Oyeniyi, et al (2016) shows that, the time it read is consistence with the information that it will always be ten minutes late but not consistence the expectation that a normal wrist watch will read the actual time. Hence,

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

depending on the perception of the researcher the concept of reliability might show slight variation in the use of constructs. But one unifying construct in the concept of reliability is 'consistency'.

Assessing secondary data reliability can entail reviewing existing information about the data, which may include interviewing officials of audited Organisation; performing simple analysis on the sample of data, including advanced electronic analysis; tracing to and from source documents; and reviewing selected system controls (Shuttleworth, 2009). This collaborate Corillo (2014) who argue that, an assessment of the reliability of data will involve an assessment of the method(s) used to collect the data. Corillo (2014) also argue that, it will depend on the source of the data been assessed. For example, for documentary source, it is unlikely that there will be a formal methodology describing how the data were collected. But in report attention is given to how the data were analysed and how the result are report.

Flintermann (2014) argued that researchers improve on the quality of their research if they can assess the reliability of the research instrument or process used to generate and collet data. But he stressed that, this does not only depend on source of data but also depends on if it is a quantitative and qualitative research. To Flintermann (2014, while researchers believe in the reliability of research instruments or process used in quantitative research, there is little or no acceptable criteria in the assessment of the reliability of qualitative research be it primary or secondary data.

According to Flintermann, (2014)
"Without the certainty of numbers and p-vetues, qualitative research expresses a loss of confidence within and outside the field. Instead of explaining how reliability can be attained and estimated; leading qualitative researchers either suggested the adoption of new criteria or argued that reliability is an issue solely belonging to the quantitative research. As much as researchers and methodologists agree upon the definition and measurement of reliability in quantitative research the less agreement exists in qualitative research. From a quantitative point of view/ reliability and its measurement is clearly defined. In qualitative research the answer to what reliability is and how to measure it is not as clear/ as many discussions exist".

Thus researchers like Golafshani (as cited in Flintermann, 2014), are of the opinion that, though the concept of reliability is used for both qualitative and quantitative research, the most important test of a qualitative study is its quality, if researchers take the idea of testing as a way of retrieving information. But to other researchers like Stenbacka (as cited in Flintermann, 2014) the concept of reliability is not applicable or pertinent or it even giving the wrong impression in qualitative research as it is difficult to differentiate between the researcher and method used. He further stressed that, the level of consistency required in quantitative research does not have any value in qualitative research. Stenbacka (as cited in Flintermann, 2014) then concluded that, rather than discussing reliability of qualitative research, it is better for researchers to make the whole process (preparation, data gathering, analysis) visible. Hence, Morse, Barrett, Mayan, Olson & Spiers (2002) identified new terms that can be introduced as parallel concepts of reliability in qualitative research. These terms are consistency, confirmability and dependability. To them, consistency

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

can be achieved when the research process can be verified from the raw data collection over data reduction to the findings. While confirmability refers to the degree to which researchers actually arrived at their research findings and interpretations or degree to, which others can confirm results (Flintermann., 2014; Koch, 2006)

Irrespective of the new terms used to define reliability and under whichever type/source of research, the above examination of the concept, of SO reliability shows the following;

- Reliability of research instruments/process is of vital concern as it is seen as a sign of generating quality data research finding and conclusion.
- Researchers determine the reliability of research instrument/process using initially generated data, with the believe that a reliable instrument will generate a reliable data
- While it is easier to assess the reliability of SD in quantitative research the reverse is the case for qualitative research.

Also, Flintermann (2014) proposed that reliability of research instrument/process especially qualitative research, can be increased if the researcher can provide an insight into how findings and interpretations were achieved, repeatability of the research (if necessary) and describing changes in procedures.

It should be noted that, since researchers determine the reliability of research instrument/process .and not the data (Wayne, 20 14), the situation is different ·for both types of data. This is because it is easier to assess the, reliability of basic primary data collection instruments like Questionnaire. Interview, Observation and Reading (Annum, 2017) because of the availability of initial run of data, But assessing an existing 'document (in the form of government publication, earlier research, personal records and clients' records, Vivek, 2011), or non-document (in the form of tape and video recordings, pictures, drawings, films and television programmes, DVO/CD, Weijun, 2008) the tools available from which secondary data are gathered may not be easily execute with direct statistical tools .. This may be assessed by the use of a step by step' method of analysis.

Theoretical Framework

The study examined various 'work from' literature 'because the theoretical framework needed to address the second and third objectives-of the study. In this' regard, 'the common errors influencing reliability, two theories (Delphi and Triangulation theories) used to improve an existing model, and some models developed to assess the reliability of secondary data are examined. The study then concluded with a list of criteria for the evaluation of reliability of SO in research as identified in research and the study proposed model.

Also as identified earlier, Flintermann (2014) opinion that available literatures have not been able to identify a suitable tool/method for the assessment of the reliability and validity of SD. This also affect the availability of enough theory on the study. This study therefore used a method of conceptual derivative and importation of theories from other field to explain existing models assessed in the study.

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

Methods of Estimating Research Instrument Reliability

As identified earlier in the study, Inter-Rater/Observer, Test-Retest, Parallel - Forms, and Internal consistency are the basic tools of estimating research instruments reliability. Leading researchers like Adefioye (2016); Bankole (2003); Oyeniyi, et al (2016) etc. believed these tools are basically used to estimate reliability of primary data research instruments. But a careful observation of the Inter-Rater/Observer method by the study observed that, it can be used to estimate the reliability of data directly and not just the research instrument. This made, it a, possible method of assessing the reliability of SD since it may not necessarily require an initial run of data.

According to most researchers Inter-Rater/Observer involved' the use of human expert as a part measurement procedure, to assess the consistency and invariably reliability of data, For example, a researcher that required and collected working capital of an organisation(s) as data for a study, can assess the data reliability. This is by estimating the consistency in the responds of two expert observers regarding the possible level of the degree of error and biasness.

William (2006) identified two methods to actually estimate inter-rater reliability First, if the measurement consists' of categories the raters check off which category each observation falls in the researcher calculates the percent of agreement between the raters. To William (2006), it may be seen to some researchers as a crude measure, but it does give art idea of how much agreement exists, and it works no matter how many categories are used for each observation. Second if the measurement is a continuous one This involve calculating the correlation between the ratings of the two observers. The correlation estimate will determine the reliability or consistency in the responds of the raters and invariabty that of the data.

In other to improve this method, two adjustments are proposed to modify the Inter-Rater/Observer method for assessment of the reliability of secondary data. The modification could be based on the Delphi theory and the triangulation theory. This is because these theories are similar to the existing inter-rater/observer method but allow the use of more than two experts. This will involve the use of' coefficient of multiple correlation statistical tool to evaluate the degree of consistency of the expert.

The Two Adjustments Proposed

As identified earlier, instead of restricting it to two expert observers, more than two observers could be used. This will involve calculating the coefficient of the multiple correlations between the ratings of the observers. The higher the result the more reliable the data. It can also be done using the triangulation theory. The theory involves the use of multiple independent source of data to establish the truth and accuracy of a claim. Hence, can be used to assess the validity of data instruments and data.

This can invariably establish the reliability of data as researchers like Oyeniyi, et. al (2016); Wayne (2014} etc. have identified that, though, a reliable data does not necessarily mean a valid data, but a valid data mean's a reliable data. Hence assessing the validity of a SD means assessing its reliability.

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

The Triangulation Theory

To Sagor (2000), the triangulation theory is similar to how legal practitioners or researchers (defense lawyers and prosecutors) convince a jury of the essential truth and accuracy (validity and reliability) of their cases. This is done through the twin processes of corroboration and impeachment. To convince a jury to believe their witnesses, another independent witnesses is brought in. As an additional witness corroborates the first witness, it increased the confidence the juror will have in the initial testimony. The more independent testimony from witnesses that support the initial witness before a jury, the more the jurors will trust the truthfulness and accuracy of the claims. Conversely, the reverse is the case if lawyers want the jury to doubt the truth and accuracy" (validity and reliability) of the other side, they try to impeach (challenge the credibility of) the testimony of as many as presented by the lawyer. Thus if as many as possible expert can pass a consistent judgment on a set of data rated/observed then the validity and reliability can be established.

The Delphi Theory

Using the Delphi theory, a structured communication technique initially developed as a systematic, interactive forecasting method. Experts will assess the data in two or more rounds. After each round, a facilitator provides an amonized summary of the experts ratings from the previous round as well as reasons for their judgments. Thus, experts are encouraged to review their rating/observation in view of the responds of the other experts. The process is deemed optimal after an acceptable level of consistency in their responds. The coefficient of multiple correlations is then used to assess the reliability of final rating/observation of the experts. The higher the result the more reliable the data.

The above tool/method of assess SD reliability is slightly different from other tools/methods developed by most researchers. This is because, the above shows a direct use of statistical tools/methods in assessing the reliability of SD. But most researchers use an indirect qualitative step by" step methods as tools for assessing the reliability of SD. Some of these tools were identified in the studies examined in subsequent paragraph of this study.

FAO in their evaluation of the quality of both the source of SD and the data itself; categorized the problems that may reduce quality as shown in table 1 below. The organisation is of the opinion that if SD is analysed for each category, then the quality of the data can be improved. FAO also presented a flow chart as shown in figure 1 depicting the decision path that should be followed when using secondary data. The flowchart has two phases. The first phase relates the relevance of the SD to the research objectives. The second phase is concern with questions about the accuracy of SD.

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

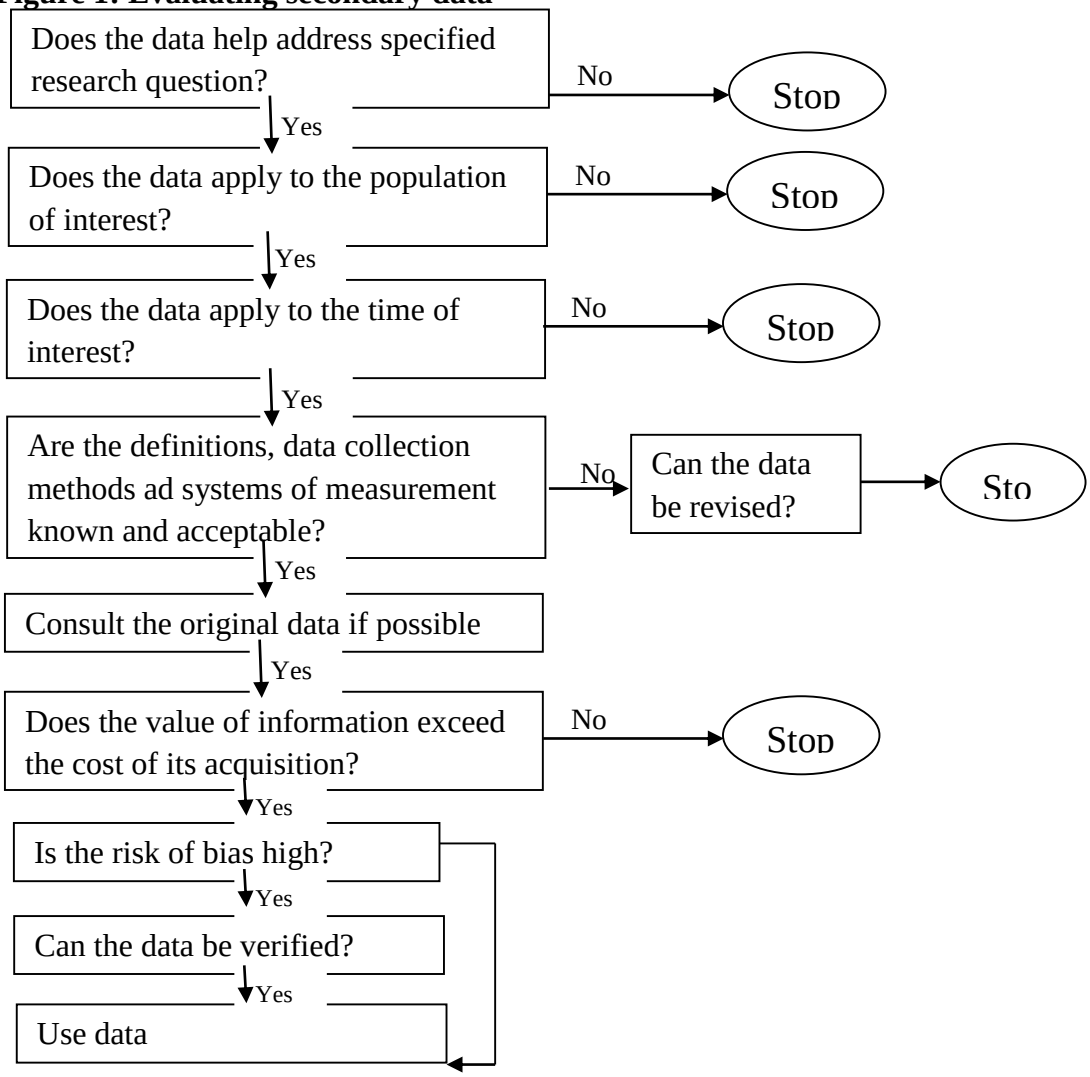
Table 1: Categories of Problems that may Reduce Quality

Definitions	Researchers have to be careful of variables definitions when making use of secondary data. For example, researchers with interest in rural communities and their average family size. If published statistics are consulted, then a check must be done on how terms such as "family size" have been defined. They may refer only to the nucleus family or include the extended family. It should be noted that definitions may change overtime and where this is not recognized erroneous conclusions may be drawn. Geographical areas may have their boundaries redefined, units of measurement and grades may change and imported goods can be reclassified from time to time for purposes of levying customs and excise duties.
Measurement error	When a researcher conducts fieldwork she/he estimates inaccuracies in measurement through the standard deviation and standard error, which are sometimes not published. They may require speaking to the individuals involved in the collection of the data to obtain some guidance on the level of accuracy of the data. The problem is sometimes not so much 'error' but differences in levels of accuracy required by decisions makers. When the research had to do with large investments in, say, food manufacturing, management will want to set very tight margins of error in making market demand estimated. In order cases, having high level of accuracy is not so critical. For instance, if a food manufacturer is merely assessing the prospects for one more flavor for a snack food already produced by the company then there is no need for highly accurate estimates in order to make the investment decisions.
Source bias	Researchers have to be aware of vested interests when they consult secondary sources. Those responsible for their compilation may have reasons for wishing to present a more optimistic or pessimistic set of results for their organisations. For example, for officials responsible for estimating food shortages to exaggerate figures before sending aid requests to potential donors. Similarly, and with equal frequency, commercial organisations have been known to inflate estimates of their market shares.
Reliability	The reliability of published statistics may vary over time. It is not uncommon, for example, for the systems of collecting data to have changed over time but without any indication of this to the reader of published statistics. Geographical or administrative boundaries may be changed by government, or the basis for stratifying a sample may have altered. Other aspects or research methodology that affects the reliability of secondary data is the sample size, response rate, questionnaire design and modes of analysis.
Time scale	Most censuses take place at 10 years intervals, so data from this and other published may be out-of-date at the time the researcher wants to make use of the statistics. The time period during which secondary data was first complied may have a substantial effect upon the nature of the data. For instance, the significant increase in the price obtained for Uganda coffee in the mid-90's could be interpreted as evidence of the effectiveness of the rehabilitation programme that set out to restore coffee estates which had fallen into a state of disrepair. However, more knowledgeable coffee market experts would interpret the rise in Uganda coffee prices in the context of large scale destruction of the Brazilian coffee crop due to heavy frost, in 1994, Brazil being the largest coffee producer in the world.
Sources of data	Whenever possible, researchers ought to use multiple sources of secondary data. In this way, these different sources can be cross-checked as confirmed of one another. Where differences occur an explanation for these must be found or the data should be set aside.

Source: Adopted from FOA Corporate Document Repository: Agriculture and Consumer Protection, (n.d.)

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

Figure 1: Evaluating secondary data



Source: Adapted from FOA Corporate Document Repository: Agriculture and Consumer Protection, (n.d.)

A critical examination of the above simple flow chart shows some simple deficiencies. One, it did not state what next if the data can be revised. Two, the researcher has to determine its tool(s) of analyzing the risk of bias. Three, it did not examine the possibilities if error in measurement of variables etc. but as identified in the last category in the study of FAO, researchers are advised to use multiple sources of secondary data. Similarly, the use of one or more other tools of analysis with the flow chart could improve the assessment of reliability of SD. For example, table 2 shows, a Flintermann, (2014) summary of errors and issues that may have an impact on the use of SD divided into different categories as identifies by (Tasic & Feruh, 2012). A combination of table 2 as an assessment tool with the flow chart can help researchers overcome some of the deficiencies in the flow chart in improving the quality of research result that use SD.

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

Table 2: Errors and Issues in Secondary Data

Errors and issues in secondary data	Caused by	
Errors that can invalidate data	Manipulation	The organization gathering data may manipulate/reorganize data to meet a purpose unknown to other. Collecting agency may want to show that the organization goal is met.
	Inappropriate, confusion or carelessness	a) Organization might collect, organize and distribute data without properly specifying the particulars of the collection process or assembly procedures b) Organization may not care about data quality or validity c) Organization's staff may not know how to collect data
	Concept error	Concept error arise because of the difference between the concept to be measured and a specific item that is used to measure a concept. Data containing error can still be use, however, only if something is known about the nature of the error.
Errors requiring data reformulation	Changing circumstances	Changes affecting data series which are not readily apparent in that data series. e.g. change in geographical boundaries, change in underlying unit of measurement
	Inappropriate transformations	Original data in secondary data sources is often presented in categories or tables that make the data more presentable or the original categories do not reflect an analyst's needs to handle the task at hand.
	Inappropriate temporal extrapolations	Secondary data often not available for intervening periods between published reports. Data for these periods need to be interpolated from the nearest two reporting years. Not knowing the true change between two these two points, any answer can be obtained for the point of time in question.
	Inappropriate temporal recognition	Arising from a misunderstanding of the time dimension of secondary data. There is always a time lag between the gathering of primary data and the time when it is made available.
Errors reducing reliability	Correct(ed) data	Data can be inconsistent form one report to another in the same published series because of errors that have been discovered, corrected and then reflected in subsequent version of the data set. Or publisher of secondary data can adjust forecasts for a decimal year against actual census numbers.
	Changes in collection procedures	Occurs due to different methods or circumstances surrounding the collection, e.g. time of collection, way of summarizing data. Generated data can be quite different from previous data in the dame data set.
	Clerical errors	Occurs because of the transposition of numbers in a series with the same number of digits or the misplacing a decimal. Outliers can be easily detected by creating diagrams or tables.

Source: Adapted from Flintermann (2014)

Flintermann (2014) is of the opinion that available literatures have not been able to identify a suitable tool/method for the assessment of the reliability and validity of SD. But based on study theoretical framework, Flintermann (2014) developed a set of criteria in five categories as shown in table 3 for the assessment of reliability of SD. The table also shows indicators of reliability or validity and the level of reliability or validity if these indicators are found.

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

Tables 3 Criteria for assessing reliability/validity in a market research report

Indicators for Reliability and Validity		**Level of reliability and validity**	**Level of validity**
Clear specification of data collection and data analysis	Detailed description on which type of research was used	High	High
	Definition of research variable	High	High
	Sources used stated	High	High
	Information on how it dealt with missing data	High	High
	Date of collection available	High	High
	Information on how quality of data used is controlled	High	High
	Transforming data (from raw data to result)	High	High
	Information about method used e.g. statistical tests	High	High
	Coding of data whether and how?	High	High
	Clear organisation of data	High	High
	Contact data presented	High	High
Clear specification about potential changes in procedure	Information about changes in methods used from one study to another	High	High
	Information about changes in sources used from one study to another	High	High
	Information about changes in definitions used from one study to another	High	High
Updates	Due to error correction – information given	High	High
	Due to new version of report	High	High
Result of comparing data collected out the research concepts with the actual research concepts data	Dataset similar to each other	High	High
Missing research variables report	Data for missing variable could be found using other sources that the research	Low	Low
	Data for missing variables could partly be found using other sources that the research	Medium	Medium
	Data for missing variable could not be found using other sources that the research	High	High

Source: Adapted from Flintermann, (2014)

Clear specification of data collection and data analysis

This category refers to how the process of data collection and analysis are described. Government recognized agencies, private organisations, or researches involve in the collection, organisation and distribution of data may poorly specify the particulars of the collection process, the data procedures (methodology), insight into the whole process (gathering, collection, analysis) and sources used. The specification of these indicators can make it easier to replicate the study using the same procedures. In qualitative research, consistency and conformability can be achieved by stating the whole process of data preparation, gathering and anaivsis. If a researcher presents the whole research process, reliability is considered to be high (Flintermann, 2014).

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

Clear specification about potential changes in procedures
This categories deal with possible changes in procedures. Reliability is at an on acceptable high level if the information on changes in methods, sources or definitions can be found and judgements on how these affect the data.

Updates
This category comprises of updates, which are either made due to error correction or new methods of data presentation or new version of a research reports. Data can be even inconsistent from one report/presentation to another in the same published series because of errors being discovered and corrected in subsequent versions. For example, different researchers using data collected from the same organisation shows obvious and unacceptable variation In input data.· Corrections of errors may also be caused by using inappropriate methods or sources or by using methods for data processing in an incorrect manner.

Comparing data collected outside of research scope to research scope
Theoretically, sources outside of research scope should provide data that is similar/equal to research scope data since reliability of an instrument is expected to yields the same results regardless of the number of repetitions. Hence, a search for data about the automotive industry using the same definitions, years and units in a research scope should provide similar findings and trends. If this is the case, reliability of the data is strengthened.

Missing Research Reports
Missing data may affect the reliability of researchers' results if data collection, processing or storage are not properly coordinated or even forgotten to be obtained. If the final data obtained for a research was reached based on missing data, chances are high that decisions or conclusions are made based on biased data. Missing data can have a significant effect on conclusions drawn. If any information about missing data or processes, sources and methods used are clearly documented and hence a replication can be done to identify whether data is missing, reliability is increased.

As identified earlier, other researchers, like Koziol & Arthur (2012); Stewart (2014); Weijun (2008) also use an indirect qualitative step by step methods as tools for evaluating the quality of SD. But most of the steps identified by their work have been included in the earlier reviewed literature of the study;

CONCLUDING REMARKS

SD provides major advantage in the use of existing data sources, with large amounts of information, at relatively cheaper cost and easily available for research purposes. Even Henrik & Jorn (1996) argue that millions of persons experience in the data bases will be available through SD, which would be impossible to collect in prospective studies, But an unreliable data could impede on the quality of research results and conclusions. The study critical examination of literature has identified tools that can aid the assessment of SD reliability.

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

Of the tools identified, the study believed that the use of the adjusted inter-raters/observer as proposed by the study will add value to the method of assessing the reliability of SD, because of its use or statistical tools to directly estimate the available data.
The study also believed that, this will serve as a base for other researchers to improve on the study of assessing the reliability of secondary data.

References

Adefioye, T. (2016). *Reliability abd validity*. Retrieved March 21, 2017, from https://www.lbs.edu.ng/sites/faculty_research/crle/Downloads/Notes%20on%20Reliability%20and%20Validity%20by%20Temilade%20Adefioye.pdf

Andrews, L., Higgins, A., Andrews, M. W., & Lalor, J. G. (2012). Classical grounded theory to analyse secondary data: Reality and reflection. *The Grounded Theory Review, 11*(1), 12-26.

Annun, G. (2017). *Research instrument.* Retrieved May 2, 2017, from http://campus.educadium.com/newmediart/file.php/137/Thesis_Repository/recds/assets/TWs/UgradResearch/ResMethgen/files/notes/resInstrsem1.pdf

Bankole, A. R. (2003). *Research methods: An introductory approach.* Lagos: Adeshina Press Production and Publication.

Barbie, E. (2010). *The practice of social research* (12th ed.). Belmon, USA: Wadsworth Cengage Learning.

Bill, G. (2015). *Is America abandoning research?* Retrieved February 7, 2017, from https://discoveryourtruenorth.org/1608-2/

Boslaugh, S. (2007). *Secondary datasources for public health: A practical guide.* London, U.K.: Cambridge University Press.

Carillo, P. R. (2014). *Secondary data analysis*. Retrieved May 10, 2017, from https://www.slideshare.net/honeyturqueza/secondary-data-analysis-report

Cnossen, C. (1997). *Secondary research: Learning paper 7*. Retrieved May 15, 2017, from https://www.coursehero.com/file/p39mt1eo/20-No-2-Cnossen-C-1997-Secondary-Research-Learning-Paper-7-The-Robert-Gordon/

Darren, L., & Gareth, H. J. (2013). *Introduction to research methods and data analysis in Psychology* (3rd ed.). London, U.K.: Pearson Education Limited.

Flintermann, B. (2014). *The quality of market research reports: The case of Marketline Advantage and the automobile industry*. Retrieved May 29, 2017, from http://essay.utwente.nl/66122/1/Flintermann_MA_MB.pdf

FOA Corporae Document Repository. (n.d.). *Marketing research and information systems*. Retrieved June 1, 2017

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The Qualitative Report, 8*(4), 597-606.

Harkness, S. (2004). *Social and political indicators of human well-being: Working paper*. Retrieved March 2, 2017, from https://www.wider.unu.edu/publication/social-and-political-indicators-human-well-being

Hassard, J. (2014). *Why are scientist abandoning their research?* Retrieved May 2, 2017, from http://www.artofteachingscience.org/why-are-scientists-abandoning-their-research/

Henrik, T., & Jorn, O. (1996). A framework for evaluation of secondary data sources for epidemiological research. *International Journal of Epidemiology, 25*(2), 50-62.

Herrera, Y. M., & Kapur, D. (2007). Improving data quality: Actors, incentives and capabilities. *Political Analysis, 15*(1), 365-386.

Houston, B. (2004). Assessing the validity of secondary data proxies for new constructs. *Journal of Business Research, 57*(2), 154-161.

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

Johnson, M. P. (2014). Secondary data analysis: A method of which time has come. *Qualitative and Quantitative Methods in Libraries (QQML), 3*(1), 619-626.

Johnson, S. A., & Houston, M. B. (2000). Buyer-supplier contracts versus joint ventures: Determinants and consequences. *Journal of Marketing Research, 37*(1), 1-15.

Kehinde, J., & Oyeniyi, O. (2017). Comment on mothly seminar series. *Lagos State University, Faculty of Management Sciences, Department of Business Administration*.

Koch, T. (2006). Establishing rigourin qualitative research: The decision trail. *Journal of Advanced Nursing, 53*(1), 91-103.

Koziol, N., & Arthur, A. (2012). *An introduction to secondary daa analysis*. Retrieved May 3, 2017, from http://r2ed.unl.edu/presentations/2011/RMS/120911_Koziol/120911_Koziol.pdf

Management Study Guide. (2016). *Secondary data*. Retrieved May 23, 2017, from http://www.managementstudyguide.com/secondary_data.htm

McCaston, M. K. (2005). *Tips for collecting, reviewing and analysing secondary data*. Retrieved May 13, 2017, from https://www.ands.org.au/__data/assets/pdf_file/0003/713235/Tips_for_Collecting_Reviewing_and_Analyz.pdf

Morse, J. M., Barrett, M., Mayan, M., Olson, K., & Spiers, J. (2002). Verification strategies for establishing reliability and validity in quantitative research. *International Journal of ualitative Methods, 1*(2), 13-22.

Norwegian Centre of Research Data. (n.d.). *Data quality and comparability*. Retrieved May 29, 2017, from http://www.nsd.uib.no/macrodataguide/quality.html

Oyeniyi, O. T., Abiodun, A. J., Moses, C. L., & Osibanjo, A. O. (2016). *Reesearch methodology: With simplified ste-by-step application of SPSS package.* Lagos: Pumark Nigeria Limited.

Peter, W. V., & Piet, J. D. (2012). *49 factors that influence the quality of secondary data sources: Paper quality and risk management.* Rotterdam, Statistics: Statistics Netherlands.

Phelan, C., & Wren, J. (2006). *Exploring reliability in academic assessment.* Cedar Falls: Iowa: University of Norhern Iowa.

Pierce, R. (2009). *Evaluating information: Validity, reliability, accuracy and triangulation.* Retrieved March 21, 2017, from https://www.sagepub.com/sites/default/files/upm-binaries/17810_5052_Pierce_Ch07.pdf

Priezkalns, E. (2016). *Revenue assurance: Expert opinions for communications providers.* Chicago: CRC Press.

Roberta, H., & Alison, T. (2015). Validity and reliabilty in quantitative studies. *ResearchGate, 18*(3), 50-62.

Sagor, R. (2000). *Guiding school improvement with action research.* Virgina, USA: Association for Supervision and Curriculum Development.

Schutt, R. (2006). *Investigation the social world.* New Jersey: SAGE Publications.

Shabnam, F. A., Zakiah, M. M., & Mohd, M. R. (2016). Detecting financial statement frauds in Malaysia: Comparing thr abilities of Beneish and Dechow models. *Asian Journal of Accounting and Governance, 7*(1), 57-65.

Shuttleworth, M. (2009). *Definition of reliability*. Retrieved May 13, 2017

Smith, E. (2008). *Using secondary data in educational and social research.* New York: McGraw-Hill Education.

Stenbacka, C. (2001). Qualitative research requires quality concepts of its own. *Management Decision, 39*(7), 551-555.

Stewart, C. (2014). *How to evaluate external secondary data*. Retrieved May 31, 2017, from https://blog.marketresearch.com/how-to-evaluate-external-secondary-data

Tasic, S., & Feruh, M. B. (2012). Errors and issues in secondary data used in marketing research. *The Scientific Journal for Theory and Practice of Socioeconomic Development, 1*(2), 326-335.

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)

Vanhatalo, E., & Kulahc, M. (2015). *Impact of autocorrelationon principal components and their use in statistical process control (SPC).* London: John Wiley & Son.

Vivek, A. (2011). *Instrument to measure primary and secondary data*. Retrieved May 29, 2017, from https://www.slideshare.net/VivekArora2/instrument-to-measure-primary-secondary-data

Vlad, M., Tulvinschi, M., & Chirita, I. (2011). The consequences of fraudulent financial reporting. *The Annals of the "Stefan cel Mare" University of Suceava. Fascicle of he Faculty of Economics and Public Administration, 11*(1), 264-269.

Watson, K. A. (2013). *Evaluating primary ans secondary sources*. Retrieved May 30, 2017, from https://www.nlj.gov.jm/rai/CSEC/Evaluating%20Primary%20and%20Secondary%20Sources.pdf

Wayne, L. D. (2014). *Terrorism, homeland security, and risk assessment via research proposal* (3rd ed.). Bloomington: Xlibris, LLC.

Weijun, T. (2008). *Research methods for business student*. Retrieved April 24, 2017, from https://eclass.teicrete.gr/modules/document/file.php/DLH105/Research%20Methods%20for%20Business%20Students%2C%205th%20Edition.pdf

William, M. (2006). *Research method knowledge base: Types of reliability*. Retrieved April 15, 2017, from https://www.socialresearchmethods.net/kb/reltypes.php

Print ISSN: 2052-6393(Print), Online ISSN: 2052-6407(Online)