Education 6610
Research on Computers in the Curriculum

A Synopsis of Data-Driven Research

Name: Rob Power
MUN ID: 9236571
Professor: Dr. Bruce Mann
Date: Nov 8, 2002

Introduction

Research is a systematic examination of natural phenomena, carried out in an effort to better understand that phenomena, and to add to our collective understanding of the world around us (Bieger and Gerlach, 1996, pp. xi-xiii; Leedy and Ormdod, 2001, pp. 3-11). While ideas or theories can spawn research proposals that involve the construction of controlled experiments, designed to collect the data needed to test those ideas, sometimes there is no need to run experiments or collect data (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68). Oftentimes, the data already exist, and the research is spawned as an effort to interpret or reinterpret that data (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68; Mann, 2002). What exactly is data-driven research, though, and how can research evolve from existing data? What constitutes good data-driven research? How is data-driven research useful, and who would use it? How, exactly, is data-driven research carried out? And what considerations must be taken when embarking on data-driven research?

What is Meant by Idea-Based Research?

There is a distinct difference between data-driven research and research that is idea-based. Idea-based research involved the construction of a scenario, or scenarios, for the purpose of collecting specific data that will be used to test the idea or theory that spurred the researcher into action(Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68). Such scenarios can involve the manipulation of variables in quantitative research, or the collection of more subjective observations, as in qualitative research. The researcher then interprets that raw data. What sets data-driven research apart from idea-based research is that the data already exists in some form, and the researcher turns to this data to carry out the research (Bieger and Gerlach, 1996, pp. 1-48; Crocker, 1998, *c*; Leedy and Ormdod, 2001, pp. 49-68; Mann, 2001).

As the phrase suggests, data-driven research can be driven by the existence of the data (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68; Mann, 2001). The researcher may look at the data and start asking questions about what it means. How are certain pieces of the data related to one another? Why are certain things happening with the data? What story does this data tell? These are all examples of questions driven by an examination of a set or sets of data.

In other instances, the data-driven researcher may be driven to find existing data to test an hypothesis or theory (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68; Mann, 2001). Or the researcher may be confronted with a specific issue or problem, and may be driven to seek out existing data to come up with answers, or to arrive at a better understanding (Bieger and Gerlach, 1996, pp. 1-48; Crocker, 1999, *c*; Leedy and Ormdod, 2001, pp. 49-68; Mann, 2001).

Although there are clear distinctions between idea-based and data-driven research, elements if the two categories are not necessarily mutually exclusive. Ideas are involved in data-driven research. A researcher may start with an idea or hypothesis, and turn to the data to either support or refute it (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68; Mann, 2001). A researcher may start with a more general idea, and use the data to help develop a clearer hypothesis or theory (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68; Mann, 2001). Or the researcher may simply examine the correlation between data to

generate an hypothesis (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68; Mann, 2001). And, as we shall see, there are similarities, as well as differences, between the uses of data-driven and idea-based research, the methodologies or carrying out such research, and the considerations that must be taken when involved in either type of research.

What Constitutes Good Data-Driven Research?

As with idea-based research, there are certain elements to consider when determining whether a research proposal constitutes good data-driven research. The first of these is that it should meet the general requirements of legitimate research (Bieger and Gerlach, 1996, pp. xi-xiii; Leedy and Ormdod, 2001, pp. 3-11). This means there must be an examination of an idea or theory, and of the relationships between data, as opposed to the simple transposing or reporting of data sets (Leedy and Ormrod, 2001, pp. 3-11).

Potential researchers should also consider such things as the availability of data to examine, the originality of their research, and the findings of previous research (Leedy and Ormrod, 2001, pp. 70-75). As with idea-based research, there is no point to expending the effort of finding and analyzing data if the answers are already known, unless the researcher is attempting to add credibility to previous findings by replicating them (Leedy and Ormrod, 2001, pp. 70-75). Replicating a study may also be warranted if the data sets to be examined have a more localized context (Crocker, 1999, *c*; Leedy and Ormrod, 2001, pp. 70-75).

Finally, researchers must consider the appropriateness of their research question, and of the use of data-driven research. Certain topics, such as comparing the effectiveness of computerized instruction to that of classroom teachers, are inappropriate (Mann, 2001). In other instances, relying upon existing sets of data may be inappropriate. The research question or problem itself may warrant an idea-based research approach, and the collection of new data where there is a manipulation of variables involved (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68).

How is Idea-Based Research Useful, and Who Would Use it?

The types of problems that can be examined through data-driven research are limited only by the needs of the researcher, and the availability of data. In fact, the availability of data can be seen as one of the more useful aspects of data-driven research. For some problems, especially in the context of education, the collection of new data may be cumbersome, impractical, and completely unnecessary (Bieger and Gerlach, 1996, pp. 1-48; Crocker, 1999, *c*; Leedy and Ormdod, 2001, pp. 49-68; Mann, 2001). Large bodies of data exist on everything from student enrollment to demographics, academic achievement, and the use of specific resources, to name a few (Crocker, 1999, *c*). For problems related to these statistics, the collection of new data may well be impractical, and the use of these large bodies of existing data can also enhance the credibility and generability of research findings (Bieger and Gerlach, 1996, pp. 1-48; Crocker, 1999, *c*; Leedy and Ormdod, 2001, pp. 49-68; Mann, 2001). This is especially true when comparing or contrasting data between different regions (Crocker, 1999, *c*).

Data-driven research, then, is extremely useful in the context of education (Crocker, 1999, *b, c*). It makes it possible to deal with the large volumes of information available on a

system as extensive as the education system. Data-driven research is thus a practical mode of research for individual educators or researchers who would otherwise be limited in time and resources (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68; Mann, 2001). It is also especially useful for educational policy makers, who must determine, from the "facts" of the system, policy on everything from funding allocations to the implementation and evaluation of new curriculum and teaching/learning strategies in the schools (Crocker, 1999, *c*). Policy makers, in many contexts, often rely upon research into existing data to justify their policy decisions, and to establish useful comparisons or benchmarks between their own policies, and these or similar groups or organizations in other areas.

Data-driven research, however, does not always involve the secondary examination of large bodies of existing data. Smaller sets of existing data are also useful and can come from numerous sources. Individual educators or researchers may want to examine data that they are already collecting or possess, from non-research purposes (Crocker, 1999, *c*; Mann, 2002). Comparing the academic performances of students on different elements of a course is one examples of this use of data-driven research. The elements of the course are known, and the results of student evaluation are already being collected. A researcher may turn to such data sets to examine whether relationships exist between certain course elements and academic achievement, or between achievement levels on different tests, assignments, or other forms of evaluation (Crocker, 1999, *c*; Mann, 2002). Aside from gaining a better understanding of issues or trends within a course, such data-driven research could be useful to help enhance the efficiency of teaching and learning methods, and to improve the overall learning experience (Crocker, 1999, *c*; Mann, 2002).

How is Idea-Based Research Carried Out?

As previously noted, data-driven research begins with two things: the existence, somewhere, in some form, of the data; and an idea, hypothesis, or problem for investigation (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68, 90-106; Mann, 2001). Either the researcher wants or needs to test an idea or hypothesis, or the researcher wants or needs to examine an issue, trend, or situation, for which a body of data already exists (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68, 90-106; Mann, 2001). From these elements the researcher would clearly define the problem to be investigated, much like a researcher would do in idea-based research. Like idea-based research, as well, the researcher may choose to define the problem in terms of more manageable sub-problems or tasks. These elements completed, the researcher is in a position to design the nature of the study itself.

While it may initially seem that the data-driven research designs are likely to be quantitative in nature, this is not always the case. Many bodies of data relied upon by data-driven researchers are qualitative in nature, and would thus require more subjective, qualitative methods of examination (Glesne, 1999, pp. 1-16). Even when dealing with strictly quantitative data, data-driven researcher must be aware of the need for qualitative approaches in their research. Meaningful examination of quantitative data may require the use of subjective, qualitative approaches in order to establish useful parameters for qualitative evaluation (Glesne, 1999, pp. 1-16; Mann, 2002). This is true, for example, when examining data that have a mix of quantitative and qualitative elements, such as class participation, or participation in online discussion forums (Mann, 2002). It is also true of policy research, where the often political goals

of the research can place subjective or qualitative parameters on scope and nature of the efforts even of impartial researchers (Crocker, 1999, *c*).

While the recruitment of research participants is a crucial stage for idea-based research (Leedy and Ormrod, 2001, 90-106), this may or may not be the case for data-driven research. In cases where the researcher is conducting secondary research on pre-existing data, it is not necessary to recruit and gain the consent of new participants (Crocker, 1999, *c*; Leedy and Ormrod, 2001, 90-106). However, in some cases, it may be necessary to gain the consent of the parties who own the original data sets (Memorial University of Newfoundland, *a, b*). And, depending upon the conditions assigned to the consent of the original research participants, it may sometimes be necessary to reaffirm their consent for the use of the data (Bieger and Gerlach, 1996, pp. 227-232; Leedy and Ormrod, 2001, pp. 107-111; Memorial University of Newfoundland, *a, b*). In other cases, the secondary data may be publicly available, or readily available within the research community, making it unnecessary to gain consent to use the data (Memorial University of Newfoundland, *a, b*).

The collection of the data itself, then would be the next step for any data-driven researcher (Leedy and Ormrod, 2001, pp. 90-106). This could be as simple a task as pulling the data from a filing cabinet, to conducting literature reviews, online searches, or requesting copies of the data from the relevant individuals, groups, or organizations.

What is of primary concern in data-driven research is the actual analysis or interpretation of the data (Crocker, 1999, *c*; Mann, 2002). This stage may involve a division of analysis into more manageable tasks, amongst researcher in many locations (one of the possibilities linked to data-driven research) (Crocker, 1999, *c*; Leedy and Ormrod, 2001, pp. 90-106). It may also involve the utilization of statistical software packages, such as Statistics Package for the Social Sciences (SPSS). Finally, it may involve the use of qualitative, subjective methods when dealing with non-quantitative data, or when establishing meaningful parameters for data interpretation (Glesne, 1999, pp. 1-16).

The last stage to conducting data-driven research is the publication of findings (Leedy and Ormrod, 2001, pp. 90-106). This may be done by submitting a research report to a scholarly publication or, in the case of policy research, to the appropriate group or organization. In such cases, the group who sponsored the research would publish the findings in the form of new policy, and documents to support that policy (Crocker, 1999, *c*).

What Considerations Must be Taken?

A number of considerations must be taken when conducting data-driven research. Like idea-based research, the first consideration here should be the type of research to be carried out (Bieger and Gerlach, 1996, pp. 35-38; Leedy and Ormrod, 2001, pp. 90-106, 148, 191-196). Will it be of a quantitative or qualitative nature. The answer depends on the type of problem under consideration, and the type of data available. It is difficult to quantify subjective non-numerical data sets without a certain degree of qualitative methodology (Glesne, 1999, 1-16). The researcher should also consider whether the available data is appropriate to the problem or issue, and whether it is extensive enough to give the problem fair treatment (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 49-68, 90-106). If not, a search for more data sources may be needed. Or the researcher may have to shift attention to idea-based research, and the collection of new data.

The type of data to be used is another consideration. Certain problems or issues may best be examined by seeking out quantified data (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 90-106). Other problems may require qualitative data, and methodologies (Glesne, 1999, 1-16).

The extent of the study is a major consideration, for several reasons. The scope of the data needed, and the extent of an idea-based study may make it unfeasible (Bieger and Gerlach, 1996, pp. 1-48; Crocker, 1999, *a*; Leedy and Ormdod, 2001, pp. 90-106). This being the case, researchers could consider turning to existing data sets, and data-driven research, in order to examine the problem or issue. When conducting data-driven research, the researcher should also consider whether the data to which they have access are extensive enough, or if they are too extensive (Bieger and Gerlach, 1996, pp. 1-48; Leedy and Ormdod, 2001, pp. 90-106). The researcher may have to seek out new data sources, or narrow the scope of the investigation.

Validity issues, and the context of the research, are important considerations. Data sets that are too extensive, or that include data from too many sources, may sacrifice internal validity, or credibility (Crocker, 1999, a; Leedy and Ormrod, 2001, pp. 90-106). The elements of *D.E.C.L.* (Mann, 2001)—the *D*elivery, the *E*nvironment, the *C*ontent, and the *L*earner—may not be homogenous across the data sources, which could be problematic if trying to make meaningful interpretations (Bieger and Gerlach, 1996, pp. 77-98). In addition, researchers must ensure that their findings have external validity, or generalizability (Bieger and Gerlach, 1996, pp. 77-85; Crocker, 1999, *a*; Leedy and Ormrod, 2001, pp. 103-106). Choosing data sources that are too small, or regionally focused, will result in findings that may not be generalizable to other regions, especially if the aspects of *D.E.C.L.* are not homogenous. Researchers must strike a balance between validity and content issues, depending on their purpose. Policy researchers may want to sacrifice content concerns when looking for benchmarks against which to gauge educational issues (Crocker, 1999, *c*). Or they may find that a lack of focus on content is inappropriate, if attempting to develop regionally specific policy (Crocker, 1999, *c*). The same can be said of data-driven research into relationships between *D.E.C.L.* aspects and localized academic performance. Again, the balance between internal and external validity, and the context of the data being examined, is affected by the specific needs of the researcher, and the users of the findings.

Finally, ethical considerations are as significant to data-driven research as they are to idea-based research (Bieger and Gerlach, 1996, pp. 227-232; Leedy and Ormrod, 2001, pp. 107-111). This is especially true in an educational context, where the relevant data involve children, and the results of the study are often used to determine policy affecting children (Bieger and Gerlach, 1996, pp. 227-232; Crocker, 1999, *c*; Leedy and Ormrod, 2001, pp. 107-111). Confidentiality of data sources, including the original research participants, is of utmost concern (Bieger and Gerlach, 1996, pp. 227-232; Leedy and Ormrod, 2001, pp. 107-111). Consent from the owners of the data, and possibly even the original research participants (and their parents or guardians) may be needed (Bieger and Gerlach, 1996, pp. 227-232; Leedy and Ormrod, 2001, pp. 107-111). In any research, there must be an effort to minimize the risk to research participants. While this may not be immediately applicable to data-driven, secondary research, it should still be taken into consideration lest the use of the existing data cause any undue harm to the original research participants (Bieger and Gerlach, 1996, pp. 227-232; Leedy and Ormrod, 2001, pp. 107-111).

Also under the subject of ethical considerations is researcher subjectivity. While the researcher's subjectivity plays a significant role in qualitative research, the existence of that

subjectivity is accounted for, controlled, and possibly even incorporated into the data analysis (Glesne, 1999, pp. 1-16, 95-112). Subjectivity becomes a serious ethical issue in policy research, even when that research is data-driven. The sponsors of the research may have political agendas when selecting researchers, and defining the scope of the research (Crocker, 1999, *c*). The researcher, whoever, should ensure that political agendas do not influence their own subjectivity, or their interpretations of data (Crocker, 1999, *c*).

Conclusion

Data–driven research has many benefits within educational contexts. It is often heavily relied upon to develop educational policy, to gauge the effectiveness of programs and practices, and to establish benchmarks between different regions and education systems (Crocker, 1999, *b, c*). It can also be used by individual educators and researchers to analyze more localized issues and trends from existing bodies of data (Crocker, 1999, *b, c*; Mann, 2002). And it is useful when the scope of data needed, or the types of data that already exist, render idea-based research either impractical or unnecessary (Bieger and Gerlach, 1996, pp. 1-48; Crocker, 1999, *b, c*; Leedy and Ormdod, 2001, pp. 90-106, Mann, 2002). The key to using data-driven research effectively is to recognize and consider the type of problem or issue that needs to be investigated, the type of data available, the appropriate research methodologies for interpreting the data, and the ethical considerations that surround all research.

References:

Bieger, George R., and Gail J. Gerlach (1996). *Educational Research: A Practical Approach*. Toronto: Delmar Publishers.

Crocker, Robert (1999). *Introduction to Quantitative Research*. [Online]. Available at http://www.webct.mun.ca:8900/45013200201/contents/lecture02.html

Crocker, Robert (1999). *The Nature of Educational Research*. [Online]. Available at http://www.webct.mun.ca:8900/45013200201/contents/lecture01.html

Crocker, Robert (1999). Policy Research. [Online]. Available at http://www.webct.mun.ca:8900/45013200201/contents/lecture08.html

Glesne, Corrine (1999). Becoming Qualitative Researchers: An Introduction, 2nd Edition. New York: Longman.

Leedy, Paul D., and Jeanne Ellis Ormrod (2001). *Practical Research: Planning and Design*. 7th Edition. Upper Saddle River, New Jersey: Prentice-Hall.

Mann, Bruce (2001). *Introductory Video to Education 6610: Research on Computers in the Curriculum*. St. John's, NF: Faculty of Education / School of Continuing Education, Memorial University of Newfoundland.

Mann, Bruce (Fall 2002). *Discussion Forum Correspondence in Education 6610: Research on Computers in the Curriculum*. St. John's, NF: Faculty of Education, Memorial University of Newfoundland.