Sources of Data for Research
A Research Primer
Walden University
Center for Research Quality

Laura Knight Lynn, PhD, Executive Director, Center for Research Quality
Leilani Endicott, PhD, Director, Center for Research Quality
Molly Lauck, PhD, Director, Center for Research Quality
Sunny Liu, PhD, Associate Director, Center for Research Quality
Louis Milanesi, PhD, Director, Center for Research Quality
Annie Pezalla, PhD, Associate Director, Center for Research Quality
Daniel Salter, PhD, Director, Center for Research Quality
Jenny Sherer, MEd, Associate Director, Center for Research Quality
Marydee Spillett, EdD, Associate Director, Center for Research Quality

For internal use only

© 2014 Walden University, LLC
# Content

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walden University Center for Research Quality</td>
<td>ii</td>
</tr>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Types of Information That Are Acceptable for Research</td>
<td>1</td>
</tr>
<tr>
<td>Systematic Methodology</td>
<td>1</td>
</tr>
<tr>
<td>Data Integrity</td>
<td>2</td>
</tr>
<tr>
<td>Data Alignment</td>
<td>2</td>
</tr>
<tr>
<td>Data Can Be Used in Three Different Phases of Research</td>
<td>3</td>
</tr>
<tr>
<td>Use 1 - Shaping the Research Question</td>
<td>3</td>
</tr>
<tr>
<td>Use 2 - Answering the Research Question</td>
<td>3</td>
</tr>
<tr>
<td>Use 3 - Elaborating on the Results</td>
<td>4</td>
</tr>
<tr>
<td>Taxonomy</td>
<td>5</td>
</tr>
<tr>
<td>Type of Data</td>
<td>5</td>
</tr>
<tr>
<td>What Is the Nature of the Data Being Used?</td>
<td>5</td>
</tr>
<tr>
<td>Original Purpose for the Data</td>
<td>5</td>
</tr>
<tr>
<td>The Data Were or Will Be Generated for What Reason?</td>
<td>5</td>
</tr>
<tr>
<td>State of the Data</td>
<td>6</td>
</tr>
<tr>
<td>Do the Data Currently Exist, or Will They Be Generated at Some Future Point?</td>
<td>6</td>
</tr>
<tr>
<td>Taxonomy of Data Sources</td>
<td>7</td>
</tr>
<tr>
<td>Research Ethics and Compliance</td>
<td>8</td>
</tr>
<tr>
<td>Examples</td>
<td>9</td>
</tr>
<tr>
<td>Existing Quantitative Research Data</td>
<td>9</td>
</tr>
<tr>
<td>Existing Qualitative Research Data</td>
<td>9</td>
</tr>
<tr>
<td>Prospective Quantitative Research Data</td>
<td>9</td>
</tr>
<tr>
<td>Prospective Qualitative Research Data</td>
<td>10</td>
</tr>
<tr>
<td>Existing Quantitative Operational Data</td>
<td>10</td>
</tr>
<tr>
<td>Existing Qualitative Operational Data</td>
<td>11</td>
</tr>
<tr>
<td>Prospective Quantitative Operational Data</td>
<td>11</td>
</tr>
<tr>
<td>Prospective Qualitative Operational Data</td>
<td>11</td>
</tr>
<tr>
<td>Resources</td>
<td>12</td>
</tr>
<tr>
<td>Methodology Training and Advice</td>
<td>12</td>
</tr>
<tr>
<td>Existing Data Collected for Research</td>
<td>12</td>
</tr>
<tr>
<td>Research Compliance</td>
<td>12</td>
</tr>
</tbody>
</table>
Introduction

The primary goal of this primer is to help researchers understand both the variety of potential sources of data for research, as well as the ways in which these data can be used in different phases of the research process. Because the research process can be varied and situation specific, this primer does not address the particular strategies to obtain access and to analyze the data from these various sources. Those skills are typically part of a scholar’s research training experience (see the Resources section at the end of this document).

Types of Information That Are Acceptable for Research

People today have access to so much information that it can be a bit overwhelming, especially when the goal is to use some of it for research. Not all types of information can be used for research, and a researcher needs to be mindful of the differences. Often, one of the biggest limitations to the utility of information in the research process is how it was or will be generated. The good news is that a researcher is typically able to exert some control over that process, either through careful analysis of existing information or rigorous collection and analysis of his or her own information.

This primer provides an examination of the types of information that can be used in research, which will be broadly referred to as “data” for ease of discussion. As will be seen, research data can take a lot of different forms—more than simply numbers in a spreadsheet. Some data already exist before a project begins, and in other cases, data will need to be generated to complete the study. Regardless of the source, data used for research are different from other types of information, and a researcher should ask the following three questions about them.

Systematic Methodology

A first question to ask is whether a systematic methodology was or will be used to generate the information. Not only do researchers need to attend to all the demands of data collection when generating their own data (e.g., appropriate sampling, standardized tools), they should have these same critical expectations for information and research produced by others. If there are no details about how the information was or will be generated, it likely cannot be used as data for research.

Use of a systematic methodology ultimately helps to assure that the results are interpretable and/or generalizable and involves most of the skills that are taught to new researchers in their coursework. For example, collecting information from a convenience sample raises questions about representativeness in quantitative studies, or inconsistent use of items on an interview protocol calls the trustworthiness of qualitative findings into question. A clear set of methods should be or should have been used to generate data for research, and these methods need to be consistent with acceptable practice and standards in research.
Data Integrity

A second question to ask concerns the integrity of the information and whether it was or will be evaluated before it is used or analyzed. Even if the information was generated with an acceptable and systematic methodology, it may still have problems that prevent its use as data in research. Running various diagnostic tests (e.g., screening for outliers or anomalies, member checking, computing reliability estimates) can often pinpoint problems that could compromise the results. If a researcher is unable to evaluate the quality of the information, then it likely cannot be used as data for research.

Data Alignment

A third question to ask is whether the information aligns to the research question(s) and the anticipated analytical strategy. Many researchers conduct a research project because the data that they need to answer a research question do not exist. In that way, they can control the alignment of their methods to the question and assure themselves (and their supervisory committee) that they are collecting appropriate data. In other instances, much information might exist around a particular group or phenomenon, but it may not align exactly to the research question. More likely, however, the available information may not exist in a form that makes it amenable to being analyzed. These topics are addressed further herein.

Alignment can be one of more difficult aspects of research to achieve and demonstrate. Basically, all the “parts and pieces” (i.e., literature, framework, approach, question, methodology, analysis, interpretation, etc.) of a study need to fit together as a whole, so that the findings address the original research problem. Some examples of misalignment in research include reviewing the research literature on children when the question concerns adults, administering an achievement test as a measurement of IQ in a study of intelligence, collecting information from managers when the question concerns employee perceptions, applying qualitative methods to test for group differences, or using an inappropriate statistical analysis strategy based on the variables in the study. For data used in research, the information contained in those data needs to align to the answer being sought in the research question.

An Example

José wants to understand the impact of employee satisfaction on productivity in the large company where he works. An informal conversation with a couple of administrative colleagues in his department would not produce information that is acceptable for a research study of employee satisfaction. José would not know whether those opinions are anything more than opinions or if they are shared by others in the company. He might also question whether administrators are actually the best source of information on how employees feel about the company.

Instead, José would want to start with a thorough examination of data reported in previous research on satisfaction and productivity. Then, one strategy to answer his research question might be to administer an anonymous, companywide survey of employees’ self-reported productivity, in conjunction with a standardized measure of satisfaction. Finally, to triangulate with those findings, José might also seek information on the business’s
performance metrics, which are typically collected in a systematic way for shareholders and governmental reporting.

Data Can Be Used in Three Different Phases of Research

Assuming a source of information is acceptable as data for research, a researcher has three distinct opportunities to use the data during the research process, depending on whether the data are more suited to shaping the research question, answering the research question, or interpreting the research question’s answer.

Another Example

Sue is conducting a study of the impact of mobile technology on the infant survival rate in countries that have a mix of regions with and without mobile phone coverage. In her literature review, she reports some statistics on mobile technology deployment in developing countries, as well as global trends in neonatal survival rates. In the methods section, she describes a design that involves collecting existing hospital records from several targeted countries, which she aligns to information that she obtains on whether the area served by the hospital has cell phone coverage. In the results section, she provides descriptive statistics for her sample districts (e.g., median incomes, number of residents), as well as the results of her regression analysis.

As can be seen in the example, Sue is using data from varying sources and in varying ways during the research process. Appreciating these differences can help a researcher to understand where to seek data and which types are appropriate at various points in the process.

Use 1 - Shaping the Research Question

Data can provide background information and support for pursuing the research project and shaping the research question and would typically be reported in the introductory sections of a published article, a doctoral capstone document (especially the literature review), or a research grant proposal. Some examples include

- Results and conclusions from other researchers, typically obtained from refereed journal articles. A good researcher always critically examines any previous findings related to the research topic.
- Statistics, which meet the abovementioned expectations for acceptable information, on the occurrence of the problem, such as data that are routinely kept by organizations and governmental agencies that collect data on behalf of the public.

Use 2 - Answering the Research Question

Certainly, data can be used to answer a research question, as discussed in the methods and results sections. These types of data are addressed in greater detail the Taxonomy section of this primer. At this point, it is important to know that researchers can rely on data that they collect themselves or data that were collected by a third party (or both!).
Use 3 - Elaborating on the Results

Data can be used to aid interpretation of results and to provide added insight into the research findings, while not being part of the main analysis. Some very typical examples include

- Descriptive statistics on the sample’s demographic variables to assure the external validity of the results, or diagnostic information to demonstrate the integrity of the data.

- Environmental referents (open-ended questions) that allowed respondents to elaborate on or clarify the responses that they gave on an instrument or survey or results from unanticipated post hoc analyses (sometimes research does not go the expected direction, regardless of how well planned it was).

Revisiting the Two Examples

In José’s project, it would be appropriate for him to report research results from published studies and public sites (e.g., the Pew Institute) in his literature review. He should have no contact with employees concerning his research project until he has Institutional Review Board (IRB) approval. Prior to that time, he can and should contact the human resources department or other key leaders in the company to discuss the feasibility of conducting the research and any requirements they have, but collection of actual data needs to occur after IRB approval.

In Sue’s project, she also has access to published research on her topic, as well as publicly available data, to argue for conducting this research project. Unlike José, she will be relying on existing data to answer her research question. Similar to José, she may want to do some background feasibility work to see if these data exist in a form that can be used to answer her research question, even though she will have to wait until she has IRB approval before she can analyze it for her study. She may also need to show IRB approval for the community partner to release the needed data.
Taxonomy

At this juncture, data for research have been distinguished from other types of information, by focusing on how they were or will be generated, evaluated, and used in the research. And, even though those criteria seem to limit a researcher’s options, a rich pool of data for research is still available. To organize the diversity of these sources, this taxonomy was developed and was built on three dimensions.

Type of Data

What Is the Nature of the Data Being Used?

One important distinction to make is whether the data are quantitative or qualitative. Although this primer assumes that readers are familiar with this basic distinction, they need to remember that validity and integrity of data can take different forms based on the nature of the data. For example, a researcher would expect to see a reliability estimate for scores from a measured variable or to see details on the coding strategies used with interview transcripts.

Original Purpose for the Data

The Data Were or Will Be Generated for What Reason?

A researcher should pay particular attention to the original intent or reason behind generating the data, especially for existing data (discussed in the next dimension of the taxonomy). When selecting data, two broad categories of reasons can have direct bearing on how the data are accessed and evaluated:

For Research Purposes

Researchers can take the necessary steps to generate the needed data themselves, of course. But, many individuals and organizations also collect and curate information with the expressed goal of supporting the scholarly production of knowledge, which can be used in the research process. The upside to this type of data is that the demands of research are usually a part of how they were or will be generated or analyzed, thus making them more amenable to analysis.

One caution: A researcher who is unfamiliar with these sources might be tempted to assume that the perfect “has-already-been-collected-for-me” dataset is sitting on a governmental agency server somewhere, just waiting to be downloaded and imported into SPSS for analysis. The reality is quite different. Sometimes, researchers need to be trained on the data management software used by the hosting agency and may spend weeks sorting through hundreds of variables to find the handful needed for their study.
For Nonresearch or Operational Purposes

Other people and organizations generate a lot of information for purposes other than research, but which has potential utility to the research process, assuming the information meets the criteria for data used in research. In some situations, the data needed to answer a research question cannot be obtained in any other way than by a third party (some examples are noted in the next section of this primer).

As with datasets generated for research purposes, the use of operational data can also present some challenges. Even though the researcher may be familiar with how the information was generated and tracked because he or she is employed by the organization, managing the dual role can present a challenge. Also, some organizations do not use the same rigor for data collection or management that is expected for research, and what seems like an acceptable dataset may be quite septic on closer analysis.

State of the Data

Do the Data Currently Exist, or Will They Be Generated at Some Future Point?

The temporal state of the data is important to how the integrity of these data is considered at different points in the research process, especially when developing a research proposal. Do the data actually exist or not? Because the research process occurs over months, it can get a little confusing to understand which types of data fit into which category. As discussed in the upcoming section on ethics and compliance, one key criterion is whether or not the research project has received institutional approval.

Existing Data That Were Generated Before Approval

Research studies using existing data are often called “secondary analyses” because the data were originally created for some other purpose (research or operations), and the current research project is a secondary purpose. Nothing about the research project is secondary in nature, as it is the researcher’s primary purpose for conducting the study.

Prospective Data That Will Be Generated After Approval

We may think of the classic research study as a process of collecting original data after the overall study has been evaluated and approved. The process for collecting, evaluating, and analyzing this type of prospective data (i.e., will exist in the future) for research purposes is what many research proposals entail, whether a doctoral capstone or a grant application. Not all research questions can be answered in this manner, and many researchers do not realize that they do not necessarily need to be the source of their prospective data.

That is, it is quite possible for a researcher to obtain approval to perform a secondary analysis of data that have not yet been generated by a third party (i.e., when the data will be primarily created for another purpose and its future use in research is secondary). This strategy does have some challenges, as sometimes the third party’s schedule for collecting the data does not always align with timing that the researcher desires. The upside is that, often, the
researcher is in a position to make sure that the prospective data are collected in the most rigorous means possible in the situation.

Using data from one’s immediate workplace (both prospective and existing) can present a number of challenges, and such a strategy is closely scrutinized for research approved by the university’s IRB. A few of the recognized challenges include the danger of losing researcher objectivity, the potential coercive relationship of the researcher to subordinates, and the balance of reporting any politically sensitive negative findings while protecting participants and the organization. If these challenges are effectively addressed during proposal development, however, these types of studies can often produce knowledge and insights that cannot be obtained through traditional means.

Taxonomy of Data Sources

<table>
<thead>
<tr>
<th>Research Data</th>
<th>Operational Data</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Existing</strong></td>
<td><strong>Prospective</strong></td>
</tr>
<tr>
<td>Data that <em>have been</em> generated for research purposes and can be accessed by the researcher for analysis</td>
<td>Data that <em>will be</em> generated for research purposes, with appropriate university approval</td>
</tr>
<tr>
<td><strong>Quantitative Examples</strong></td>
<td><strong>Quantitative Examples</strong></td>
</tr>
<tr>
<td>● Public/controlled databases</td>
<td>● Psychological tests</td>
</tr>
<tr>
<td>● Published research results</td>
<td>● Surveys/perceptual measures</td>
</tr>
<tr>
<td>● Archived research data</td>
<td>● Personal attributes (BMI, age, income)</td>
</tr>
<tr>
<td><strong>Qualitative Examples</strong></td>
<td><strong>Qualitative Examples</strong></td>
</tr>
<tr>
<td>● Public/private repositories</td>
<td>● Interviews</td>
</tr>
<tr>
<td>● Published research studies</td>
<td>● Observations</td>
</tr>
<tr>
<td>● Raw qualitative data from other researchers</td>
<td>● Written responses</td>
</tr>
</tbody>
</table>

**Existing**

Data that *have been* generated for an organization’s *operational* purposes and can be accessed by the researcher for analysis

**Quantitative Examples**

● Operational records
● Client data (deidentified)
● Public records

**Qualitative Examples**

● Past student work products
● Internal documents/policies
● Recordings and minutes
● Photos and video

**Prospective**

Data that *will be* generated for an organization’s *operational* purposes, which can be accessed by the researcher with appropriate university approval

**Quantitative Examples**

● Outcomes after an operational change.
● User statistics and data
● Future tracked information
● Byproducts of future intervention

**Qualitative Examples**

● Future student work products
● Artwork
● Media coverage
● Future online dialogue
Research Ethics and Compliance

One overriding consideration of data used in research concerns the ethics of how data are obtained. Because of its importance, a specific discussion of research ethics and compliance is offered here, with a focus on the role of the IRB. To review, data can be used to shape the question, to answer the question, and/or to elaborate on the results. Whether IRB approval is needed depends on what type/source of data is being used for which purpose.

The use of prospective data in any of the three ways needs prior approval from the university’s IRB. Researchers must never generate or collect new data without prior IRB approval, regardless of to which research phase the results are to be applied.

When needed, however, researchers are allowed to do some exploratory work on the feasibility of using or accessing the intended data without specific IRB approval. In some situations, a pilot study may be necessary before the main research can occur. Researchers should email IRB@waldenu.edu to find out whether prior IRB approval is needed for the pilot they have in mind or to access and/or use a dataset.

Accessing existing, publicly available data does not require prior IRB approval for Use 1 (shaping the research question) or Use 3 (elaborating on results). This allowance includes all information in research publications and public, nonsecured Web sites. It does not include private records held by organizations, even if a researcher has access to such information via his or her job. In those situations, IRB approval should be sought.

All individuals conducting research under the auspices of the university, including doctoral students and researchers seeking grant support, are required to complete an IRB review step for Use 2 (answering the research question). This step\(^1\) is applicable, even if only to confirm that any existing public data a researcher wishes to analyze are indeed acceptable for inclusion in a study. The IRB will determine whether any permissions are needed for the data that will address the research question.

Researchers should note that the IRB review is streamlined when accessing existing data to answer the research question, especially when accessing existing public data (i.e., only serves to confirm that the data are indeed public and therefore eligible for research use without any permissions).

---

\(^1\) All doctoral projects go through a formal IRB review after the supervisory committee and the University Research Reviewer have formally approved the proposal.
Examples

The following descriptions focus specifically on each of the eight types of data sources in the taxonomy. The listings of possible sources are by no means exhaustive and are meant to get researchers to start thinking beyond any preconceived notions that they may have about data. Examples are also provided to demonstrate how that source of data might be applied specifically to answer a research question (Use 2).

Existing Quantitative Research Data

This source includes quantitative data that have been generated for research purposes and that can be accessed by researchers for further analysis.

- Public databases (e.g., researchers or agencies that post entire datasets)
- Controlled access databases
- Published research results (for literature reviews, meta-analysis)
- Archived research data from other researchers (e.g., the Inter-University Consortium for Political and Social Research)
- Data from a previous study conducted by the researcher

Example: The researcher conducts a regression analysis of protective factors’ prediction of youth outcomes in a large multicity dataset that was originally only used to report relationships between risk factors and youth outcomes.

Existing Qualitative Research Data

This source includes qualitative data that have been generated for research purposes and that can be accessed by other researchers for analysis.

- Public/private research repositories
- Published research studies (for literature reviews, meta-synthesis)
- Research data made available by other researchers
- Data from a previous study conducted by the researcher

Example: A faculty member has conducted numerous studies of online discussion dynamics using a very large dataset of several years’ worth of deidentified discussion board postings and releases the dataset to the student researcher, who wishes to analyze the postings along a new dimension not previously examined.

Prospective Quantitative Research Data

This source includes quantitative data that will be generated for research purposes. These types of studies follow what might be considered a more traditional model of quantitative research conducted by researchers, although some examples of secondary analysis do exist (see Prospective Quantitative Operational Data below). Remember: All uses of prospective data require prior approval of the university’s IRB.
• Psychological tests
• Surveys and perceptual measures
• Personal attributes (BMI, age)
• Quantitative data from an approved in-progress research project at another site.

Example: A grant-funded faculty researcher uses a causal-comparative design to compare the differences in trait anxiety and research self-efficacy in doctoral students who video-chatted regularly with their chair in the first term of capstone completion and doctoral students who used the phone.

Prospective Qualitative Research Data

This source includes qualitative data that will be generated for research purposes. These types of studies follow the more traditional model of qualitative research conducted primarily by doctoral students, although some examples of secondary analysis do exist (see the last example in this list). Remember: All uses of prospective data require prior approval of the university’s IRB.

• Interviews
• Observations
• Written responses
• Qualitative data from an approved in-progress research project at another site.

Example: In a metropolitan community where the funding for several social services agencies was recently realigned, the student researcher conducts interviews of the agency directors and their direct reports to understand how they responded to this change.

Existing Quantitative Operational Data

This source includes quantitative data that have been generated for operational (nonresearch) purposes, which can be accessed by the student researcher for analysis.

• Operational records from any organization tracking outcomes, expenditures, predisposing factors, personnel history, etc.
• Client data (deidentified)
• Public records (that were created as part of some nonresearch operation)
• A Web site’s past usage statistics and user data
• Information generated on any organization’s clients, patients, students, personnel, etc. (as long as it is deidentified before being shared with researcher)

Example: To understand the difference in schools that participated in a statewide initiative and those that did not, a researcher conducts a repeated-measures study using archived state-administered achievement tests to compare group differences in the change in student scores over 5 years.
Existing Qualitative Operational Data

This source includes qualitative data that have been generated for operational (nonresearch) purposes, which can be accessed by the student researcher for analysis.

- Internal documents and policy statements
- Client notes
- Recordings and minutes of meetings or other events
- Photos and video
- Any artifacts created as a byproduct of an organization’s operations

Example: To develop an understanding of how “trolls” impact perceptions of an emergency event, the researcher performs a content analysis of Twitter “Tweets” from three recent high-profile emergencies in an urban setting.

Prospective Quantitative Operational Data

This source includes quantitative data that will be generated for operational (nonresearch) purposes and can be accessed by the student researcher for analysis. Remember: All uses of prospective data require prior approval of the university’s IRB.

- Outcomes that can be tracked in the future, as a new practice is implemented (can be tracked using the organization’s standard recordkeeping or using a modified recordkeeping practice, if organization approves)
- A Web site’s usage statistics and user data that will be observed in the future
- Information tracked in the future on any organization’s clients, patients, students, personnel, etc. (as long as it is deidentified before being shared with the researcher).

Example: The researcher uses regression to examine the predictive power of users’ demographics on their frequency of utilizing particular resources on a Web site, which will be supplemented with an analysis of past data.

Prospective Qualitative Operational Data

This source includes qualitative data that will be generated for operational (nonresearch) purposes, and will be accessed by the researcher for analysis. Remember: All uses of prospective data require prior approval of the university’s IRB.

- Internal documents and policy statements
- Client notes
- Recordings and minutes of meetings or other events
- Any artifacts created as a byproduct of an organization’s operations (could include student work products, artwork, press coverage, etc.)

Example: A pending policy change provides an opportunity for the student researcher to analyze recordings and minutes of a board’s meetings, in addition to public reactions in the press, in order to analyze the relationship between the new policy and the board’s upcoming decisions and relations with the public.
Resources

Methodology Training and Advice

Much discussion in this primer concerns the use and application of systematic methodologies. These topics can range from selecting a sample, to designing a survey, to choosing a statistical analysis that is appropriate for which type of variables. Developing these skills and abilities is part of a researcher’s training.

Sometimes, when examining his or her options with a source of data, a researcher may need a refresher on a topic, more information on a new technique, or even some advice from a more experienced researcher. The Research Resources page on the Center for Research Quality’s (CRQ) Web site has a broad array of these types of resources—from just getting started to publishing one’s findings.

Faculty can also feel free to email methodologyadvice@waldenu.edu if they need support with a student’s project or with their own research (more information is on the Research Resources page). This service is for faculty only, so students’ best resources for methodology advice are the research faculty members at the university, especially the members of their capstone supervisory committee.

Existing Data Collected for Research

The use of existing data collected to support the research process has grown noticeably in the past few years at Walden. So much so, that the CRQ has recently launched a new page devoted to this topic: Data Resources and Support. Even if a researcher is not planning on a secondary analysis of this type of data, many of the linked sites provide descriptive data that might be helpful in Use 1, shaping the research question, described earlier.

Research Compliance

Many questions related to research ethics and compliance can be found on the CRQ’s Office of Research Ethics and Compliance page. Researchers with specific questions that are not answered in those materials are encouraged to email the board at IRB@waldenu.edu. Just keep in mind that the IRB review comes at a particular point in the research process. That is, if a researcher has not settled on a topic and a methodology, its likely premature to seek IRB approval for a particular sample.