Advancing education research through mixed methods with existing data

Daphne C. Watkins[a] **and Natasha C. Johnson**[b], [a] University of Michigan, School of Social Work, Ann Arbor, MI, United States; and [b] Columbia University School of Social Work, New York, NY, United States

Glossary

Conceptual framework A collection of concepts and empirical findings from the literature, used to show relationships among ideas and how they relate to research.

Descriptive research The second level of the knowledge-level continuum that uncovers descriptive information on the topic of interest using qualitative and/or quantitative methods.

Existing data Data already collected by you or someone else that can be used in a new research project. Also referred to as "secondary data."

Existing qualitative data Text, visual, or recorded (e.g., audio, video) data collected by you or someone else that can be used to address a new research question.

Existing quantitative data Numerical data collected by you or someone else that can be used to address a new research question.

Explanatory research The highest level of the knowledge-level continuum and is conducted using quantitative methods of inquiry because previous data exists on the topic of interest.

Exploratory research The first level of the knowledge-level continuum; includes research that begins to fill the substantial gap in the existing literature related to your topic of interest.

Knowledge-level continuum A framework for describing the progression in the levels of scientific inquiry. If there is little to no pre-existing knowledge on your topic of interest, an exploratory level of investigation should be used. If some pre-existing knowledge exists, a descriptive level should be used. If there is a robust amount of pre-existing knowledge, an explanatory level should be used.

Mixed methods The rigorous and epistemological application and integration of qualitative and quantitative research approaches to draw interpretations based on the combined strengths of both approaches to influence research, practice, and policy.

Mixed methods with existing data Identifying, evaluating, and incorporating one or more existing (i.e., secondary) qualitative or quantitative data sources in a project.

Primary data Data collected in real-time to achieve the goals outlined in a research project, also called "original" or "raw" data.

Secondary data Data already collected by you or someone else that can be used in a new research project. Also referred to as "existing data."

Theory An explanation of the "how" or "why" of phenomena, typically referencing existing facts. A theory is usually based on an existing framework and is frequently tested.

Introduction

Funders and publishers are instrumental in establishing the guidelines and processes we use for sharing data (Chawinga and Zinn, 2020). Education researchers—particularly those who use existing data (or "secondary" data) sources in mixed methods—can benefit from the recent increase in accessible data, as the advantages of using existing data in education research are far and wide (Logan, 2020; Smith, 2011). One of the most fundamental benefits of existing data is its ability to extend the knowledge about a research topic above and beyond the current state of the science. For example, a growing body of literature and data sources considers the impact of the global COVID-19 pandemic on students' learning and educational experiences (Dorn et al., 2020; Grooms and Childs, 2021). Existing data sources that include COVID-19-related items or information about student experiences during the global pandemic can be used to extend the field's understanding of educational disparities and shifting academic norms. Existing data provide an opportunity to think critically about a research question outside its initial parameters, using data collected by you or someone else. Furthermore, having a chance to work with existing data gives researchers an insider view of a complete dataset, its strengths and weaknesses, and its ability to address a research question associated with their research interests. An insider view of a project can also teach budding researchers how to lead their own research projects.

This chapter describes how researchers can use existing data in mixed methods projects to advance education research. We begin by defining existing data for mixed methods using an education research context. We then cover the advantages and disadvantages of using existing data broadly and in mixed methods projects. Next, we present the knowledge-level continuum as a way to frame the exploratory, descriptive, and explanatory research conducted in educational settings. Then, we use examples from education research to describe existing qualitative and quantitative data in mixed methods and how to "mix" these data to answer research questions. Easy-to-follow instructions do not always accompany existing data sets. Therefore, sifting through data records and making sense of the available information associated with existing data sets is an essential part of the process. We end with a discussion of how existing data can support the deepening of knowledge and the breadth of educational topics researchers and practitioners can explore.

What are existing data in mixed methods?

Mixed methods are the rigorous and epistemological application and integration of qualitative and quantitative research approaches to draw interpretations based on the combined strengths of both approaches to influence research, practice, and policy (Plano Clark and Ivankova, 2016; Creswell, 2015; Watkins and Gioia, 2015). John Creswell (2015) defines mixed methods as:

> An approach to research in the social, behavioral, and health sciences in which the investigator gathers both quantitative (closed-ended) and qualitative (open-ended) data, integrates the two, and then draws interpretations based on the combined strengths of both sets of data to understand research problems (p. 2).

Previous definitions of mixed methods clearly distinguish between quantitative and qualitative data. However, they rarely note whether these data can come from existing data sources or are newly collected data specific to each mixed methods project (Watkins, 2022). Given the importance of this distinction, we insist that one might use at least two types of data in their mixed methods: primary and secondary data. *Primary data* are collected in real-time to achieve the goals outlined in a research project. Other names for primary data are "original" or "raw" data. *Secondary data*, or *"existing" data*, are the data that have already been collected by you or someone else that can be used in a research project. Various data sources exist, ranging from large, publicly available data (e.g., US Census data) to smaller, privately-owned data (e.g., local social service agency data of the client population). Whether they include publicly available case studies, data points, papers, artifacts, or electronic documents, existing data are valuable in education research (Logan, 2020). Scholars have gone to great lengths to define mixed methods studies apart from qualitative and quantitative single-method studies (Bazeley, 2018; Creswell and Plano Clark, 2018; Hesse-Biber, 2010; Mertens, 2009, 2017; Morse, 2009). Mixed methods have also been described across various disciplines (Curry and Nunez-Smith, 2015; Johnson and Christense, 2013; O'Cathain, 2010) and professions (Andrew and Halcomb, 2009; Haight and Bidwell, 2016; Johnston, 2012; Magee et al., 2006; Sheperis et al., 2016; Watkins and Gioia, 2015). Yet, sparse resources provide clear guidelines for what mixed methods with existing data involve.

Existing data can be either qualitative or quantitative. Smith's (2008) book *Using Existing Data in Education and Social Research* provides a concise account for using existing qualitative data in social science research. The author affirms that existing data "… can be numeric or non-numeric. Non-numeric, or qualitative secondary data can include data retrieved second hand from interviews, ethnographic accounts, documents, photographs, or conversations" (pp. 4–5). Though most existing data sources tend to be quantitatively oriented, Smith reminds us of the possibilities of existing qualitative data in our education and social research endeavors. Let's consider an example pertinent to education research. Charlie is a graduate student interested in examining K-12 students' sense of belonging within the virtual learning environment since the COVID-19 pandemic. Currently, the knowledge base on virtual learning since the start of the pandemic is still relatively new, so there is little to no data on students' sense of belonging in the widespread but new virtual learning environment. Charlie's first step is to finalize the research question, then search for existing

data sources to address the research question. What if Charlie helped collect data on a similar topic for their professor when COVID first began? Does this still count as "existing data?"

Adding a time element to the definitions of primary and existing data is critical when distinguishing between the two because you can collect primary data now that you use for secondary purposes later (Watkins, 2022). There is no hard and fast rule to the length of time that must pass before primary data can be considered "existing" data. The only real distinction involves whether someone has a fundamental research question they want to answer and whether the data to address that question exists. For example, if Charlie collected data 2 years ago and those data can help answer a different research question now, Charlie can use those data to answer the new, different research question. Whether the new research question is related or unrelated to the original study, Charlie can use those existing data for a secondary purpose.

Education research can benefit from combining mixed methods with existing data methods (Fig. 1). Mixed methods involve collecting and analyzing both quantitative and qualitative data to address a research question. Whereas existing data methods involve evaluating quantitative and qualitative data that have been previously collected and deciding how to use them for a secondary purpose. Bringing the two methods together results in mixed methods with existing data. Hence, doing *mixed methods research with existing data* involves identifying, evaluating, and incorporating one or more existing data sources into a mixed methods project (Watkins, 2022). Incorporating existing data into mixed methods takes the definition of mixed methods to a new level, as mixed methods are ways to collect, analyze, and integrate qualitative and quantitative data in rigorous and theoretically sound ways to encompass the breadth and depth of a topic (Creswell, 2015; Guetterman et al., 2015; Tashakkori and Teddlie, 2010).

Using existing data in mixed methods also expands the purpose of secondary data analysis, which is to further analyze an existing data source by addressing a research question like (or distinct from) that for which the original data were collected (Hewson, 2006). Increasing our use of existing data sources adds to the unique possibilities for the future of mixed methods in education research. Previous studies suggest existing data are typically used at the beginning (or for the first phase) of a mixed methods study (Hewson, 2006; Johnston, 2014). However, as qualitative and quantitative data sources become more robust and sophisticated, the potential for implementing entire mixed methods studies using education data previously collected for a different purpose is promising.

Using data collected by you or someone else to answer your research questions is common in education research and academic settings (Davis-Kean and Jager, 2017; Hewson, 2006; Panchenko and Samovilova, 2020). It can maximize a project's resources (e.g., time, resources, and energy) beyond those used for its original purpose. The purpose statement, research questions, and theory should guide the decision to use existing data in a mixed methods project. Further, applying or generating theory in mixed methods with existing data will vary depending on whether the research prioritizes the quantitative or qualitative phase of the study. For instance, let's revisit the education research example presented above. Suppose Charlie wants to investigate students' sense of belonging within the virtual learning space, and there is little pre-existing knowledge about students' sense of belonging specific to virtual learning. In that case, Charlie must prioritize a research approach that allows them to gather more comprehensive insight into how students describe their connectedness and belonging to peers and educators within the virtual space. Charlie might use mixed methods with existing data for this study. Moreover, Charlie should also review the advantages and disadvantages of using mixed methods with existing data before proceeding.

Advantages and disadvantages of using existing data in mixed methods

There are several advantages to using existing data in mixed methods. For example, using existing data to address a mixed methods question may save time, money, and the resources necessary to collect primary data. Also, researchers might decide to use existing data because they have access to a high-quality and robust data source. Students are drawn to existing data for their mixed methods because the data can extend their current scholarship. Deep diving into someone else's data can expand a researcher's training and scope. Though one's education, training, and skills are unique to them and allow them to view the world through a particular lens, this uniqueness might also limit the researcher's contribution to research. In short, using existing data can enhance a researcher's scope and abilities.
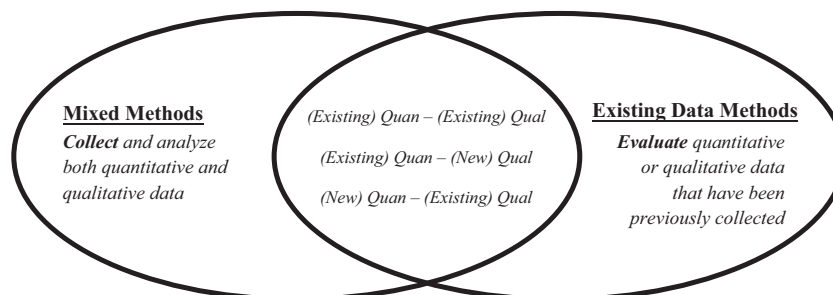

Fig. 1   The merging of methods: mixed methods with existing data.

Despite the advantages of using existing data in mixed methods, there are also some challenges. For example, some researchers might find using someone else's data mysterious and unclear. Budding researchers may also question their authenticity because they did not undergo all the steps in the research process (e.g., idea conceptualization, IRB application, recruitment, instrument development, data collection, data analysis, reporting, etc.) to achieve their project goals. Then there is the matter of missing information. Sometimes existing data files are incomplete, leaving researchers who want to use the data for secondary purposes confused about the procedures undertaken by the original research team and uncertain about the best way to repurpose the data to answer a new research question. Despite the challenges associated with existing data, they can be valuable for education research and can help streamline the path to answering the research questions. However, it is best to start at the beginning of scientific inquiry to address a research topic to do this well. This process involves deepening your understanding of your topic and the ways mixed methods with existing data can enhance exploration in an educational context. To do this, researchers must examine the topic, and the possibility of using existing data for a mixed methods study in the context of the knowledge-level continuum.

Existing quantitative data in mixed methods

Though the popularity of existing qualitative data has increased (Beck, 2019; Hughes and Tarrant, 2020; Largan and Morris, 2019), historically, education researchers who have used existing data in their studies have been guided by the quantitative data. In other words, the quantitative data are the dominant or prioritized phase in the mixed methods. Theory will guide the variables used if the researcher prioritizes the quantitative study component in mixed methods research. In sequential designs, prioritized phases occur before other phases of the study. Suppose the prioritized quantitative step involves using an existing dataset. In that case, the theory used by the original investigators will guide your understanding and selection of the pre-existing variables in those data to help answer your research questions.

Existing quantitative data can come in many forms, from thousands of individual high school student cases to a dozen or so school administrative cases in your office building. Existing quantitative data are everywhere and include population-based studies, which use methods to survey all people within a specific population (e.g., US census). Other examples of existing quantitative data are cohort-based and longitudinal surveys, such as the National Survey of American Life, conducted from 2001 to 2003 (Jackson et al., 2004). Though previously collected sources exist, researchers still need to request access to the data or investigate if the data are publicly available. A Google search of "data repository" would help you start your journey to identify publicly available existing quantitative data.

Gauge the quantitative "breadth" of a topic, then proceed qualitatively

Existing quantitative data in a mixed methods study can help you gauge the breadth of a topic before you proceed with the qualitative inquiry. Quantitative inquiry uses deductive reasoning to test and confirm (or refute) a theory (Grinnell and Unrau, 2018). This process usually begins with a theory and then establishes a plan to confirm or refute that theory. Given the sheer number of cases that can be observed using quantitative methods, there are apparent advantages to analyzing existing quantitative data. One clear advantage is gauging the quantitative breadth of a topic before deciding how to proceed qualitatively. For example, let's assume you are a data analyst at Booker T. Washington High School. Your assistant principal has asked you to explore possibilities with the school's confidential student survey data. You decide to use these existing quantitative data to understand what adolescents think about the social-emotional learning services offered by Booker T. Washington High School. After you get descriptive data about students' feelings about social-emotional services, this information could be strengthened by a small, qualitative study in which you assemble groups of students who represent the demographics reflected in your existing quantitative analysis and then ask them to expound on your quantitative findings with new qualitative responses.

Existing qualitative data in mixed methods

In qualitative research, inductive reasoning identifies patterns that generate theory. Compared to mixed methods with a dominant quantitative component, theory plays a different role with a dominant qualitative component. Rather than guide the selection of variables to test the theory, when using existing data in mixed methods, the data generate theory, or in most cases, a conceptual framework that can be tested with quantitative data. In this context, a *theory* explains the "how" or "why" of phenomena, typically referencing existing facts (Kimmons, 2022). A theory is usually based on an existing framework and is frequently tested. On the other hand, a *conceptual framework* is a collection of concepts and empirical findings from the literature used to show relationships among ideas and how they relate to research. For the purposes of this chapter, conceptual frameworks are "preliminary theories" you hope to evolve into testable explanations for the how and why of a particular phenomenon. When beginning mixed methods with a dominant qualitative component, you might consider building a preliminary conceptual framework (to test as a theory in a subsequent study). Whether you plan to generate theory (e.g., qualitative analysis) or test theory (e.g., quantitative analysis), the theory has a role in mixed methods that should be embedded throughout your research.

Existing qualitative data (e.g., text, image, etc.) provide insight and help address some of your unanswered qualitative research questions. Though not as popular as their quantitative counterpart, existing qualitative data has gained some traction from previous scholars (Bishop, 2007, 2009; Fielding and Fielding, 2000; Hammersley, 2009; Hinds et al., 1997). In particular, and given the increased respect, rigor, and credibility of qualitative data in recent years, more scholars consider its use in an existing analysis

capacity. In fact, in 2019 and 2020 alone, three books were published on secondary analysis of qualitative data (Beck, 2019; Hughes and Tarrant, 2020; Largan and Morris, 2019). Despite this progress, a comprehensive appreciation for existing qualitative data lags that of existing quantitative data. This seems to be due to the plethora of resources and guidelines for using existing quantitative data compared to the sparse resources available for using existing qualitative data. Some examples of existing qualitative data may include video recordings, audio recordings, and text documents (i.e., transcripts) of previous interviews with students.

Gauge the qualitative "depth" of a topic, then proceed quantitatively

Qualitative research uses inductive reasoning, beginning with observations and ending with theory. Therefore, if information from an existing qualitative data source aligns with the information needed to address your research question, you should consider using these existing qualitative data as a source for your mixed methods. There is an implied depth associated with qualitative research; therefore, you can immerse yourself in the data, conduct a rigorous qualitative analysis, then generate a conceptual framework (i.e., preliminary theory) that connects relevant concepts about your topic of interest. Deep inquiry into existing qualitative data can serve as a springboard for building a new quantitative study component into your mixed methods project.

Let's consider another example where your mentor owns a qualitative data source on educational outcomes of court-appointed youth, and you are interested in exploring these data further. You and your mentor study this topic, but he collected focus group data three years ago and recently received a Department of Education grant to which he needs to turn his attention for the next 5 years. He offers the data to you to help advance your scholarship in this area. Using your mentor's existing qualitative data can help you gauge the depth of the educational outcomes of court-appointed youth so that you can know how you want to proceed quantitatively. Namely, you could analyze the focus group data to generate a list of concepts and themes for which a conceptual framework can be created. After which, you could collect new quantitative data (including items that measure the concepts and themes generated from your qualitative analysis) to generate hypotheses and test your conceptual framework. These ideas are ways to enhance the use of your mentor's existing qualitative data by repurposing them so that you have some direction in your research.

Both quantitative and qualitative data have advantages and disadvantages, but sometimes researchers are reluctant to use qualitative data as an existing data source. This is unfortunate given the plethora of possibilities for open-ended inquiries into the human experience. For example, there are benefits to assessing social and cultural contexts, connections, relationships, and subjective processes that result from social phenomena (Watkins and Gioia, 2015). Furthermore, uncovering the factors that influence decision-making and the opinions of research participants will strengthen your research, not weaken it.

The knowledge-level continuum in mixed methods

To understand how mixed methods with existing data can be helpful in education research, start by making decisions about your study based on the current research. This framing of education research can help because it addresses two fundamental questions when using existing data: (1) What is the existing *knowledge* on the topic? (2) What are the existing *data* on the topic? But before we ask these questions, we should first ask, what is the knowledge-level continuum and how is it helpful in guiding mixed methods with existing data in education research?

The *knowledge-level continuum* is a framework (Fig. 2) for describing the current state of science in an area of inquiry and the progression of science therein (Watkins, 2022). It outlines existing knowledge, and the ways existing knowledge merges with the organization, creation, and dissemination of new knowledge (Watkins, 2022). Previously, the knowledge-level continuum has been used in other disciplines, such as information technology (Vassallo, 1999) and social work (Grinnell and Unrau, 2018; Watkins and Gioia, 2015). Research inquiry can fall anywhere along the continuum depending on how much is already known about the topic. The knowledge-level continuum encourages researchers to think about research as being a direct by-product of the information about a topic and the purpose of the study (Grinnell and Unrau, 2018; Watkins, 2022). In short, if there is little to no preexisting knowledge of your topic of study, an exploratory level of inquiry should be used. If some pre-existing knowledge exists, a descriptive level of investigation should be used. If there is a robust amount of pre-existing knowledge, an explanatory level of inquiry should be used (Watkins, 2022).

Let's revisit the previous example with Charlie to understand how the knowledge-level continuum can be applied to education research. Charlie is interested in examining students' sense of belonging in virtual learning spaces. Specifically, Charlie is interested in reviewing the factors that support and limit students' sense of belonging in the new virtual learning environment. Let's say Charlie completes a thorough literature search, looking through dozens of peer-reviewed article databases. Unfortunately, very little published research examines students' sense of belonging within virtual learning spaces for K-12 students.

Additionally, of the very little information published on the topic, Charlie found it to be of poor quality, leading to challenges regarding Charlie's trust in the rigor and quality of the research. Given the limited amount of pre-existing research and the poor quality of the research, if Charlie were initiating a research project in this area, they would use an exploratory level of inquiry. *Exploratory research* is the first level of the knowledge-level continuum. It includes research that begins to fill the substantial gap in the existing literature related to the topic of interest. Exploratory research can be conducted using qualitative methods of inquiry. For the example described, Charlie may begin with interviewing all students at a specific school who participated in virtual learning and then asking about their sense of belonging.
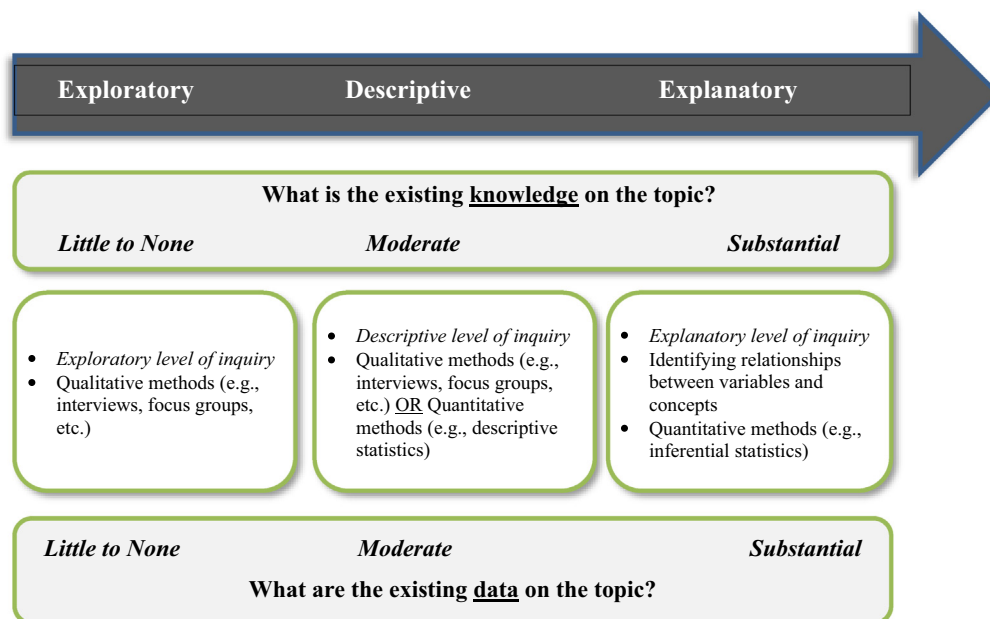
Fig. 2 The knowledge-level continuum as a framework for mixed methods with existing data.

Considering the same example above, let's say Charlie continued their search for peer-reviewed articles that demonstrate students' sense of belonging in virtual learning spaces and cannot locate articles. However, Charlie's faculty advisor sends them eight peer-reviewed articles she found while working on a grant. As Charlie reviews the articles, they are excited to see the articles synthesize a lot of the information they hoped to find. The articles also describe preliminary findings on college students participating in online classes and their sense of belonging. However, none of the articles include results for K-12 students and their sense of belonging in virtual learning spaces during the unprecedented time of the COVID-19 pandemic. In short, there are still some gaps in research that focus on this topic.

In reviewing the knowledge-level continuum, Charlie acknowledges that while *some* information exists on the topic of interest, the existing literature does not extend to cover the specific topic Charlie was hoping to investigate. Given this, Charlie decides that a descriptive level of inquiry would be best. *Descriptive research* is the second level of research within the knowledge-level continuum. It involves expounding on the pre-existing knowledge by describing the sample and concepts of interest in more detail. Descriptive research can be conducted using either qualitative or quantitative methods of inquiry. In the case above, Charlie may use an existing sense of belonging psychometric tool to measure students' sense of belonging in the virtual learning environment.

If the existing literature included substantial knowledge of sense of belonging, specifically for K-12 students that participated in virtual learning during the unprecedented 2020–2021 academic year, then Charlie would use an explanatory level of inquiry for the research. *Explanatory research* is the highest level of the knowledge-level continuum and is conducted using quantitative methods of inquiry because previous data exists on a topic. If Charlie conducted an exploratory study (i.e., first level of the knowledge level continuum) to understand the students' sense of belonging in virtual learning environments and then conducted a descriptive study (i.e., secondary level of the knowledge continuum) to collect more detailed information about students and educators' sociodemographic background and characteristics, then existing data exists on the topic! Therefore, an explanatory study can help Charlie understand the associations between concepts identified in the exploratory and descriptive studies on this topic. Explanatory research studies produce inferential statistics that Charlie can generalize to the larger population.

Applying the knowledge-level continuum to mixed methods with existing data

The knowledge-level continuum can be used as a framework for mixed methods research with existing data because it both informs, and is informed by, how much is already known about the topic of interest. Let's consider an example where a researcher, Marcus, is trying to decide whether to use a convergent, exploratory sequential, or explanatory sequential mixed methods design for a study with existing data (Creswell, 2015; Watkins, 2022). In this example, Marcus wants to use existing quantitative data on SAT scores for Deferred Action for Childhood Arrivals (DACA) students who attend ivy league colleges and universities in the United States.

The first step is for Marcus to search for previous knowledge and data on the topic. He might decide to check Google Scholar, Psych Info, EBSCOhost, and Science Direct for articles that include keywords in the title and abstract. Then, he might finalize his research question (*How do SAT scores for DACA students who attend ivy league colleges and universities compare to SAT scores for other*

DACA students at other colleges and universities at their colleges and universities?). After the research question is finalized, Marcus should consider whether the question is suitable for integrating his existing data with a second, qualitative data source. Marcus must determine which mixed methods design (Fig. 3) will help address his research question. He can do this by examining the research question, the current state of science (i.e., previous knowledge and data on the topic), and the existing data on the subject (i.e., existing data on DACA students' SAT scores).

If Marcus finds little to no existing research on the topic, he has the option to conduct an exploratory sequential design. However, given the variables in his existing dataset, he could likely build a convergent design where the quantitative phase involves analyzing the existing data on SAT scores for DACA students at ivy league colleges and universities. His qualitative phase could include online interviews with a select number of DACA students at ivy league colleges and universities across the United States. Though his existing quantitative data are more descriptive, and he could also conduct an explanatory sequential design, he would prefer to do a convergent mixed methods study where the new qualitative findings can be merged with the existing quantitative data results.

Three questions when choosing existing data in mixed methods research

Before incorporating existing data in your mixed methods research, you must first identify the most appropriate existing data for your project. You may be asking yourself: How do I choose the best existing data for my mixed methods project? Below, we provide three questions you should ask yourself as you review potential data sources.

Question 1: How will you define the terms in your research question?

The first question you should ask yourself will help outline the plan of action for your mixed methods project. When looking at your research question, what does each word mean? In other words, how will you define the terms in your research question for your mixed methods with existing data? Essentially, what you need to uncover is what is under investigation in your study. As you choose specific terminology for your research question, be sure to determine how you will define and measure these terms. The words you use to develop your research question will outline an action plan for your mixed methods with existing data. In the earlier example, Charlie can perform a thorough literature review of students' sense of belonging, then choose terms for their research question that signal the study's exploratory nature. Therefore, Charlie's research question could be *What are the barriers and facilitators to students' sense of belonging in virtual learning spaces?* Looking closely at this research question, we would advise Charlie to highlight words that need a definition and are directly connected to the success of their study. For example, terms such as "barriers," "facilitators," "students," "sense of belonging," and "virtual learning spaces" all need clear definitions. They also need units of measurement as the success of Charlie's mixed methods is contingent upon a deeper understanding of and the ability to measure these terms.

Question 2: How do you plan to find the existing data?

There may be field-specific data repositories for existing data; however, you must actively seek out information to help locate existing data sources regardless of your field of study. The next question you ask yourself should point you toward existing data that will help answer your research question. Begin by asking your mentors, advisors, professors, and classmates for recommendations for data sources related to your well-defined research question. Additionally, use institutional resources, such as academic librarians, to strategize the methods used to search the web and online databases. It is important to note that the more defined your research question is, including your concepts and variables, the easier it is for others to assist with your search. You want to develop a plan or strategy to locate the data at this stage. Which search terms will you use? What strategies (e.g., filters, Boolean connectors) will you use to search across data repositories? What data depositories can you access? In Charlie's case, they could retrieve existing
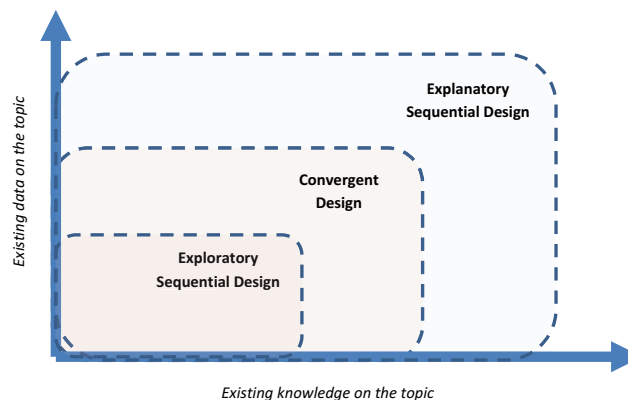

Fig. 3 Choosing mixed methods designs that inform, and are informed by, existing data and the knowledge-level continuum.

data from their professor, so thankfully, Charlie did not need to look very far to locate existing data. However, it will be necessary for Charlie to closely examine the appropriateness of their professor's existing data to ensure that the research question (*What are the barriers and facilitators to students' sense of belonging in virtual learning spaces?*) can be addressed. This leads us to the final question.

Question 3: Do the existing data help answer the research question?

Now that you have defined the terms in your research question and developed a plan for locating existing data, you must address the final question regarding the appropriateness of the existing data for your mixed methods project. Evaluate the existing data to determine if it has adequate information to answer your research question (Kiecolt and Nathan, 1985; Smith, 2008, 2011). This final question aims to ensure you have thoroughly assessed the fit of the existing data for answering the research question *before* incorporating it into your mixed methods. Go through the existing data, supplemental materials, and any previously published reports on the data so you can learn about the primary study, from conceptualization to evaluation. So Charlie will need to conduct a rigorous review of the variables (if quantitative), concepts reviewed (if qualitative), and all the professor's supplemental materials on file. If there are materials the professor does not have, Charlie should try to locate these materials to help them develop a comprehensive knowledge base about the data. Charlie should be able to answer questions such as: What was the original purpose for the data? How were participants recruited? How were relevant concepts measured? How was missing data handled?

In sum, there are three questions to address when choosing to include existing quantitative or qualitative data in your mixed methods study. This strategic process begins with you asking yourself questions that will help you explicitly define your research question. Next, you must develop a plan to locate existing data that includes the variables or concepts captured by your research question. Finally, once you locate the existing data, you must assess if they meet the needs of your research question. These steps provide a roadmap and an action plan for selecting existing data.

Conclusion

This chapter provides an overview of how mixed methods with existing data can advance education research. The knowledge-level continuum can serve as a framework for your mixed methods study. It describes the research process for all studies, depending on how much is already known about the topic of interest. The knowledge-level continuum is a direct by-product of the existing information on a topic, the purpose of the research, and your anticipated next steps. Using existing data in mixed methods includes identifying, evaluating, and incorporating one or more existing data sources to serve as one or more data phases of your mixed methods. It acknowledges not only the purpose of mixed methods but also the purpose of existing data analysis, which is to further the investigation of an existing data source by addressing a research question similar to (or distinct from) that for which the original data were collected.

Using existing data is advantageous in mixed methods because it allows you to expand your knowledge and understanding of a topic, gauge the qualitative depth of a topic so you know how to proceed quantitatively, and gauge the quantitative breadth of a topic so you know how to proceed qualitatively. Almost any research question in education research would benefit from an existing data source. Access to an existing study's protocols, guidelines, training materials, codebooks, and field notes can make you feel like you were part of the original research team. Therefore, consider locating an existing data source before agreeing to collect data for a new mixed methods project.

References

Andrew, S., Halcomb, E.J. (Eds.), 2009. Mixed Methods Research for Nursing and Health Sciences. Wiley-Blackwell Publishing, Chichester, UK.

Bazeley, P., 2018. Integrating Analyses in Mixed Methods Research. Sage Publications, Thousand Oaks, CA.

Beck, C.T., 2019. Secondary Qualitative Data Analysis in the Health and Social Sciences. Routledge, London, UK.

Bishop, E., 2009. Ethical sharing and reuse of qualitative data. Aust. J. Soc. Issues 44 (3), 255–272.

Bishop, L., 2007. A reflexive account of reusing qualitative data: beyond primary/secondary dualism. Socio. Res. Online 12 (3), 43–56.

Chawinga, W.D., Zinn, S., 2020. Research data management at an African medical university: implications for academic librarianship. J. Acad. Librariansh 46 (4), 102161.

Creswell, J.W., 2015. A Concise Introduction to Mixed Methods Research. Sage Publications, Thousand Oaks, CA.

Creswell, J.W., Plano Clark, V.L., 2018. Designing and Conducting Mixed Methods Research, third ed. Sage Publications, Thousand Oaks, CA.

Curry, L., Nunez-Smith, M., 2015. Mixed Methods in Health Sciences Research: A Practical Primer. Sage Publications, Thousand Oaks, CA.

Davis-Kean, P., Jager, J., 2017. Using secondary data analysis for educational research. In: Wyse, D., Selwyn, N., Smith, E., Suter, L.E. (Eds.), The BERA/SAGE Handbook of Educational Research: 2 Volume Set, first ed. Sage Publications, London, UK, pp. 505–522.

Dorn, E., Hancock, B., Sarakatsannis, J., Viruleg, E., 2020. COVID-19 and Student Learning in the United States: The Hurt Could Last a Lifetime. McKinsey & Company, p. 1.

Fielding, N.G., Fielding, J.L., 2000. Resistance and adaptation to criminal identity: using secondary analysis to evaluate classic studies of crime and deviance. Sociology 34 (4), 671–689.

Grinnell Jr., R.M., Unrau, Y.A., 2018. Social Work Research and Evaluation: Foundations of Evidence-Based Practice, eleventh ed. Oxford University Press, New York, NY.

Grooms, A.A., Childs, J., 2021. "We need to do better by kids": changing routines in U.S. schools in response to COVID-19 school closures. J. Educ. Stud. Placed A. T. Risk 26 (2), 135–156.

Guetterman, T.C., Fetters, M.D., Creswell, J.W., 2015. Integrating quantitative and qualitative results in health science mixed methods research through joint displays. Ann. Fam. Med. 13, 554–561.

Haight, W., Bidwell, L.N., 2016. Mixed Methods Research for Social Work: Integrating Methodologies to Strengthen Practice and Policy. Oxford University Press, Oxford, UK.

Hammersley, M., 2009. Can we re-use qualitative data via secondary analysis? Notes on some terminological and substantive issues. Socio. Res. Online 15 (1), 47–53.

Hesse-Biber, S.N., 2010. Mixed Methods Research: Merging Theory With Practice. Guilford Press, New York, NY.

Hewson, C., 2006. Secondary analysis. In: Jupp, V. (Ed.), The Sage Dictionary of Research Methods, first ed. Sage Publications, London, pp. 274–275.

Hinds, P., Vogel, R., Clarke-Steffen, L., 1997. The possibilities and pitfalls of doing secondary analysis of a qualitative data set. Qual. Health Res. 7 (3), 408–424.

Hughes, K., Tarrant, A., 2020. An introduction to qualitative secondary analysis. In: Hughes, K., Tarrant, A. (Eds.), Qualitative Secondary Analysis, first ed. Sage Publications, London, UK, pp. 3–18.

Jackson, J.S., Torres, M., Caldwell, C.H., Neighbors, H.W., Nesse, R.M., Taylor, R.J., Trierweiler, S.J., Williams, D.R., 2004. The national survey of American life: a study of racial, ethnic, and cultural influences on mental disorders and mental health. Int. J. Methods Psychiatr. Res. 13 (4), 196–207.

Johnson, R.B., Christensen, L.B., 2013. Educational Research: Quantitative, Qualitative, and Mixed Approaches. Sage Publications, Thousand Oaks, CA.

Johnston, M.P., 2012. School librarians as technology integration leaders: enablers and barriers to leadership enactment. Sch. Libr. Res. 15 (1), 691–693.

Johnston, M.P., 2014. Secondary method: a method of which the time has come. Qual. Quan. Methods Librar. 3 (3), 619–626.

Kiecolt, K.J., Nathan, L.E., 1985. Paper Series on Quantitative Applications in the Social Sciences: Secondary Analysis of Survey Data. Sage Publications, Newbury Park, CA.

Kimmons, R., 2022. Education Research Across Multiple Paradigms. EdTech Books.

Largan, C., Morris, T., 2019. Qualitative Secondary Research: A Step-by-Step Guide. Sage Publications, London, UK.

Logan, T., 2020. A practical, iterative framework for secondary data analysis in educational research. Aust. Educ. Res. 47 (3), 129–148.

Magee, T., Lee, S.M., Giuliano, K.K., Munro, B., 2006. Generating new knowledge from existing data: the use of large data sets for nursing research. Nurs. Res. 55 (2), S50–S56.

Mertens, D., 2009. Transformative Research and Evaluation. Guilford Press, New York.

Mertens, D.M., 2017. Evaluation in Practice Series: Mixed Methods Design in Evaluation. Sage Publications, Thousand Oaks, CA.

Morse, J.M., 2009. Mixing qualitative methods. Qual. Health Res. 19 (11), 1523–1524.

O'Cathain, A., 2010. Assessing the quality of mixed methods research: toward a comprehensive framework. In: Tashakkori, A., Teddlie, C. (Eds.), SAGE Handbook of Mixed Methods in Social & Behavioral Research, second ed. Sage Publications, Thousand Oaks, CA, pp. 531–558.

Panchenko, L., Samovilova, N., 2020. Secondary Data Analysis in Educational Research: Opportunities for Ph.D. Students, pp. 1–7. SHS Web of Conferences 75.

Plano Clark, V.L., Ivankova, N., 2016. Mixed Methods Research: A Guide to the Field. Sage, Thousand Oaks, CA.

Sheperis, C., Young, J., Daniels, M., 2016. Counseling Research: Quantitative, Qualitative, and Mixed Methods, second ed. Pearson, Boston, MA.

Smith, E., 2008. Using Secondary Data in Educational and Social Research. McGraw-Hill Education, New York, NY.

Smith, E., 2011. Special issue on using secondary data in educational research. Int. J. Res. Method Educ. 34, 219–221.

Tashakkori, A., Teddlie, C. (Eds.), 2010. SAGE Handbook of Mixed Methods in Social & Behavioral Research, second ed. Sage Publications, Thousand Oaks, CA, pp. 581–611.

Vassallo, P., 1999. The knowledge continuum - organizing for research and scholarly communication. Internet Res. 9 (3), 232–242.

Watkins, D.C., 2022. Secondary Data in Mixed Methods Research. Sage Publications, Thousand Oaks, CA.

Watkins, D.C., Gioia, D., 2015. Pocket Guides to Social Work Research Methods Series: Mixed Methods Research. Oxford University Press, New York.