

Exploratory Data Analysis in Schools: A Logic Model to Guide Implementation

Matthew B. Courtney, *Kentucky Department of Education*

Abstract

Exploratory data analysis (EDA) is an iterative, open-ended data analysis procedure that allows practitioners to examine data without pre-conceived notions to advise improvement processes and make informed decisions. Education is a data-rich field that is primed for a transition into a deeper, more purposeful use of data. This article introduces the concept of EDA as a necessary structure to be embedded in school activities by situating it within the literature related to data-driven decision-making, continuous school improvement systems, and action research methodologies. It also provides a succinct six-part framework to guide practitioners in establishing EDA procedures.

Keywords: continuous improvement; exploratory data analysis; data driven decision-making; data use; research use

Introduction

Exploratory data analysis (EDA) is an iterative process that allows users to examine a large volume of data quickly and meaningfully to better understand and utilize that data for decision-making. Originally conceptualized by the renowned statistician John Tukey (1977), EDA utilizes statistical calculations and data visualization methods to examine data with an open mind. Contrary to scientific data analysis

Matthew B. Courtney. (2021). Exploratory Data Analysis in Schools: A Logic Model to Guide Implementation. *International Journal of Education Policy & Leadership* 17(4). URL: <http://journals.sfu.ca/ijepl/index.php/ijepl/article/view/1041> doi:10.22230/ijepl.2021v17n4a1041

IJEPL is a joint publication of PDK International, the Faculty of Education at Simon Fraser University, the College of Education and Human Development at George Mason University, and the University of Delaware. By virtue of their appearance in this open access journal, articles are free to use, with proper attribution, in educational and other non-commercial settings 90 days after initial publication. Copyright for articles published in IJEPL is retained by the authors. More information is available on the IJEPL website: <http://www.ijepl.org>



(also called confirmation data analysis), where the user is analyzing a set of collected data to answer a pre-determined question, the EDA process is a discovery process, where the user gathers information to inform a hypothesis that may be tested or explored later on (Bezerra, Silva, Guedes, Silva, Leitão, & Saito, 2019; Selwyn, Henderson, & Chao, 2017). It is a process that allows the user to observe unexpected patterns and consider the impact of those patterns (Behrens, 1997).

EDA is an especially useful tool for practitioners due to its more implicit foundational principles—namely the fact that EDA is process oriented over theory oriented, is rooted in relatively simple mathematical concepts, and is flexible and iterative in nature (Tukey, 1993). There is no right or wrong way to conduct an EDA process. The key is to keep an open mind and to test different modeling techniques until new information about the data is uncovered. To quote John Tukey, “EDA is an attitude, a state of flexibility, a willingness to look for those things that we believe are not there, as well as those we believe to be there” (Tukey quoted in Jones, 1986, p. 806).

As educators continue to use data to drive decision-making, reflective practitioners should embrace the principles of data science (Bill & Melinda Gates Foundation, 2015; Daniel, 2019). Researcher and data scientist Vasant Dhar (2013) provides a succinct definition of data science: “data science is the study of the generalizable extraction of knowledge from data” (p. 64). Rather than following traditional research protocols to collect data related to a specific question, data science relies on the analysis of vast quantities of data that are strategically and continually collected by an organization to answer in-the-moment questions. EDA is a safe entry point into data science for practitioners looking to enhance their data use. In the sections that follow, this article will make the case that the education field is primed for an expansion into the world of data science and that this expansion will help remedy existing data use issues within established systems and provide leadership with greater insight into the impact of school and district activities.

Data use within education

Education systems have been systemically collecting data for both reporting and continuous improvement purposes for many years (Cannata, Redding, & Rubin, 2012; Data Quality Campaign, 2012). While these systems have been built to meet the needs of their organizations, they have also been built to meet the requirements of effective data science systems. Effective data science requires data collection systems that include a large volume of data, varied data points, and a high velocity of new information (Wang, 2017). The education system has an incredible volume of data due to efforts to collect and archive data for public dissemination. In the United States, this initiative is called ED*Facts* and results in thousands of annually archived data points about schools from all fifty states (U.S. Department of Education, 2020). Similar initiatives have been undertaken by other governments, including in the United Kingdom (United Kingdom Data Service, 2020) and Australia (Australian Bureau of Statistics, 2020). These same functions ensure that the education data available is sufficiently varied and collected at a high velocity. Schools generate, collect, and archive a wide variety of data for local purposes in addition to federally mandated reporting. This body of data includes achievement data, poverty data,

race/ethnicity data, benchmark data, program evaluation data, and even social media data (Makela & Hoff, 2018; Selwyn et al., 2017; U.S. Department of Education, 2020a; U.S. Department of Education, 2020b). Existing educational data systems are already set up for the implementation of regular EDA procedures.

Educators do not just collect this data, they use it. The field is already steeped in a culture of data-driven decision-making (DDDM) (Cohen-Vogel & Harrison, 2013). The ability to deploy DDDM methodologies is widely considered to be a pre-requisite skill for education leaders (Wang, 2019), a valuable component of professional learning community (PLC) meetings (Dufour, DuFour, Eaker, & Many, 2006), and a way of driving instructional change (Bill & Melinda Gates Foundation, 2015). DDDM is, however, fraught with conflict, leading some leaders to question its underlying moral implications (Wang, 2019) and others to question whether educators have the necessary skills to navigate the data (Bill & Melinda Gates Foundation, 2015).

One of the biggest concerns with current thinking on DDDM is that theoretical models lack specificity (Carrier & Whaland, 2017) and leaders lack the assessment literacy to complete DDDM tasks on their own (Bill & Melinda Gates Foundation, 2015; Crum, 2009). In schools where DDDM systems are in place, they are generally led by district-level administrators and tied to specific goals, outcomes, or the formal accountability structures stemming from high-stakes testing (Carrier & Whaland, 2017; Noyce, Perda, & Traver, 2000). Left untended, DDDM procedures have little impact on student outcomes, but some research has shown that, with training, impacts on student achievement can be seen (Carrier & Whaland, 2017; Crone, Carlson, Haack, Kennedy, Baker, & Fine, 2016; Keuning, van Geel, Visscher, & Fox, 2016). In a best-practice scenario, DDDM models can be effectively deployed when leadership makes it a priority and develops an open and collaborative school culture (Lange, Range, & Welsh, 2012). Collaborative teams should be reviewing a wide range of data, including student demographics, learning data, process data, and perception data (Crone et al., 2016; Lange et al., 2012).

EDA is itself a DDDM framework that provides a structured method to help education leaders dig deeper into their data. Rather than focusing on merely interpreting the results of the past, EDA and other data science techniques merge a wide range of skill sets—from basic analytics and statistics to complex machine learning and artificial intelligence—to predict the future impact of a decision (Dhar, 2013). The value of EDA as a DDDM framework lies in its open-endedness and iterative nature. It is designed to answer questions that the practitioner has not even thought to ask (Tukey, 1977). Most important to the effective use of EDA is the underlying practitioner knowledge that gives the results meaning and drives the necessary change (Bill & Melinda Gates Foundation, 2015; Dhar, 2013; Tichnor-Wagner, Wachen, Cannata, & Cohen-Vogel, 2017).

EDA within continuous improvement constructs

The term “continuous improvement” describes a state in which an organization seeks to improve outcomes through rigorous and systemic evaluation, reflection, and adjustment (Park, Hironaka, Carver, & Nordstrum, 2013). It is not merely the writing of an annual improvement plan, rather it is the continuous monitoring of inputs and outputs,

and repeated tests of small changes to existing processes, that lead to lasting and sustainable change in an organization (Park et al., 2013; Tichnor-Wagner et al., 2017).

When it comes to the implementation of continuous improvement systems, leadership and stakeholder engagement appear to be key (Mac Iver, Sheldon, Epstein, Rice, Mac Iver, & Simmons, 2018; Park et al., 2013), and successful implementation mirrors the same needs as successful DDDM structures. Leaders must work to develop and maintain a learning mindset and remain flexible to change throughout the continuous improvement process (Park et al., 2013). Teachers also play a key role in managing the continuous improvement process in schools, and their investment is a prerequisite for the successful implementation of continuous improvement systems (Cannata et al., 2012; Devaney, Smith, & Wong, 2012). Improvement systems work best when communication between stakeholders is clear and practitioner knowledge is blended with scientific skills to measure meaningful outcomes (Park et al., 2013; Tichnor-Wagner et al., 2017). When built and maintained with intentionality, continuous improvement systems have been shown to lead to lasting change in school processes and student outcomes (Park et al., 2013; Smith, Akiva, Blazeviski, Devaney, & Pelle, 2008; Vaszauskas, 2011).

Similar to DDDM structures, the concepts of continuous improvement are not new; however, they remain difficult for many to implement effectively (Kaufman, Cash, Coartney, Ripley, Guy, Glenn, Mitra, & Anderson, 2020). One key element to the successful implementation and sustainability of continuous improvement is organizational infrastructure (Park et al., 2013); this has led to several models designed to make continuous improvement more approachable. One common continuous improvement model is the plan-do-study-act (PDSA) cycle (Deming, 2000; Tichnor-Wagner et al., 2017; Vaszauskas, 2011). In a PDSA cycle, leaders or teachers identify a problem, plan an intervention, implement an intervention, study the impact of the intervention, and act on the results of their study to improve the impact of the intervention. The cycle then begins again, with the team planning for a new intervention period that implements their improvements.

The successful implementation of the PDSA model sets the stage for effective data science systems discussed earlier. When implemented with fidelity, these systems continuously and rapidly produce a large volume of varied data that can be deeply analyzed. However, as has been discussed, educators are largely unequipped to use that data to effectively inform their work (Bill & Melinda Gates Foundation, 2015; Dhar, 2013; Tichnor-Wagner et al., 2017). The addition of flexible EDA procedures may help solve this issue by providing a guidepost for practitioners as they seek to apply their data skills.

EDA as an action research methodology

Action research has emerged as a powerful continuous improvement tool that is frequently deployed in the education field. In his book, *Guiding School Improvement with Action Research*, Richard Sagor (2000) defines action research as “a disciplined process of inquiry conducted by and for those taking the action. The primary reason for engaging in action research is to assist the ‘actor’ in improving and/or refining his or her actions” (n.p.).

This definition succinctly captures the essence of action research. It is research conducted by practitioners for the purpose of informing a decision. In the context of a continuous improvement process, that decision could be an attempt to understand the underlying factors that contribute to an identified problem, a proposed innovation to solve an identified problem, or the evaluation of an innovation after it has been fully implemented.

Many authors have proposed models for the implementation of action research projects (Keegan, 2016; Levesque, Fitzgerald, & Pfeiffer, 2015; Sagor, 2000; Stringer, 2008). While these models vary in the number of steps and the necessary duration of research projects, action research models generally follow the same basic process as more formal experimental research designs. First, the practitioner identifies a problem to investigate. Next, they use existing literature to inform the specific research questions that will be examined. Finally, data is systemically gathered and analyzed before the reflective practitioner determines their next steps. As with the other constructed frameworks discussed in this article, a key feature of action research is its cyclical nature, where a practitioner's determined next steps become a research problem to be explored later (Keegan, 2016).

Action research has a multitude of benefits for the education practitioner beyond the continuous improvement process. Action research helps to build capacity in educators by bridging the gap between the world of academia and the practical experience earned in the field (Amir, Mandler, Hauptman, & Gorev, 2017). When conducted systemically and with fidelity, action research projects allow practitioners to test and verify research findings conducted in different settings and help leaders at the classroom and organization level establish a larger understanding of the field (Carver & Klein, 2013; Lee, Sachs, & Wheeler, 2014). Action research projects have been shown to improve teacher and leader reflective practice (Amir et al., 2017; Carver & Klein, 2013) and reduce feelings of isolation (Meyer-Looze, 2015). Action research can be conducted alone or as a leadership team. When teams of practitioners collaborate on action research projects, schools and districts can begin to develop a meaningful, system-wide focus on continuous improvement that establishes a lasting culture of evidence-based decision-making and teaching (Sato & Loewen, 2019; Zambo, 2011). Some research also suggests that students can benefit from directly participating in the action research process with their teachers (Martin & Bridgmon, 2009).

The successes documented by action research procedures provide promising evidence that a set of steps can guide education practitioners toward more thoughtful and deliberate decision-making. While action research fits neatly into the continuous improvement cycle, it is rigid in nature and seeks to answer a pre-determined question. EDA, by contrast, is open-ended and does not seek to answer pre-determined questions. Rather, it seeks to expose the answers to questions that practitioners may have never thought to ask in the first place (Tukey, 1993).

Therein lies the value of EDA. This article has established that education practitioners have access to a varied and voluminous data set that is ever growing through existing continuous improvement efforts. It has also established that, while the field has a variety of existing structures that rely on data analysis, educators struggle to effectively use data to inform their work without a clear framework to guide them. If action re-

search is accepted to be a successful model for rigid, formal program evaluation, then, this article posits, EDA is its flexible, iterative, informal, and necessary counterpart.

An EDA framework for schools

Before schools can begin to effectively implement EDA procedures, a framework for implementation must be established. A clear framework must begin with an examination of expected timelines. While limited research exists on the benefits of the regular, open-ended type of data analysis proposed here, research into the PDSA cycle suggests that a cycle runs for roughly ninety days (Tichnor-Wagner et al., 2017). In the United States, the average school year is 180 days (U.S. Department of Education, 2020). If this information is applied to the EDA logic model proposed here, schools should build time for regular EDA processes twice each school year. Presumably, the ideal time to perform these analyses would be at the start of each semester, although there is no current research to confirm the effectiveness of that timeline.

The proposed framework would also benefit from a brief discussion about the stakeholders that should be involved in the process. Previously cited research suggests that relevant frameworks (continuous improvement systems, action research projects, and DDDM processes) benefit from the involvement of a wide range of stakeholders, including system- and building-level administrators, classroom teachers, parents and community stakeholders, and students. It is unrealistic to expect that each of these four categories of stakeholders would be directly involved in the data-analysis process, so stakeholder engagement must be more clearly defined. It is suggested here that the act of completing the EDA process is done at the building-administrator level, and the results of the analysis are communicated in detail to classroom teachers and summarized for students, parents, and community stakeholders.

There is precedent and literature support for this stakeholder involvement pattern. The role of leadership in implementing new changes is vital to the success of those changes (Park et al., 2013). Leadership must engage meaningfully in the process and research already suggests that principals are performing this function and find value in data analysis (Militello, Bass, Jackson, & Wang, 2013). Once data has been analyzed, a deep-level analysis should be shared with the classroom teachers. Research has demonstrated that teacher data use can be cultivated but must be done so in a focused and intentional way that includes guidance and coaching by leadership (Huguet, Marsh, & Farrell, 2014; Van Gasse, 2019). Finally, anonymized summary data should be shared with the community, reflecting existing school report card protocols.

Having established a timeline and stakeholder involvement scheme, the final component of the framework involves outlining the analytic steps necessary to complete an EDA process. Data scientists usually perform EDA processes using a suite of sophisticated data tools, including R, Python, or STATA software. While these tools are very powerful, the average school-level practitioner needs to work with the data analysis tools available in most spreadsheet software. Fortunately, EDA methods are flexible; the framework is adaptable to the technical skill level of the practitioner. While the specific steps taken by practitioners will vary from building to building, this article will highlight six steps that all education EDA procedures should include

1) gathering necessary data, 2) exploring categorical variables, 3) calculating descriptive statistics, 4) creating data visualizations, 5) examining correlations, and 6) interpreting the results for decision-making.

Gathering the necessary data

The first step in completing an EDA process is to gather and prepare the necessary data for analysis. Ideally, practitioners would take time to secure a wide variety of data around the subject they wish to examine. For example, if an EDA process is to examine the students in a school building, the data set should include all available data related to each student, such as their demographics, attendance, behavior, or achievement. Data collected to explore an issue may look different. If a practitioner wants to know more about the state of English language learners, for example, they will be better served by accessing a variety of data related to immigration, international performance measures, language acquisition, and student demographic datasets housed in federal databases.

Care should be taken to prepare the data in a way that is clean and can be easily interpreted. Hadley Wickham (2014) proposes a set of principles he calls “tidy data.” Tidy data can be summed up by three elements: 1) each variable forms a column, 2) each observation forms a row, and 3) each observational unit forms a table. In the example presented before of analyzing students in a building, the observational unit is the individual student; the variable is the demographic data, behavior data, or assessment data linked to that student; and the observational unit is the actual score associated with both the student and the variable.

Exploring categorical variables

Having gathered the necessary data and ensured its *tidiness*, the education practitioner should begin their EDA process by examining the categorical variables housed within the dataset. Categorical variables, sometimes called discrete variables, are those that can be counted (Martin & Bridgmon, 2012). In school-level data sets, these variables typically include demographic values, such as gender, grade level, race/ethnicity, and eligibility for programs for special education, migrant, or economically disadvantaged populations. Examining categorical variables and groups can be easily done in most spreadsheet software with the =COUNTIF function. This function makes it possible to quickly and easily count and sort variables by telling the spreadsheet to count the number of times a word or phrase shows up in a data set.

An important component of school-level data analysis is the comparison of performance between groups. Before those comparisons can be made, the practitioner must first determine who the groups are. This is easily done in spreadsheets using the =COUNTIFS function. Again, this function counts the number of times a word or phrase appears in a data set, but it allows the user to add conditions to the search. This is an especially useful tool for creating demographic summary tables. Practitioners should begin their EDA process by creating tables that allow them to summarize their data set. An example of this table is included in Table 1, which reports the number of students at each grade level by race/ethnicity. Tables such as this could be created to report on any combination of groups.

Table 1. Sample demographic summary table

Grade	African/American	Asian	Hispanic/Latino	White
Six	25	11	32	95
Seven	29	9	15	84
Eight	22	7	27	102

Once variables have been counted, practitioners should convert those counts to percentages or ratios. This will help the practitioner to get a better understanding of the makeup of their schools or classrooms. These ratios will provide valuable information for the practitioner later in the EDA process, when elements related to student achievement are explored.

Calculate descriptive statistics

After having completed a thorough summary of the categorical variables contained within their data set, the education practitioner should examine the continuous variables with descriptive statistics. Continuous variables exist along a spectrum; in education they tend to be variables such as grades, test scores, behavior referrals, or absences (Martin & Bridgmon, 2012). Descriptive statistics provide a summary of a distribution of scores. Generally, an EDA process should include the calculation of the mean, median, mode, range, and standard deviation; although other descriptive statistics may be useful to more advanced practitioners (Martin & Bridgmon, 2012).

Just as with the examination of continuous variables, the calculation of descriptive statistics is relatively simple in most spreadsheet software. Built in functions, such as =AVERAGE or =STDEV, are easily applied to user-generated tables. Some software also includes add-in tools to calculate multiple descriptive statistics at once, such as the Analysis ToolPak in Excel or the XLMiner Analysis ToolPak in Google Sheets. Practitioners should use these tools to build summary tables that allow them to see an overview of their distributions. Practitioners will also be well served by using these functions to build tables that allow them to see how the previously identified student groups performed when compared to one another. In most spreadsheet software, this can be easily done by adding "IF" to the end of the function, such as =AVERAGEIF. This function will take the average of all the numbers in a distribution, but only if they meet a specific condition that has been set. The goal is to gain as much information about the data set as possible. The practitioner is not looking for answers but mining the data for new information.

Data visualizations

Data visualizations are an important part of EDA. Graphic displays have been shown to help develop a user's understanding of data and effectively communicate new information. These techniques have been common in other fields, such as psychology and business analytics, for many years (Alhadad, 2018; Diamond & Mattia, n.d.). Practitioners should start by creating data visualizations for their summary table. Common plots, such as pie charts, bar charts, scatter plots, and line graphs, can be used to better understand the meaning of summary charts. These visualizations are easy to make in most spreadsheet software.

Two other data visualizations that can be helpful are box-and-whisker plots and histograms. These two plots apply statistics to the visualization process by describing the distribution. Box-and-whisker plots help to depict the full range of a distribution by clearly showing the highest and lowest score in relation to the median and upper and lower quartiles. Similarly, histograms place scores into bins and graph them with bars (Martin & Bridgmon, 2012). These plots can be used to help practitioners see relationships between groups and identify instructional inequities. During this exploratory process, practitioners should plot all continuous variables to examine underlying relationships and look for gaps.

Examining correlations

The final step in the proposed EDA framework is the examination of correlations. Correlations identify the strength of relationships between variables. While there are multiple ways to calculate correlations, the formula built into most spreadsheet software is the Pearson product moment correlation. This formula produces a correlation score on a range of negative one to positive one, with zero representing no correlation. The closer to the extremes the score, the stronger the relationship. Correlations can only report the strength of a relationship and do not report on causality. It is important for practitioners to consider the results in light of that fact (Martin & Bridgmon, 2012).

Practitioners with a sufficiently robust data set should find value in the creation of a correlation matrix, again performed quite easily in most spreadsheet software. These matrices place the correlations between all variables in one chart with a few simple clicks. Practitioners should build a correlation matrix using all variables. In some cases, practitioners may need to transform string variables (non-numeric variables such as text) into numerical variables through a coding process.

Interpreting results for decision-making

In the end, the individual results of an EDA procedure are meaningless unless they are interpreted and used to inform decision-making. The goal of EDA is not to find answers to questions. Rather, it is to identify previously unknown points of information (Tukey, 1977). Practitioners should examine the results of each analytic test at face value and look for triangulation in the data. Triangulation is a concept applied in the social sciences that involves examining an issue through multiple tests and lenses. The underlying premise is that trends that can be spotted in multiple places have more meaning than trends identified in only one outcome (Given, 2008). When the practitioner begins to triangulate findings, they can be sure the information is meaningful to their situation and will prove worthy of further exploration or decisive action.

Creating a culture of data analysis

While there are many benefits to deploying EDA procedures within the context of continuous school improvement, there are undoubtedly many challenges as well. Leaders seeking to implement EDA procedures in their schools should consider these challenges and be proactive in addressing them.

As with any continuous improvement initiative, school leaders must foster a culture in which EDA can thrive. Data use must be a regular part of the school's continuous improvement conversation and be embedded in the underlying beliefs of the school before it can become meaningful to staff (Day & Sammons, 2013; Gerzon & Guckenburg, 2015; Schildkamp & Datnow, 2020). The first step in this process is to approach data work with an open and collaborative mindset. Teachers must feel comfortable engaging in the work, and a risk-free environment should be cultivated (Danley, 2020; Lange, Range, & Welsh, 2012; Schildkamp, Poortman, Ebbeler, & Pieters, 2019; Schwanenberger & Ahearn, 2013). It is important that analysis teams have protected time to do their work and that school leaders support their data teams by providing clear expectations, structures, and resources to guide their efforts (Gerzon & Guckenburg, 2015; Kekahio & Baker, 2013; Schildkamp et al., 2019). Above all else, school leaders must ensure that EDA work focuses on continuous improvement over accountability. Leaders should refrain from using data to create blame, guilt, or hostility. A focus on compliance and accountability over improvement and learning will not lead to the authentic use or meaningful adoption of the protocols (Park, Daly, & Guerra, 2012; Schildkamp et al., 2019).

In building a culture of data use, school leaders should ensure that appropriate steps are taken to acknowledge and address confirmation bias. Teachers often rely on data that they were not a part of gathering or may not fully understand. As has been discussed, they also frequently lack structures to lead their analytic process. This forces them to rely on their own prior knowledge or experiences when interpreting data (Schildkamp, 2019; Vanlommel, Vanhoof, & van Petegem, 2016). Educators should seek to complete EDA processes with an open mind. The protocols outlined in this logic model are constructed in such a way as to intentionally prevent the analyst from seeking answers to specific questions. By keeping the process fluid and open-ended, practitioners can help to limit occurrences of confirmation bias.

Finally, the acquisition of necessary technical skills is vital in the development and successful implementation of EDA protocols. A 2015 report released by the Bill and Melinda Gates Foundation found that current data systems are often overwhelming to teachers due to the large amounts of data from multiple sources and the frequent incompatibility of tools and data sets. Other research previously discussed in this article examines similar challenges (Dhar, 2013; Schildkamp & Datnow, 2020; Tichnor-Wagner et al., 2017). The beauty of EDA is that it can yield quick insights with little skill, as it is process-focused rather than theory-focused (Tukey, 1977). Educators can easily learn the handful of entry-level analysis techniques listed here (descriptive statistics, correlations, and basic visualizations) in a short amount of time. As educators continue to deploy EDA methodologies, further professional learning will allow them to dig deeper and uncover new insights (Gerzon & Guckenburg, 2015; Lange, Range, & Welsh, 2012).

Conclusion

This article used existing literature to provide a rationale and implementation logic model for the incorporation of EDA as a routine school activity. While this article is heavily rooted in existing literature, it is unable to say with certainty that the systemic

implementation of an EDA process will improve educational decision-making or school outcomes. Future research will need to be conducted to examine the validity, replicability, and impact of this model. The model is also not without limitations. Its success hinges on the analytic ability of the practitioner and their underlying understanding of the potential ethical concerns of this type of data use, including student privacy (Selwyn et al., 2017; Wang, 2017).

References

- Alhadad, S.S.J. (2018). Visualizing data to support judgement, inference, and decision making in learning analytics: Insights from cognitive psychology and visualization science. *Journal of Learning Analytics*, 5(2), 60–85. doi:10.18608/jla.2018.52.5
- Amir, A., Mandler, D., Hauptman, S., & Gorev, D. (2017). Discomfort as a means of pre-service teachers' professional development — an action research as part of the “Research Literacy” course. *European Journal of Teacher Education*, 40(2), 231–245. doi:10.1080/02619768.2017.1284197
- Australian Bureau of Statistics. (2020). *Education*. Canberra, AU: Australian Bureau of Statistics. Retrieved March 26, 2020, from <https://www.abs.gov.au/education>.
- Bezerra, A., Silva, I., Guedes, L.A., Silva, D., Leitão, G., & Saito, K. (2019). Extracting value from industrial alarms and events: A data-drive approach based on EDA. *Sensors*, 19(12). doi:10.3390/s19122772
- Behrens, J.T. (1997). Principles and procedures of exploratory data analysis. *Psychological Methods*, 2(2), 131–160.
- Bill & Melinda Gates Foundation. (2015). *Making data work for teachers and students*. Seattle, WA: Bill & Melinda Gates Foundation.
- Cannata, M., Redding, C., & Rubin, M. (2016). *Continuous improvement in action: Educators' evidence use for school improvement*. [Abstract.] Washington, DC: SREE Spring 2016 Conference.
- Carrier, L.L., & Whaland, M. (2017). Left behind by policy: A case study of the influence of high stakes accountability policy on data-based decision making in one small, rural New Hampshire school. *The Rural Educator*, 38(3), 12–26.
- Carver, C.L., & Klein, C.S. (2013). Action research: A tool for promoting faculty development and continuous improvement in leadership preparation. *NCPEA International Journal of Educational Leadership Preparation*, 8(2), 162–177.
- Cohen-Vogel, L., & Harrison, C. (2013). Leading with data: Evidence from the National Center on scaling up effective schools. *Leadership and Policy in Schools*, 12(2), 122–145. doi:10.1080/15700763.2013.792934
- Courtney, M.B. (2020). The education agency as research intermediary. *Journal of Educational Issues*, 6(1), 1–9.
- Crone, D.A., Carlson, S.E., Haack, M.K., Kennedy, P.C., Baker, S.K., & Fine, H. (2016). Data-based decision-making teams in middle school: Observations and implications from the middle school intervention project. *Assessment for Effective Intervention*, 41(2), 79–93. doi:10.1177/1534508415610322
- Crum, K. (2009). Building the foundation for data-based decision making: Creating consensus on language and concepts. *International Electronic Journal for Leadership in Learning*, 13(5), 1–23.
- Daniel, B.K. (2019). Big data and data science: A critical review of issues for educational research. *British Journal of Educational Technology*, 50(1), 101–113. doi: 10.1111/bjet.12595
- Danley, A.J. (2020). Providing teacher candidates experience in data teams: Analyzing literacy data. *SRATE Journal*, 29(1), 1–8.
- Data Quality Campaign (2012). Quality data is key to improving education. *Journal of Applied Research on Children: Informing Policy for Children at Risk*, 3(2), 1–3.
- Day, C., & Sammons, P. (2013). *Successful leadership: A review of the international literature*. Reading, UK: CfBT Educational Trust.

- Deming, W.E. (2000). *The new economics for industry, government, and education* (2nd ed.). Cambridge, MA: MIT Press.
- Devaney, E., Smith, C., & Wong, K. (2012). Understanding the “how” of quality improvement: Lessons from the Rhode Island program quality intervention. *Afterschool Matters, Fall*(16), 1–10.
- Dexter, B., & Seden, R. (2012). “It’s really making a difference”: How small-scale research projects can enhance teaching and learning. *Innovations in Education and Teaching International, 49*(1), 83–93. doi:10.1080/14703297.2012.647786
- Dhar, V. (2013). Data science and prediction. *Communications of the ACM, 56*(12), 64–73. doi:10.1145/2500499
- Diamond, M., & Mattia, A. (n.d.). Data visualization: An exploratory study into the software tools used by businesses. *Journal of Instruction Pedagogies, 17*, 1–7.
- DuFour, R., DuFour, R., Eaker, R., & Many, T. (2006). *Learning by doing: A handbook for professional learning communities at work*. Thousand Oaks, CA: Corwin Press.
- Gerzon, N., & Guckenburg, S. (2015). *Toolkit for a workshop on building a culture of data use* (REL 2015-063). Washington, DC: U.S Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast & Islands. Retrieved March 23, 2020, from <http://ies.ed.gov/ncee/edlabs>.
- Given, L.M. (2008). *The SAGE encyclopedia of qualitative research methods*. Thousand Oaks, CA: Sage Publications.
- Huguet, A., Marsh, J.A., & Farrell, C.C. (2014). Building teachers’ data-use capacity: Insights from strong and developing coaches. *Education Policy Analysis Archives, 22*(52), 1–31. doi:10.14507/epaa.v22n2.2014
- Jones, L.V. (Ed). (1986). *The collected works of John W. Tukey: Philosophy and principles of data analysis 1949–1964*. London, UK: Chapman & Hall.
- Kaufman, E.K., Cash, C.S., Coartney, J.S., Ripley, D., Guy, T.M., Glenn, W.J., Mitra, S., & Anderson, J.C. (n.d.). Planning to create a culture of continuous improvement with the Department of Defense education activity. *Education Planning, 26*(4), 5–19.
- Keegan, R. (2016). Action research as an agent for enhancing teaching and learning in physical education: A physical education teacher’s perspective. *The Physical Educator, 73*(2), 255–284. doi:10.18666/TPE-2016-V73-I2-6236
- Kekahio, W., & Baker, M. (2013). *Five steps for structuring data-informed conversations and action in education* (REL 2013-001). Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Pacific. Retrieved March 23, 2020, from <http://ies.ed.gov/ncee/edlabs>.
- Keuning, T., van Geel, M., Visscher, A., & Fox, J-P. (2016). *The effects of a schoolwide data-driven decision making intervention on elementary schools’ student achievement growth for mathematics and spelling*. Washington, DC: SREE Spring 2016 Conference.
- Lange, C., Range, B., & Welsh, K. (2012). Conditions for effective data use to improve schools: Recommendations for school leaders. *International Journal of Educational Leadership Preparation, 7*(3). *The Connexions Project*, 1–11.
- Lee, J.S., Sachs, D., & Wheeler, L. (2014). The crossroads of teacher leadership and action research. *The Clearing House, 87*(5), 218–233. doi:10.1080/00098655.2014.924896
- Levesque, K., Fitzgerald, R., & Pfeiffer, J. (2015). *A guide to using state longitudinal data for applied research*. Washington, DC: US Department of Education, Institute of Educational Science, National center for Education Evaluation and Regional Assistance, Analytic Technical Assistance and Development. Retrieved March 23, 2020, from <https://ies.ed.gov/ncee/pubs/20154013/>
- Mac Iver, M.A., Sheldon, S., Epstein, J., Rice, E., Mac Iver, D., & Simmons, A. (2018). Engaging families in the high school transition: Initial findings from a continuous improvement initiative. *School Community Journal, 28*(1), 37–66.
- Makela, J.P., & Hoff, K. (2018). Career outcomes data from social media: Examining quality in current practices. *The Career Development Quarterly, 67*(3), 220–235. doi:10.1002/cdq.12192

- Martin, W.W., & Bridgmon, K.D. (2012). *Quantitative and statistical research methods*. San Francisco, CA: Jossey-Bass.
- Meyer-Looze, C.L. (2015). Creating a cycle of continuous improvement through instructional rounds. *NCPEA International Journal of Educational Leadership Preparation*, 10(21), 29–45.
- Militello, M., Bass, L., Jackson, K.T., & Wang, Y. (2013). How data are used and misused in schools: Perceptions from teachers and principals. *Education Sciences*, 2013(3), 98–120. doi:10.3390/educsci3020098
- Noyce, P., Perda, D., & Traver, R. (2000). Creating data-driven schools. *Educational Leadership*, 57(5), 52–56.
- Park, S., Hironaka, S., Carver, P., & Nordstrum, L. (2013). *Continuous improvement in education*. Stanford, CA: Carnegie Foundation for the Advancement of Teaching.
- Park, V., Daly, A.J., & Guerra, A.W. (2012). Strategic framing: How leaders craft the meaning of data use for equity and learning. *Educational Policy*, 27(4), 645–675. doi:10.1177/0895904811429295
- Sagor, R. (2000). *Guiding school improvement with action research*. Alexandria, VA: ASCD.
- Sato, M., & Loewen, S. (2019). Do teachers care about research? The research-pedagogy dialogue. *ELT Journal*, 73(1), 1–10.
- Schildkamp, K. (2019). Data-based decision-making for school improvement: Research insights and gaps. *Educational Research*, 61(3), 257–273. doi:10.1080/00131881.2019.1625716
- Schildkamp, K., & Datnow, A. (2020). When data teams struggle: Learning from less successful data use efforts. *Leadership and Policy in Schools*. doi:10.1080/15700763.2020.1734630
- Schildkamp, K., Poortman, C.L., Ebbeler, J., & Pieters, J.M. (2019). How school leaders can build effective data teams: Five building blocks for an e wave of data-informed decision making. *Journal of Educational Change*, 2019(20), 283–325.
- Schwanenberger, M., & Ahearn, C. (2013). Teacher perceptions of the impact of the data team process on core instructional practices. *NCPEA International Journal of Educational Leadership Preparation*, 8(2), 146–162.
- Selwyn, N., Henderson, M., & Chao, S.H. (2017). The possibilities and limitations of applying open data principles in schools. *Cambridge Journal of Education*, 47(2), 167–187. doi:10.1080/0305764X.2016.1143449.
- Smith, C., Akiva, T., Blazeovski, J., Devaney, T., & Pelle, L. (2008). *Final report on the Palm Beach quality improvement system pilot*. Ypsilanti, MI: Center for Youth Program Quality.
- Stringer, E.T. (2008). *Action research in education* (2nd ed.). Saddle River, NJ: Pearson/Merrill/Prentice Hall.
- Tichnor-Wagner, A., Wachen, J., Cannata, M., & Cohen-Vogel, L. (2017). Continuous improvement in the public school context: Understanding how educators respond to plan-do-study-act cycles. *Journal of Education Change*, 18(4), 465–494. doi:10.1007/s10833-017-9301-4.
- Tukey, J.W. (1977). *Exploratory data analysis*. Boston, MA: Addison-Wesley.
- Tukey, J.W. (1993). *Exploratory data analysis: Past, present, and future* [Technical report No. 2.]. Princeton, NJ: Department of Statistics: Princeton University.
- U.S. Department of Education. (2020). *EdFacts initiative*. Washington, DC: U.S. Department of Education. Retrieved March 26, 2020, from <https://www2.ed.gov/about/inits/ed/edfacts/index.html>.
- U.S. Department of Education. (2020a). *Common core of data*. Washington, DC: U.S. Department of Education. Retrieved March 26, 2020, from <https://nces.ed.gov/ccd/>.
- U.S. Department of Education. (2020b). *National Center for Education statistics, schools and staffing survey* (SASS) [Public school data file]. Washington, DC: U.S. Department of Education.
- United Kingdom Data Service. (2020). *Education*. Colchester, UK: United Kingdom Data Service. Retrieved March 26, 2020, from <https://www.ukdataservice.ac.uk/get-data/themes/education.aspx>.
- Van Gasse, R. (2019). The effect of formal team meetings on teachers' informal data use interactions. *Frontline Learning Research*, 7(2), 40–56. doi:10.14786/frl.v7i2.443

- Vanlommel, K., Vanhoof, J., & van Petegem, P. (2016). Data use by teachers: The impact of motivation, decision-making style, supportive relationships, and reflective capacity. *Educational Studies*, 42(1), 36–53. doi:10.1080/03055698.2016.1148582
- Vaszuskas, J. (2011). Education change, Baldrige, and Schlechty. *Scholarly-Practitioner Quarterly*, 5(2), 114–134.
- Wang, Y. (2017). Education policy research in the big data era: Methodological frontiers, misconceptions, and challenges. *Education Policy Analysis Archives*, 25(94), 1–21. doi: 10.14507/epaa.25.3037
- Wang, Y. (2019). Is data-driven decision making at odds with moral decision making? A critical review of school leader's decision making in the era of school accountability. *Values and Ethics in Educational Administration*, 14(2), 1–9.
- Wickham, H. (2014). Tidy data. *Journal of Statistical Software*, 59(10), 1–23.
- Zambo, D. (2011). Action research as signature pedagogy in an education doctorate program: The reality and hope. *Innovative Higher Education*, 36(4), 261–271.