CIPP

Evaluation During Disruptions:
Course Corrections and
Other Considerations

Hadley Moore
Sarah Heinemeier
Margaret Gillis
Debra Shaver
Jennifer Schaaf
Jill Lammert

Evaluation During Disruptions:
Course Corrections and Other Considerations

What should you do when a large-scale disruption makes your evaluation impossible to implement as planned? Evaluators across the nation are facing this problem as a result of the extraordinary circumstances associated with the COVID-19 pandemic. However, disruptions can also occur due to hurricanes, earthquakes, wildfires, floods, as well as more localized events like a roof collapse, extreme temperatures, or large-scale administrative changes. With disruptions like these, you may lose participants or an entire intervention site; be unable to gather formative, fidelity, or outcome data; or otherwise have your evaluation stymied. Program and project (hereafter referred to as "project") activities may be disrupted in other ways too, presenting additional challenges and considerations depending on when the disruption occurs.

This document offers guidance to evaluators and staff whose projects have been disrupted, with a particular focus on Office of Special Education Programs (OSEP) grantees. It is organized into four sections:

1. Find Your Bearings
2. Assess the Available Options
3. Select the Best Options for Your Evaluation
4. Interpret and Report Results Accurately

We include a detailed discussion of the available options and creative solutions that might be available to you (and their associated limitations), as well as criteria you can use to select the options that are best for your evaluation. In addition, we highlight important considerations for interpreting and reporting evaluation data based on activities that have been corrected midcourse. Finally, we conclude with three hypothetical examples of disruption to OSEP-funded projects with different evaluation designs, offering suggestions and considerations depending on the timing and severity of the disruption and the groups of participants affected by the disruption.

FIND YOUR BEARINGS

In the face of a large-scale disruption, it can take some time and careful consideration to determine the implications for your evaluation or the project overall. This document focuses primarily on disruptions and adjustments to evaluation activities, but it is also important to consider possible changes to a project's implementation and/or participants since those changes might affect the evaluation. Several considerations and guiding questions are provided below.

HAS PROJECT IMPLEMENTATION CHANGED?

In considering how the disruption affected (or might affect) your evaluation, you first need to consider any changes to project activities. These changes are largely project-specific and relate to the stage at which the disruption occurred. In thinking through the ways in which these changes might affect your evaluation, we recommend considering several questions:

When and how did the disruption affect project implementation?

- Were any activities delayed?
- Did any activities shift from in-person to remote?
- Did any activities end earlier than planned, without the possibility of resumption?
- Did any project activities change in some other way that might affect your evaluation?
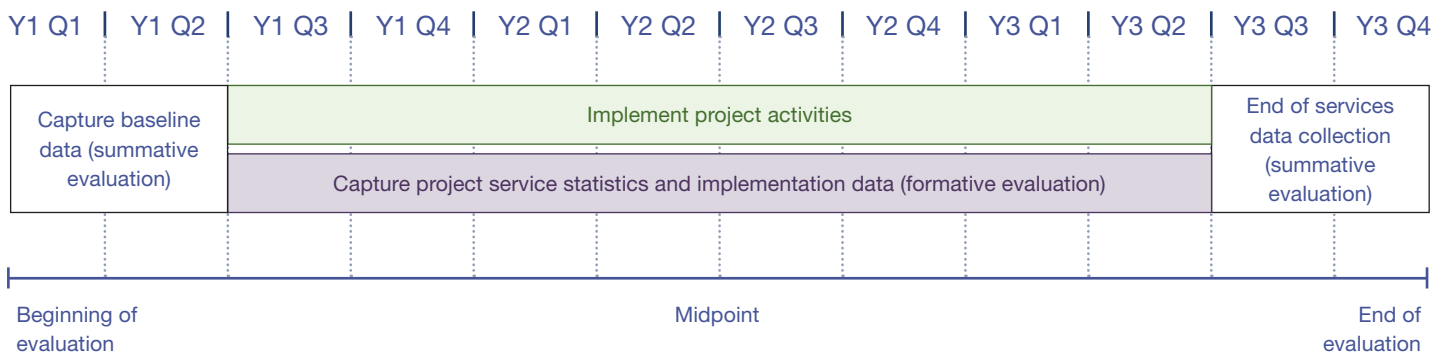
Do the project activities have distinct stages or phases?

- If so, were any stages fully implemented?
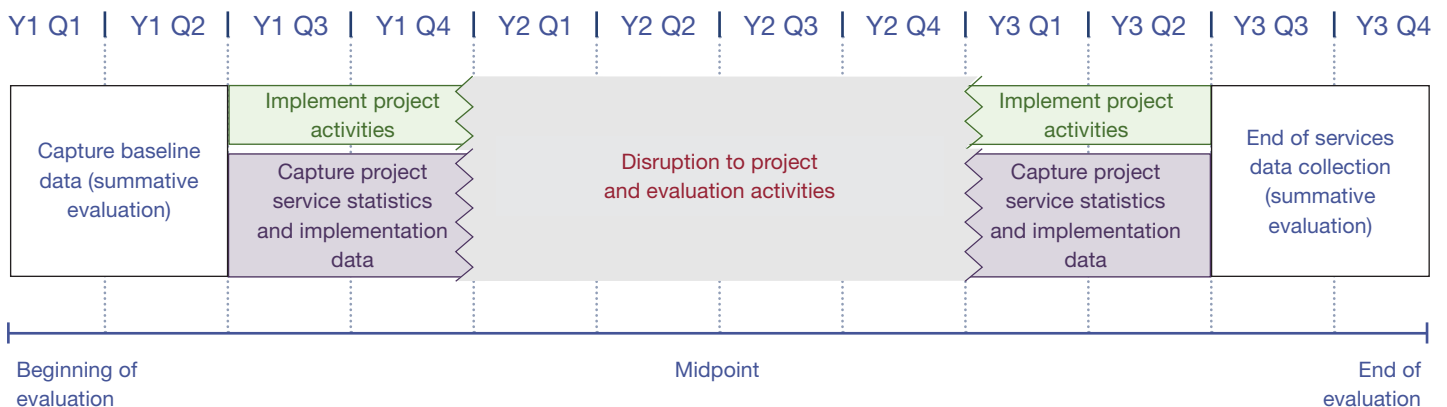- Or, were any stages not implemented at all?

Process maps or flow charts can help illustrate how the project experiences may have changed for participants and how evaluation activities were disrupted. See Exhibit 1 for a process map of an original evaluation design without a disruption, followed by a depiction of the same evaluation design with a disruption.

Exhibit 1. Comparison of original and disrupted design

Original Design, without Disruption

| Y1 Q1 | Y1 Q2 | Y1 Q3 | Y1 Q4 | Y2 Q1 | Y2 Q2 | Y2 Q3 | Y2 Q4 | Y3 Q1 | Y3 Q2 | Y3 Q3 | Y3 Q4 |

Capture baseline data (summative evaluation)

Implement project activities

Capture project service statistics and implementation data (formative evaluation)

End of services data collection (summative evaluation)

Beginning of evaluation — Midpoint — End of evaluation

Disrupted Design

| Y1 Q1 | Y1 Q2 | Y1 Q3 | Y1 Q4 | Y2 Q1 | Y2 Q2 | Y2 Q3 | Y2 Q4 | Y3 Q1 | Y3 Q2 | Y3 Q3 | Y3 Q4 |

Capture baseline data (summative evaluation)

Implement project activities

Capture project service statistics and implementation data

Disruption to project and evaluation activities

Implement project activities

Capture project service statistics and implementation data

End of services data collection (summative evaluation)

Beginning of evaluation — Midpoint — End of evaluation

HAVE YOUR PROJECT PARTICIPANTS BEEN AFFECTED?

When a large-scale disruption occurs, it is also important to think about the project participants' experiences. Here are some important questions to consider:

- How might the disruption affect participants' engagement with project or evaluation activities?
- How might the experience affect participants' responses to data collection activities, such as survey questions or interviews?
- How might trauma from the disruption affect participants' willingness (or ability) to participate in the project or evaluation?
- How might those who continue with their participation differ meaningfully from those who do not continue?
- Did the trauma affect all participants in the project or evaluation equally?

HAVE THE FUNDAMENTAL COMPONENTS OF THE EVALUATION CHANGED?

Next, equipped with answers to the questions above, you should revisit the project logic model, asking the following:

- Have the inputs, activities, or outputs changed?
- Given current realities, does the logic model still reflect the hypothesized relationships between inputs, activities, outputs, and outcomes?
- Are the outcomes, or the timelines for achieving the outcomes, reasonable given any changes to the project?
- Have circumstances changed enough to warrant changes to anticipated project outputs and outcomes?

The project logic model serves as a critical foundation for your evaluation plan. Therefore, if changes to the logic model are required, you will likely need to change the evaluation plan. Examine your evaluation questions to determine which are still applicable and whether you need to remove some and/or add others. In some cases, the goals of your evaluation may change; for example, you may need to pivot to examine the impact of the disruption on project activities and participants. As you consider changes to the evaluation plan, it is important to think about the resources available for the evaluation. Prioritize allocating resources to answer those questions that are most critical, if necessary. Talk with the project team, other members of the evaluation team, and, possibly, key stakeholders to gather input on the revisions you are considering. Make sure to distribute the revised logic model and evaluation plan to these groups. If you need to make changes, talk with the project director to ensure that they obtain approval for the changes from the OSEP Project Officer or other funder, as applicable.

For More Information

For more information on logic models and how to connect them to evaluation plans, see Linking Expectations to Evaluations: Using Your Logic Model to Create Your Evaluation Plan.

ASSESS THE AVAILABLE OPTIONS

You can use different strategies to address the disruption to your evaluation, depending on your specific situation. In this section, we offer the following possible strategies to consider in the face of disruption:

- Maintain your original evaluation design, with limited changes to the evaluation
- Maintain your original evaluation design but collect your data differently
- Maintain your original evaluation design but analyze your data differently
- Change your evaluation design

In the next section, we provide some considerations to help you select the best solutions for your particular context.

CAN YOU MAINTAIN YOUR ORIGINAL EVALUATION DESIGN WITH LIMITED CHANGES?

In some cases, when faced with a disruption, you may still be able to maintain your original evaluation design, with some modifications. One common modification would be using a smaller number of participants (e.g., families, students, related services providers, teachers, classes, schools, districts) in the evaluation than you originally planned. For example, if a subgroup of your original participants experienced project activities as intended and is able to continue with the evaluation as planned, you may be able to use data collected from this subgroup of the original population for your analyses and put aside data for those who did not. Potential limitations of this approach are that the subgroup that remains in the evaluation may no longer be representative of the population of interest or that any statistical analyses you conduct may become underpowered (see sidebar). Still, collecting and analyzing some data is often preferable to having no data at all.

Another modification you can make when keeping your original design is to evaluate shorter periods of implementation than you had originally planned. For example, if teacher training was planned for a full school year—with pre- and posttest data collection scheduled at the beginning and end of the school year—you might be able to collect your posttest data shortly after the training ended due to the disruption, resulting in a 4-month treatment period, rather than one that spanned 9 months.

Regardless of your design, keep an eye out for attrition! One consideration that has important implications for your evaluation—no matter your design—is attrition of evaluation participants. Attrition can result from changes to project implementation or from participants not being available during the data collection period, lacking access to or understanding of remote modes of data collection, or simply being unwilling to participate and withdrawing from the evaluation. If you make a midcourse correction and move from in-person to remote data collection, you may also experience attrition if participants do not have a computer, internet access, or a webcam.

> **Quick Summary: Analytic Power**
>
> Large decreases in sample sizes will reduce the analytic power of statistical analyses—the ability to detect a statistically significant effect if one exists. If the evaluation is trying to examine a project's or intervention's effects, a power analysis can help you determine if you have a large enough sample size remaining after the disruption to have a reasonable chance for detecting effects.

When attrition is high, or when attrition rates are different for different subgroups of the population, evaluation findings may not represent the population being evaluated (i.e., your data will not be representative). If this occurs, consider conducting an attrition analysis to understand any potential bias in the results so you can accurately report on the limitations associated with your findings. You can conduct an attrition analysis by comparing the characteristics of those who participated with those who were lost to attrition (for example, did those who stayed in the evaluation receive services not received by those who dropped out?). If you find that individuals who participated in the evaluation differed in meaningful ways from those who did not, your findings may not be representative of, or generalizable to, the population of interest. If you have a comparison group as part of your evaluation (e.g., your design contains a treatment and a comparison group), you can also compare attrition rates for each group. If there is a significant difference in attrition between the two groups, you may have an issue with attrition bias. Document the results of any attrition analyses you conduct so you can inform the readers of your evaluation report how attrition might have affected your results.

CAN YOU MAINTAIN YOUR ORIGINAL DESIGN BUT COLLECT DATA DIFFERENTLY?

Altering data collection methods, timing, or locations may allow you to continue collecting valuable data for the evaluation when faced with a disruption. Some options—such as moving from in-person to virtual interviews—may require minimal changes to the original evaluation plan, while other options may have implications for the evaluation design, analysis, and interpretation of findings. Disruptions to project or evaluation activities may also provide opportunities for new data collection activities that would not have been possible originally.

Remote data collection

One strategy to keep your evaluation on track is to shift to remote data collection. You might conduct data collection activities, originally planned to occur in person, through telephone, videoconference, or online methods. For instance, you might be able to conduct interviews and focus groups through a videoconference or complete individual assessments via videoconference or web-based administration. Exhibit 2 presents ideas for shifting from in-person to remote administration for several common data collection methods along with key questions and considerations.

Important Consideration: Safety Protocols

Keep in mind that in some cases, such as with the COVID-19 pandemic, you may need to implement new safety protocols for data collection. For example, during a pandemic, if in-person data collection is required and will be ongoing, you would need to obtain and use appropriate personal protective equipment for participants and data collectors.

Exhibit 2. Key questions and considerations for remote data collection options

In-Person Data Collection Methods	Remote Data Collection Options	Key Questions and Considerations
One-on-One Interviews	– Telephone – Videoconference	**Key questions:** Can you conduct interviews over the telephone or videoconference? **Considerations:** Confidentiality may be an issue if respondents do not have access to a private location. Consider the nature of the interviews and whether they need to be conducted in private.
Focus Groups	– Conference call – Videoconference	**Key questions:** Can focus group participants access a computer with reliable internet and a video camera? If not, is a conference call an option? **Considerations:** This method may exclude participants without reliable internet or telephone service.
Individual Assessment	– Videoconference – Online	**Key questions:** For assessments that require interaction between the assessor and the person being assessed, can you use a videoconference? For assessments that can be completed independently and do not require proctoring, is online administration an option? **Considerations:** Feasibility may depend on the participant's age (it may be difficult for young children to focus), the type of assessment (e.g., an assessment requiring the student to demonstrate a complex skill may not work), and ability to fully view participants. Even written assessments may need to be administered by videoconference to allow for test instructions and proctoring.
Surveys	– Online – Mail-in	**Key questions:** For surveys that were originally planned to be administered in person, can participants complete them independently through an online platform or via mailed paper surveys? **Considerations:** Internet or smartphone access is required for participation in an online survey. While mailed paper surveys with return envelopes and postage paid may be appropriate for populations with limited technology access, there are issues associated with addresses for paper surveys. For example, in the case of large-scale disruption the respondents or the evaluation team might be displaced from their regular mailing addresses.
Observations	– Video observation	**Key questions:** Can participants' behavior be fully observed via video observation? **Considerations:** Observations of virtual classrooms depend on participants' willingness and ability to turn on their camera. Video observations of in-person classrooms or events may not capture as much information as in-person observations, as cameras might not reach all corners of the room, and visual barriers and/or insufficient resolution due to internet bandwidth or other issues may make it difficult to see all participants in the room.

Technological and logistical considerations

In assessing your technology options, consider factors such as evaluation participants' and data collectors' access to and familiarity with required technology. Also consider logistical issues like consent processes, communication with participants, and revisions to evaluation protocols.

Technology. A consideration for all remote options is whether data collectors and evaluation participants have access to dependable technology, including a computer or smartphone, internet, and videoconference software. In considering remote methods, think about the needs and skills of data collectors and evaluation participants. Are they likely to have access to the required technology and have the skills necessary to use it? If not, can your project provide the hardware, software, or internet access needed, as well as technical training and support? In locations where internet access is limited, are there facilities such as schools or libraries where individuals can

access internet services, if needed, to participate in the evaluation? If not, is data collection using smartphones possible? Will you need to hire additional staff or staff with different skills to conduct remote data collection?

Training and guidance documents. Regardless of your data collectors' experience with remote data collection methods, they will need guidance and training on your evaluation's data collection methods and procedures. In your revised evaluation plan, include details about procedural changes to data collections and plans to offer training and support to data collectors and evaluation participants, as needed. This might include revised data collection protocols, revised process guidelines, and training and support on data collection methods, technology, and online norms and procedures. Training and guidance will be critical if data collection methods shift. Consider how you will do the following:

- Establish inter-rater reliability with remote data collectors
- Ensure the accurate transfer of data into your data management system
- Complete data quality checks that now include the effects of remote data collection

Logistics. Changes in data collection may have other logistical implications. Will you need to change the consent process or how you communicate with evaluation participants? Evaluators often rely on on-site staff to distribute information about evaluation activities and obtain participant consent. For example, a teacher might be responsible for sending printed evaluation information home with students and collecting consent forms. If this is no longer an option, work with on-site staff to determine how, when, and by whom information will be provided to participants through remote mechanisms such as email or a website. You may also need to provide updated information and training to on-site staff who are involved in the evaluation.

Validity and reliability considerations

If you plan to collect data remotely, you will need to consider data collection approaches as well as the validity and reliability of the instruments you use. Some important questions to consider include the following:

- Will the remote format change how participants interact with and respond to the instruments?
- Will instruments that were developed to produce valid and reliable information when used in person provide the same valid and reliable information in a virtual setting?

For validated instruments, you may be able to contact the instrument developer for guidance. If this isn't an option, you can use the steps below to gather data on the validity and reliability of your instruments. Ideally, you should do this prior to selecting and administering the instrument; however, if this is not possible, you can also do it after data collection to identify the limitations of your evaluation and provide guidance on interpretation of findings.

Conduct a pilot test. If possible, test remote administration of the instrument with individuals or sites, preferably those that are not involved in the evaluation. A pilot test serves as a valuable practice run that will give you important information about how feasible remote administration actually is. Be specific about the type of feedback you would like to gather, including how long it takes participants to complete instruments, ease of use/functionality of the technology, and any issues or problems participants encounter.

Assess inter-rater reliability. Using the new remote data collection methods, determine whether different observers or scorers are consistent in their ratings, if applicable. As with original testing procedures, continue training and retesting until you achieve sufficiently high reliability (e.g., 80% or higher). If inter-rater reliability of remote administration remains lower than that of in-person administration, data quality will be lower, and it also provides evidence that the remote format is not fully comparable with the in-person format.

Determine whether deleting measures will affect validity. If due to the disruption, part of the intervention or project was discontinued, outcome measures related to that component may no longer be relevant for the evaluation. If those measures are part of a validated instrument, however, removing them may affect the instrument's overall validity. This is particularly important if the overall instrument has other outcome measures that are still relevant to your evaluation. In such cases, consider administering those less relevant measures anyway and provide instructions on what to do with that data. For example, a classroom observation may have measures of classroom activities that cannot occur in virtual instruction (e.g., teachers walking by students during instruction), along with other measures of activities that should be taking place. To preserve an instrument's validity, the evaluators may instruct data collectors to complete the entire instrument but note in the analysis and reporting that a particular activity or behavior could not be implemented because of the disruption.

> **Helpful Hint**
>
> If you need to alter validated data collection instruments due to changes in project activities, consult with the instrument developer to ensure that modifications will not weaken the instruments' validity.

Compare data collected through in-person versus remote methods. After collecting some or all data through remote methods, examine the data to look for comparability with data collected in person using the same instrument. This often takes the form of psychometric analysis, but you can also get an idea of differences in the data by running descriptive analyses or looking at differences in qualitative responses. You may need to access published research for data to inform psychometric analyses if you cannot collect any data in person due to the disruption. Conducting psychometric analyses on an instrument employed during in-person versus remote administration is a good way to determine whether there are systematic differences in reliability and validity for the different data collection formats. For example, for instruments with multiple items that measure a single construct such as a survey, compare internal consistency of items for data collected during in-person versus remote administration.

It is important to clearly document your remote data collection procedures, the extent to which administration of outcome instruments differed from those administered in person, and any data quality concerns—addressing whether or not you think data from the remote administration can be comparable to data collected as originally planned. Even if you cannot fully answer that question, the OSEP Project Officer/funder, stakeholders, and other audiences for your evaluation need to understand the potential limitations of your evaluation results.

Delayed data collection

In some situations, a delay in data collection may be a viable option for dealing with disruption to project or evaluation activities. Depending on the nature and timing of the disruption, you may be able to delay collection of baseline, implementation, and/or followup data. For example, you might be able to collect followup data 8 months after the project ends instead of 6 months, as originally planned. Factors to consider include the following:

- OSEP/funder reporting requirements
- Length of the delay
- Whether project implementation is continuing during the disruption
- Availability of data collectors to conduct delayed data collection
- Willingness of sites or individuals to participate in data collection at a different time
- Extent of potential attrition of evaluation participants due to the delay
- Whether timing adjustments are possible with external assessments, for example, if you are using state achievement test scores as an outcome measure.

Delayed collection of baseline data

If you are collecting baseline data for your evaluation—and if any interruption to project activities occurs early in the project—delaying baseline data collection may be a good option, provided the project can proceed as planned after the disruption. If the disruption occurs after baseline data was collected for some sites but not for others, you may still be able to group sites into cohorts for analysis and reporting. However, if participants had substantial exposure to project activities before the rescheduled baseline data collection, delayed baseline data may not represent the true baseline conditions. This may affect statistical estimates of project effects (e.g., if participants have already achieved growth in outcome measures, statistical effect sizes may be smaller than if the baseline data collection occurred prior to exposure to project activities). If possible, compare baseline scores and key participant characteristics of individuals in the delayed and non-delayed baseline assessment cohorts to determine if the data collected truly represent a "baseline."

Delayed collection of implementation data

If project activities stopped due to the disruption, you may be able to simply pause your evaluation and resume data collection when implementation restarts. However, if you must delay data collection while project implementation is ongoing, you could use interviews or other sources of data to assess the extent and nature of implementation during the delay. This might include determining if all project activities were implemented and if all sites or individuals participated as expected. This data will help you understand project implementation during the disruption and provide context for findings once you move forward with your evaluation. Relatedly, you should try to determine whether project implementation changed in any way during the delay in evaluation activities. Even if the project's implementation did not *officially* change, the disruption might have unofficially changed the way activities were implemented. If possible, collect additional information on implementation to help you and others understand whether deviations from desired implementation were related to the disruption and what implications any changes might have for the project and the evaluation findings.

Delayed collection of followup data

When the disruption interrupts project activities and/or the evaluation toward the end of the project period, you may be able to collect followup or posttest data later, when conditions improve. Keep in mind that the meaning of this delayed followup data may change. For example, delays in followup data collection can potentially weaken the ability of the evaluation to detect treatment effects, as can changes to the overall project or context during the delay. In addition, it may sometimes be hard to know how participants may have changed with respect to key variables (e.g., motivation, self-regulation), even if project activities were not implemented during the delay. Consider the example of a district's school-wide positive behavioral supports initiative being put on hold during the disruption. If discipline referral data is one of your outcome measures, when school resumes, you may see changes in this data that are not attributable to project activities but rather to experiences students had while project activities were on hold. Therefore, when interpreting findings, you should carefully consider whether the delayed outcome data truly reflects the effects of the treatment. It is also important to note that if the project period has ended by the time followup data collection is feasible, your ability to use this option will also depend on the availability of resources to support evaluation activities after the project is over.

Data collection at a different site

Depending on the project's implementation stage, it may be possible to recruit replacement sites for the evaluation if the sites you had originally selected are no longer able to participate. This is primarily an option for regional or local disruptions, such as a fire or a hurricane, not large-scale disruptions such as the COVID-19 pandemic. Choose sites with similar characteristics to the original sites, and use the same selection criteria used to select original sites (e.g., characteristics of the population, readiness for the project activities/intervention, leadership commitment). This option may be best suited for replacing comparison or control group sites or intervention group sites at early stages of project implementation, where evaluation investments have been minimal. For disruptions at later stages of implementation, starting over with new sites may be cost- and time-prohibitive. In all cases, assess the implications for project timelines and the costs for delays. Think about the time and resources needed—for initial communications with sites, obtaining permissions, following Institutional Review Board (IRB) processes, informing and obtaining consent from participants, training site personnel, etc.—and determine if the delays and additional costs associated with moving to new sites fit within the project and evaluation resources. Also, remember to ensure that the project team clears these changes with the OSEP Project Officer/funder.

Helpful Hint

Collecting followup data well after the conclusion of an intervention or activity may provide less robust information on treatment effects, but it can still provide important information on the extent to which intervention components or project activities were implemented as intended and whether and to what extent efforts are being sustained.

Important Consideration: Evaluation Budget

If you decide to collect data at different sites, you may need to reprint materials, conduct screening interviews, provide tangible resources or materials, or provide training or guidance to new participants, all of which have an associated cost. Also, if resources were provided (such as project materials for school use), will you attempt to retrieve them? And if so, how?

Collecting New Data during the COVID-19 Pandemic

During the COVID-19 pandemic, some funders asked evaluators to pivot to answer new questions aimed at understanding how educators, students, and families were coping during the pandemic. Evaluators added new data collection activities, such as virtual focus groups, to collect data to inform policymakers and funders on communities' needs and their implications for future instruction or services.

Collection of new types of data

Pay attention to possible new evaluation questions emerging from the disruption. For example, you may want to better understand changes in the educational context in which the project is being implemented, so you might add data collection activities, such as focus groups or interviews. Or, if project implementation is altered, you may want to understand the changes that occurred due to the disruption, which would require additional implementation measures, such as a modified implementation checklist or fidelity instrument. In addition, you may want to consider collecting more qualitative data to better understand participants' experiences of the disruption and gain insight into how, if at all, those experiences might be affecting project outcomes.

New remote data collection methods may present different data collection opportunities from those that you had previously planned. For example, you could supplement your original data collection activities with virtual interviews or focus groups with project implementers or participants. Of course, the ability to adjust measures or launch new data collection efforts will depend on guidance from your funder, timing, and evaluation resources.

CAN YOU MAINTAIN YOUR ORIGINAL EVALUATION DESIGN BUT ANALYZE YOUR DATA DIFFERENTLY?

Sometimes, a disruption hits the project or evaluation so late that nothing can be done about the evaluation design or the data collection activities. What are your options in those cases? While it might be tempting to try to adjust the way you analyze your data, we strongly recommend caution. When considering updates to your data analysis plan because of a disruption to your evaluation, it is important to think critically about the rationale behind these changes. In other circumstances, when data is missing, for example, you might use statistical adjustments such as multiple data imputation, regression imputation, maximum likelihood, or others. However, statistical adjustments such as these will not usually be appropriate in the context of a disruption because the *disrupted* data is unlikely to be missing at random. In many situations associated with disruption, the missing data might reasonably be expected to have higher or lower values than the non-missing data. For example, as has been seen in the COVID-19 pandemic, unequal access to computers means that more affluent students, or students located in places with good internet connection, had more opportunities to engage in online learning than students who did not have access to computers or the internet.[1] So, analysis of data collected during remote instruction would have to take this unequal access to technology into account when dealing with missing data.

---

[1] Van Lancker, W., and Parolin, Z. (2020). COVID-19, school closures, and child poverty: A social crisis in the making. *The Lancet: Public Health*, 5(5), E243- E244.

The following considerations can help guide your examination of potential changes to your analysis plan.

1. First, using your logic model, make sure you have a solid understanding of the types of results you would have expected if the disruption had not occurred, or if the disruption had not substantially affected project implementation. For example, if the project had three trainings designed to result in increases in teachers' knowledge of high-quality math instruction for students with disabilities, you would likely anticipate that the teachers who attended all three trainings would show greater knowledge increases than teachers who attended just one.

2. Next, consider the ultimate beneficiaries of your project (e.g., students with disabilities) and how similar or different (i.e., heterogeneous) you would have expected their outcomes to be if the disruption had not occurred. For example, if you are looking at math outcomes for students with disabilities, how variable is this data during typical times?

3. From here, consider the impact of the disruption on project activities. How consistently were activities implemented, and do you have fidelity of implementation data? If you have data showing that project activities were implemented with fidelity (e.g., teachers were able to attend all three trainings and demonstrated increases in their knowledge of high-quality math instruction for students with disabilities), you can be more confident in your analysis of differences in ultimate outcomes (e.g., math outcomes among the students with disabilities served by those teachers). If, however, you do not have fidelity data, or if that data shows that implementation varied substantially across sites or contexts, then you need to be careful in your interpretations of any differences in outcomes.

4. Also think about the experiences of the ultimate beneficiaries of your project (e.g., students with disabilities) during the disruption. For example, if you were planning to examine student academic outcomes during the COVID-19 pandemic, differential learning loss due to missed or ineffective instructional time is something that you would need to consider. Could you use findings on summer learning loss to estimate the effects of your disruption (see callout box)? If you want to incorporate information about participant experiences into your analysis, you will need to think creatively about how to measure or account for the changes your participants experienced, given the specific context of your unique disruption.

> **Innovative Idea: Measuring Learning Loss**
>
> Can you use existing data and established findings related to student summer learning loss during a particular time period (e.g., the 2019-20 school year) to estimate the effects of student learning loss due to your disruption?

CAN YOU CHANGE YOUR EVALUATION DESIGN?

Depending on how the disruption affects the project and its evaluation, you may opt to change your evaluation design to accommodate a new reality. Be sure to make a design change that can use existing data or employ data that you can still feasibly collect. You may be able to use a new design to complement data that you already collected with your original design (while considering how you would analyze and interpret the different types of data). Exhibit 3 presents some hypothetical examples of changes in evaluation design to spur your thinking and creativity.

Exhibit 3. *Hypothetical examples of changes in evaluation design*

Instead of your original design...	... you might consider using this new design
Non-equivalent pre-post control group design	Interrupted time series
Quasi-experimental design	A series of case studies
Simultaneous implementation across all participants/sites	Staggered cohort design

When deciding whether to change your evaluation design, two primary considerations are the length of the project (and the evaluation) and the timing of the disruption. Disruptions that occur later in the project and evaluation timeline make changing your evaluation design much more difficult, especially if you've already completed any major evaluation activities. And any changes to an evaluation design—even one that was in the beginning phases of implementation—will likely mean a loss of baseline data. In this case, you might consider collecting new baseline data or collecting administrative or extant data from an earlier time point to use as your baseline (see the "Select the Best Options for Your Evaluation" section for options).

INTEGRATION OF EXTANT DATA

Extant data (also sometimes referred to as administrative data) is often publically available and, across datasets, includes a wide selection of outcome data including national-, state-, and district-level student achievement, educational settings, teacher qualifications, and demographic data. Using these types of data that are already available can reduce logistical issues for your evaluation and minimize the burden to your potential evaluation participants during what may be a very difficult time for them. For example, in the case of a local disruption, you could possibly substitute statewide outcome data for previously planned primary data collection of student outcomes in the district.

There are several important factors to consider related to the use of extant data. First, it can be difficult to achieve a perfect match between your project's anticipated outcomes and the outcome data available in extant datasets. In addition, you do not have any control over the format of the data; and sometimes the way the data is organized limits the types of analyses you are able to conduct. Finally, if the disruption affecting your project is widespread enough, it may also disrupt other extant data collections that would normally occur (for example, COVID-19 caused statewide assessments to be unavailable at the end of the 2019-20 school year).

New natural experiments

As you are considering changes to your evaluation design, reflect on whether the disruption gives you an opportunity to answer more questions than you had originally anticipated. Sometimes, natural

Potential Sources of Extant Data

There are many sources of extant data, some of which you may discover through the course of your work. Several sources that OSEP-funded project evaluators may find relevant include the following:

- Data collected under the Individuals with Disabilities Education Act (IDEA)
- Common Core of Data (CCD)
- National Assessment of Educational Progress (NAEP) data
- State-level early intervention and special education data
- District-level early intervention and special education data

experiments—those that arise naturally when participants are randomly exposed to different conditions due to acts of nature or other factors in the environment—result from disruptions. For example, if flooding closes some of the schools in a particular area, a natural experiment might result if teachers from those schools attend fewer professional development sessions related to a certain reading intervention than teachers whose schools did not close. Then later in the year, the evaluation team might be able to assess whether students whose teachers received less professional development performed similarly on reading outcome measures to the students whose teachers received more professional development. While the evaluation team might not have explored this question originally, the data from the natural experiment might reveal that the project team can reduce the number of professional development sessions required for teachers in future years. Exploring options for natural experiments arising from disruptions could be relevant in situations where implementation is interrupted or delayed for some participants or locations, but not others, or in situations where implementation is of varying lengths of time in different locations or for different participants.

SELECT THE BEST OPTIONS FOR YOUR EVALUATION

Obtaining high-quality and meaningful data for project evaluations can be challenging under the best of circumstances. When unexpected events disrupt the evaluation, it can be daunting to select the best approach for moving forward. Below, we outline a structure you can use to think through the available options, including critical issues to consider and key questions to ask as you proceed through the decision-making process.

HOW CAN YOU DECIDE WHICH OPTIONS TO CHOOSE?

As mentioned earlier, when faced with a significant disruption, start by revisiting the evaluation's goals and logic model, and then take a systematic approach to assessing the feasibility of the different options available to you. This will help you to clearly identify how you will address any challenges with implementation of your evaluation plan. Engage a variety of stakeholders—the OSEP Project Officer/funder, project staff, research methodologists and statisticians, project participants, and families—in the process. They can help you identify priorities for the evaluation going forward, pinpoint additional sources of data, and consider the benefits and drawbacks of options. Additionally, involving them can build understanding of and commitment to any changes in plans.

Review the evaluation's goals and objectives. The primary goals of most evaluations are to provide key stakeholders with formative and summative data about a project's or intervention's progress and outcomes. As you consider what might need to change in your evaluation due to the disruption, think about the project's most critical information (data) needs. Have the information needs changed?

- Given how the disruption has affected the project and the evaluation, what information will best help key stakeholders understand project outcomes?
- What information do funders/OSEP Project Officers want and/or need? Is that different from what it was prior to the disruption?
- What information is needed to meet reporting requirements?
- Do you have flexibility in determining the timing and content of evaluation reports?

Use your assessment of what has happened due to the disruption (see the "Find Your Bearings" section earlier) to reassess and document the evaluation's (new) goals and priorities.

Examine the sources of data that are (still) available. What data have you already collected for the evaluation? Have some data sources become unavailable to you? What are potential new sources of data? Are there sources of relevant extant data or new possibilities for data collection you can embed into other activities? Work with the project and broader evaluation teams to brainstorm potential data sources.

Assess the feasibility of options for collecting the needed data. For all the potential options for dealing with the disruption, determine whether they are truly feasible for your evaluation, given the other responsibilities of the evaluation team, site staff, and evaluation participants; access to participants and/or sites; and the remaining evaluation resources, including the money and time available. Consider implications for the budget, timeline, and deliverables. Then lay out possible solutions to each challenge, taking into consideration the budgetary and time constraints, as well as any existing requirements from the project or the funder.

Specify criteria for selecting the best options. Using the revised goals of your evaluation plan, update your logic model and determine the amount of resources and time that you have available as these will be important selection criteria. Additional criteria will depend on the original evaluation design and feedback from your OSEP Project Officer/funder. For example, in a major disruption in which site personnel, evaluation participants, and families may be experiencing trauma, funders and other stakeholders may be focused on minimizing burden on participants.

Develop a revised timeline and budget. Before making final decisions, determine the budget and timeline for the final options to ensure that they are acceptable to your OSEP Project Officer/funder and other stakeholders. Some considerations might include the following:

- Staff costs for redesigning some element of the evaluation as well as cost related to IRB updates and approvals, as required;

- Costs associated with transferring to technology-based or remote data collection (Will you need to purchase hardware, software, or internet service for evaluation participants or data collectors? If not, will remote data collections be possible?); and

- Costs of pilot-testing or validating instruments and using altered data collection approaches.

After evaluating feasibility, costs, timeline, benefits, and challenges, determine if you have options that will provide meaningful information to key stakeholders. If not, work with your OSEP Project Officer/funder to investigate whether the evaluation resources can be repurposed to answer new evaluation questions such as what sites and participants need during the disruption.

INTERPRET AND REPORT RESULTS ACCURATELY

When it comes time to report data for an evaluation that was affected by a disruption, be as transparent as possible in the report about what aspects of the evaluation and the project were affected by the disruption, how the evaluation and project teams resolved any problems, and the potential limitations of those solutions. This provides important context for readers and helps ensure that they have a full understanding of what happened and how the results can be interpreted and used.

WHAT WAS DISRUPTED?

Describe the type of disruption and its effects on the evaluation. Consider effects on the design, data collection, analysis, and interpretation of results. Include the following in your description.

The status of the evaluation prior to the disruption. Had evaluation data collection started? Was baseline data collected? Was partial data collected? Were you toward the end of the project with multiple years of evaluation data already collected?

Timing and duration of the disruption. Disruptions vary in duration, from quite short (no heat during a particularly cold week) to extended periods of time (such as when a hurricane causes extensive flooding damage to school buildings). Some disruptions, such as a hurricane or wildfire, are abrupt, and it is easy to pinpoint when the disruption occurred or began and to estimate when it might end. For other disruptions, such as a pandemic, it is harder to estimate the beginning in a particular area and to determine when it might end. Try to identify the date the disruption began and how long it lasted as best as you can, and assess the effects it had based on the project and evaluation timelines.

Effects of the disruption on implementation of project activities and the evaluation. In what ways did project implementation change? Did project activities pause, completely stop, or were they reconfigured in some way? How many components of the project changed? What made it difficult to proceed with the evaluation as originally planned? Which aspects of the evaluation needed to change? Did any aspects of the evaluation stay the same?

HOW DID YOU ADDRESS THE DISRUPTION?

In the evaluation report, give details about the solutions you implemented to address the disruption to your evaluation. Describe these solutions and specify what changed relative to what you originally planned. Consider and report on changes you made in each area of your evaluation, including the following:

- Design
- Participants
- Data collection methods
- Instruments
- Timelines
- Analysis

WHAT ARE THE LIMITATIONS OF THE RESULTS?

In reporting results for an evaluation affected by a disruption, it is especially important to specify any and all possible limitations. You may want to revisit the section "Assess the Available Options" to review the limitations presented for each possible solution, and then consider which limitations apply to the specific solutions you implemented. In general, it is important to reflect on whether the changes made to the evaluation affect the confidence you can place in the evaluation findings. Some key questions to consider include the following:

- Is the data collected post-disruption comparable to data collected previously?
- Were the project activities pre- and post-disruption sufficiently similar to make comparisons?
- Were treatment and comparison or control groups similarly affected by the disruption?
- Were sites similarly affected?
- How did the disruption itself affect outcomes?

See Exhibit 4 for a checklist of reporting considerations for potential limitations.

Exhibit 4. *Checklist of potential limitations to document in evaluation reports*

Characteristics of the disruption	Addressed in Reporting/Interpretation
Status of the evaluation prior to the disruption	☐
Timing of the disruption	☐
Duration of the disruption	☐
Effects of the disruption on implementation of project activities and the evaluation	☐
What changes were made to implementation activities?	☐
What made the evaluation unfeasible to complete as planned?	☐
Which aspects of the evaluation needed to change?	☐
Did any aspects of the evaluation stay the same?	☐
Changes to the evaluation	**Addressed in Reporting/Interpretation**
Design	☐
Participants	☐
Data collection methods	☐
Instruments	☐
Timelines	☐
Analysis	☐
Other (specify: _____)	☐
Limitations of the results	**Addressed in Reporting/Interpretation**
Design-specific limitations	☐
Comparability of data collected before and after the disruption	☐
Varying impacts of the disruption across sites	☐
Confounding factors (e.g., trauma)	☐
Representativeness of participants	☐

A NOTE ABOUT TRAUMA

In addition to the considerations presented above, it is important to keep in mind that during the disruption, project or evaluation participants may have experienced a trauma that may change the outcomes of your evaluation, regardless of any disruption to project activities. If the disrupting event was traumatic, it may lead to a loss of skills, an increased stress response, and other negative effects that might affect participant performance on outcome measures or even interfere with potential positive effects of a project or intervention.[1] For example, a child who has lost his or her house in a hurricane may perform worse on an assessment than would have been the case otherwise, regardless of the quality of the received related services.

Moreover, major events can inhibit people's recall. Positive events tend to expand people's window for remembering specific details around the event, while negative events tend to shorten it.[2] Consequently, people experiencing trauma may have difficulty remembering activities undertaken before the event or find it hard to accurately remember their feelings, knowledge, or practices or project or system policies and procedures prior to the event.

Finally, the disruption may alter priorities for participants—a participant's pre-disruption goals or planned participation in project activities may no longer be the same. For example, if a participant, pre-disruption, was hoping to make knowledge or skills gains as a result of participation and the disruption caused a change such as death in the family or food insecurity, the participant's motivation and immediate needs may be altered.

The COVID-19 Pandemic and Recall

Events like the COVID-19 pandemic, which take place over an extended period, can be harder to pinpoint, report on, and understand than events that occur at a particular point in time, such as an earthquake or the terrorist attack on September 11, 2001. A prolonged disruption like a pandemic may lead to changes in policy and procedures that may or may not remain in place after life returns to normal. Since people use salient events as anchors, events such as COVID-19 will likely make it difficult (or perhaps unwise) for evaluators to use post-pandemic surveys to ask recall questions about typical pre-pandemic practices.

A QUESTION OF REPRESENTATIVENESS

In your consideration of potential limitations, it is important to ask whether the participants in your evaluation are still representative of the target population. Consider whether changes in the evaluation methods may have affected your ability to recruit a non-biased and representative group of participants. Were certain people or populations left out of the data collection because of the changes you made? For example, if you switch to virtual collection of student outcome data, but you are unable to collect data from those without reliable access to the internet, students who are in lower income environments—or those who live in rural areas with limited internet access—may be disproportionately underrepresented in your evaluation sample. As such, the evaluation sample may no longer reflect the total population of students in the target population. Determine any potential limitations related to representativeness in your evaluation. Specifically, note how those participating in your evaluation during or immediately following the disruption varied from the target population and whether conclusions can be drawn about the target population given the differences identified.

---

[1] Substance Abuse and Mental Health Services Administration (SAMHSA) (n.d.). *Understanding child trauma*. Available at:
     https://www.samhsa.gov/child-trauma/understanding-child-trauma.

[2] Kensinger, E.A. (2009). Remembering the details: Effects of emotion. *Emotion Review*, 1, 99-113.

HYPOTHETICAL EXAMPLES

Below, we present examples and solutions for three evaluation designs that are common for OSEP projects: one-group pre-post design (Exhibit 6), interrupted time series design (Exhibit 7), and non-equivalent pre-post control group design (Exhibit 8). In each example, we consider when the disruption occurred, how severe the disruption was, and who the disruption affected (e.g., all participants or a subgroup of participants). Severity can be defined as duration, magnitude of disruption to operations, and participant availability. However, every project will need to assess severity independently, as there is no uniform standard. The tables that follow the descriptions of the examples provide possible solutions that the evaluation might use, depending on the timing and severity of the disruption. When considering solutions such as these, be sure that the project team consults with the OSEP Project Officer/funder before making any substantial changes to the evaluation or project implementation plans!

For More Information

For more information about these designs, see the following resources available on the OSEP IDEAs That Work Website:

- Compared to What? Using One-Group Pre-Post Designs
- Compared to What? Using Single-Case Designs
- Compared to What? Using Nonequivalent Pre-Post Control-Group Designs

EXAMPLE 1: ONE-GROUP PRE-POST DESIGN

Original design: A Parent Training and Information (PTI) Center worked throughout a state to increase its outreach to parents and families living in rural areas. After 6 months of intensive outreach, the PTI's evaluation team analyzed the demographic data of families who received technical assistance from the center and compared it to the demographic data obtained in the previous 6 months to see if they were actually providing services to an increased number of rural families.

Disruption effect: In this example, the data analyzed (i.e., project statistics on demographic characteristics of the families served) is directly related to the PTI's ability to provide services. The treatment period is relatively modest: 6 months. Given this information, two key considerations rise to the surface:

- Because the treatment period is relatively short, a disruption of even a few weeks can affect the PTI's ability to provide the needed services to rural families, as well as the families' ability to engage in those services.
- Generally speaking, making comparisons to prior time periods may not be appropriate, given the disruption. Thus, the evaluation team will have to make a judgement call as to whether comparisons of data over time are possible.

Exhibits 5 and 6 illustrate how the disruption affects the project implementation and conclusions that can be drawn from the evaluation, and outline possible ways to address disruption to this design.

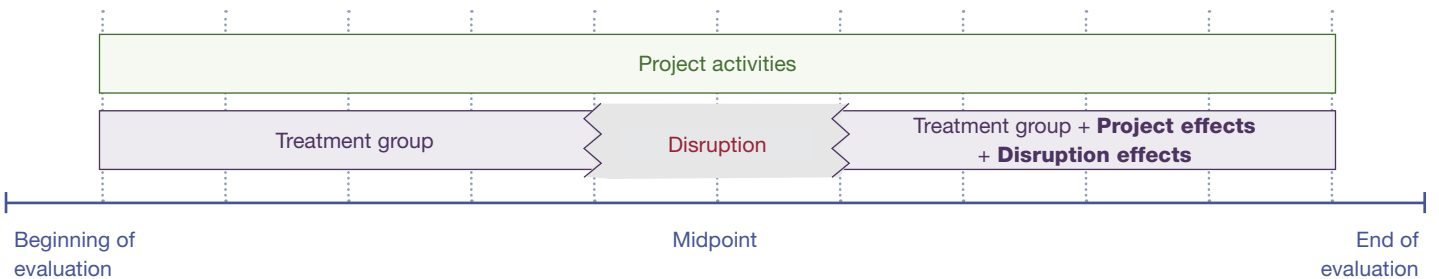*Exhibit 5.* *The effects of different types of disruption on one-group pre-post designs*

No Disruption

At the end of the evaluation period, the team will examine the project effects that the treatment group experienced.

Project activities	
Treatment group	Treatment group + **Project effects**

Beginning of evaluation   Midpoint   End of evaluation
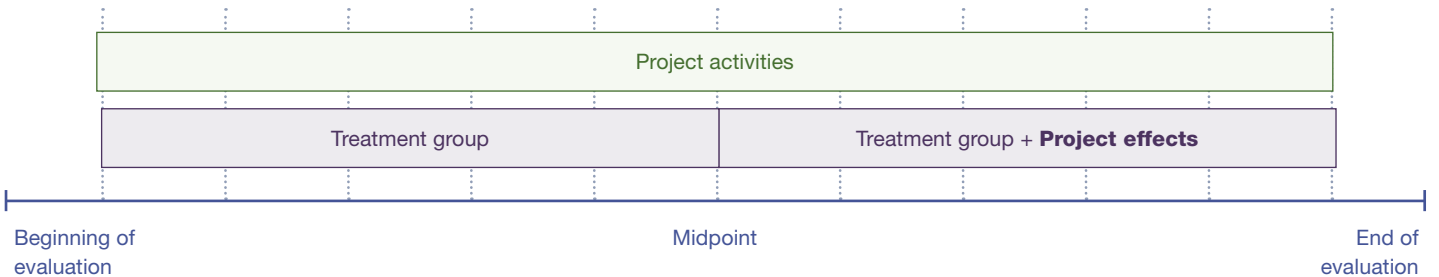
Universal Disruption

A disruption affected all the participants in the treatment group. The team will need to consider the project effects and the disruption effects.

Project activities		
Treatment group	Disruption	Treatment group + **Project effects** + **Disruption effects**

Beginning of evaluation   Midpoint   End of evaluation

Partial Disruption

The disruption affected some participants, but not all. The team will need to consider the project effects and the disruption effects for the groups separately.
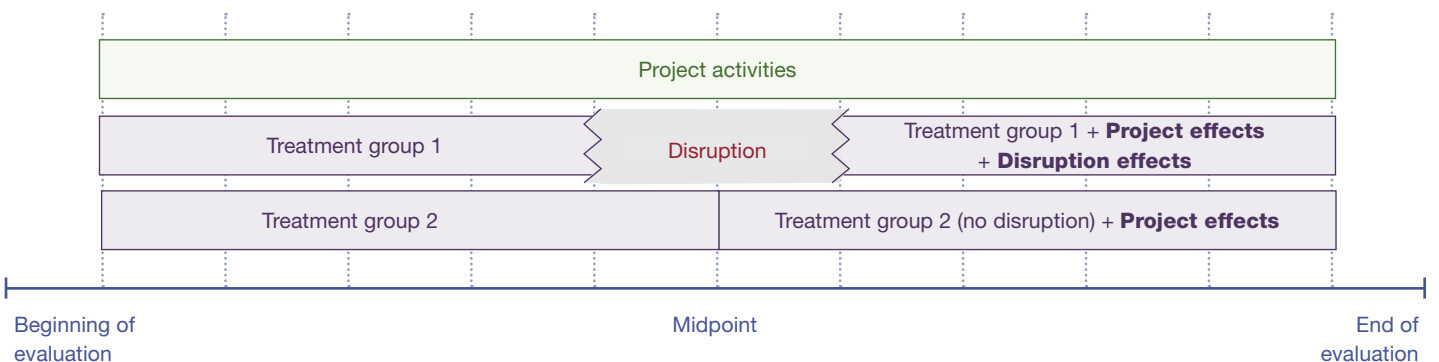
Project activities		
Treatment group 1	Disruption	Treatment group 1 + **Project effects** + **Disruption effects**
Treatment group 2		Treatment group 2 (no disruption) + **Project effects**

Beginning of evaluation   Midpoint   End of evaluation

Exhibit 6. Ways to address disruption in a hypothetical one-group pre-post design

Original Design: A PTI Center did intensive outreach to rural families. After 6 months, they assessed whether they served more rural families during the treatment period than in the 6 months prior.

Who is Affected	Severity of Disruption	Possible Solutions Based on Timing of the Disruption		
		Early Disruption	**Midpoint Disruption**	**Late Disruption**
All rural families	**Mild**	PTI delayed intensive outreach until the disruption was over ▸ adjust timeline for collecting demographic data	PTI suspended intensive outreach until disruption ended ▸ collect posttest data regardless of disruption and add focus groups to identify any effects on post data collection	PTI ended intensive outreach early ▸ end evaluation at the point of the disruption and compile data from the intensive outreach up to the disruption
	Severe	PTI delayed intensive outreach until the disruption was over ▸ adjust timeline for collecting demographic data Disruption affected only some areas of the state ▸ incorporate the disruption as a new element ▸ incorporate the disruption as a change to the environment or context for some treatment or comparison groups	PTI suspended intensive outreach until disruption ended ▸ collect remote qualitative data on participants' experiences with the outreach and plans to engage going forward ▸ assess the validity of planned comparisons after the project activities resumed, given the length of the disruption and any changes to implementation	PTI ended intensive outreach early ▸ end evaluation at the point of the disruption and compile data from the intensive outreach up to the disruption

EXAMPLE 2: INTERRUPTED TIME SERIES

In three districts in each of three states (a total of nine districts), a Technical Assistance and Dissemination (TA&D) Center provided intensive technical assistance (TA) focused on reducing the rate of suspensions and expulsions for students with disabilities in schools with high rates. The evaluation team decided to use an abbreviated interrupted time series design to examine the effects of training offered as part of the TA. The training sessions were implemented in a staggered fashion in each state, with one district in each state being trained in Year 1 of the project, a second district in Year 2, and the third in Year 3. The training always occurred midway through the school year. Because monthly discipline data was available for each district, the evaluation team could obtain the number of suspensions and expulsions for students each month for 4 months before the training and 4 months after. The fact that implementation was staggered and occurred in different locations helped control for history and selection bias. Evaluators compared trend lines across districts and conducted regression analyses to determine the significance of changes in trend lines from before to after the training sessions.

In this example, the evaluation team plans to study the effects of project services delivered to sequential cohorts. There are several project and evaluation factors to consider:

- The disruption may affect one, two, or all three states included in the evaluation and may affect some or all districts within each state.

- The disruption may affect the timing of the data collection.

- The timing and severity of the disruption could significantly interfere with the delivery of the training, which was to be provided midway through the school year. This might make it impossible for the project to implement the project with fidelity to the original plan.

- Depending on its length, the disruption could affect a single cohort (Year 1, 2, or 3) or multiple cohorts.

The team had planned to use data already collected by schools (suspensions and expulsions). The team will need to consider whether schools can reliably capture that data and how meaningful it will be, given the nature and extent of the disruption. Note that if the disruption is mild, it will likely only occur in one year and not in all of the states in the evaluation. If the disruption is severe, it could affect one or more of the states and districts for an extended period of time. If this is the case, the project team will need to discuss whether the original project model is viable or if a new TA model is warranted. If the project needs to consider a new TA model, then the project will likely need a new evaluation design. These changes should be cleared with the OSEP Project Officer/funder. Exhibit 7 presents possible ways to address disruptions to this design.

Exhibit 7. *Ways to address disruption in a hypothetical interrupted time series design*

Original Design: A TA&D Center used an abbreviated interrupted time series design to examine the effects of training on suspension and expulsion rates. Training sessions were staggered, with one district in each state trained midway through Years 1, 2, and 3. The team evaluated monthly suspension and expulsion data for 4 months before and 4 months after the training.

Who is Affected	Severity of Disruption	Possible Solutions Based on Timing of the Disruption		
		Early Disruption	**Midpoint Disruption**	**Late Disruption**
Cohorts 1-3 in one or more states	**Mild**	Center conducted training on schedule ▸ continue with the original evaluation plan Disruption shortened the period of data collection ▸ use statistical adjustments to make the data across cohorts and reporting periods comparable	Center delayed training briefly ▸ continue with original evaluation plan Center delayed training briefly in some states ▸ use only data from states that were not disrupted	Center completed training, but post-training data collection ended early ▸ create new evaluation group and compare data from states that were not affected to states that were ▸ compare data across comparable cohorts and reporting periods
	Severe	Center delayed training until after the disruption ▸ delay evaluation Center continued training for unaffected locations ▸ limit evaluation to unaffected sites	Center delayed training in some states ▸ use only data from states that were not disrupted Center revised training to address the new context ▸ select new outcome measures because of remote learning	Center completed training but did not collect post-training in one or more locations ▸ revise evaluation design to use qualitative data collected remotely instead of suspension and expulsion data

EXAMPLE 3: NON-EQUIVALENT PRE-POST CONTROL GROUP DESIGN

As part of a Personnel Development Program (PDP) project evaluation, faculty rated student performance during their practicum using the Teacher Rating and Assessment Instrument for Teachers of Students with Significant Disabilities (TRAIT-SD). The TRAIT-SD uses classroom observation, teacher interview, and document review to rate a teacher on 37 specific skills. The evaluation team used the TRAIT-SD to rate the scholars participating in the PDP project at the beginning and end of their practicum experience to document gains over the course of the school year. At the beginning and end of the school year, the faculty also used the TRAIT-SD to observe a cohort of scholars who did not participate in the PDP project. Scholars were matched on key demographic variables and their current GPA. Baseline equivalence on the pre-assessment was also calculated. When analyzing the data, the evaluation team adjusted the changes in scores from pre-assessment to post-assessment for the treatment and control groups and compared them to determine if the PDP project contributed to significant student gains on the TRAIT-SD.

In a non-equivalent pre-post control group design, sample size becomes an important consideration in assessing changes to the evaluation. In this example, we will assume that there were 50 scholars in the treatment group and another 50 in the control group. Another important factor in this example is the type of data collection that was originally planned (e.g., a rating assessment that used observations, interviews, and document review). If the evaluation team could only conduct interviews and review documents, and had to do virtual classroom observations, they would need to determine whether the TRAIT-SD data would still be valid and reliable with virtual observations.

Exhibit 8 illustrates how the disruption affects the project and the evaluation. Exhibit 9 provides possible solutions that the PDP's evaluation team might use, depending on the timing and severity.

Exhibit 8. *The effects of different types of disruption on non-equivalent pre-post control group designs*
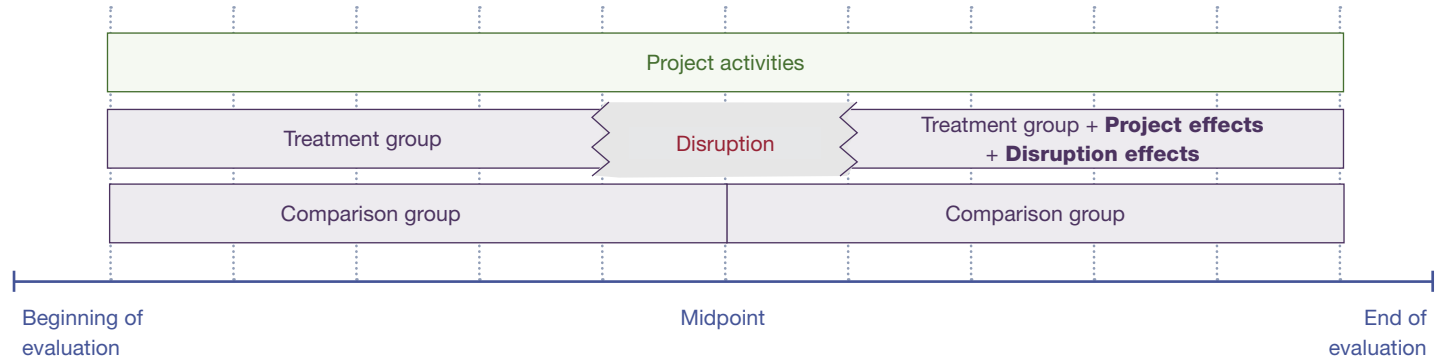
Universal Disruption (treatment and comparison)

A disruption affected both the treatment and comparison groups. The team will need to consider how the disruption affected or altered the groups and in what ways.
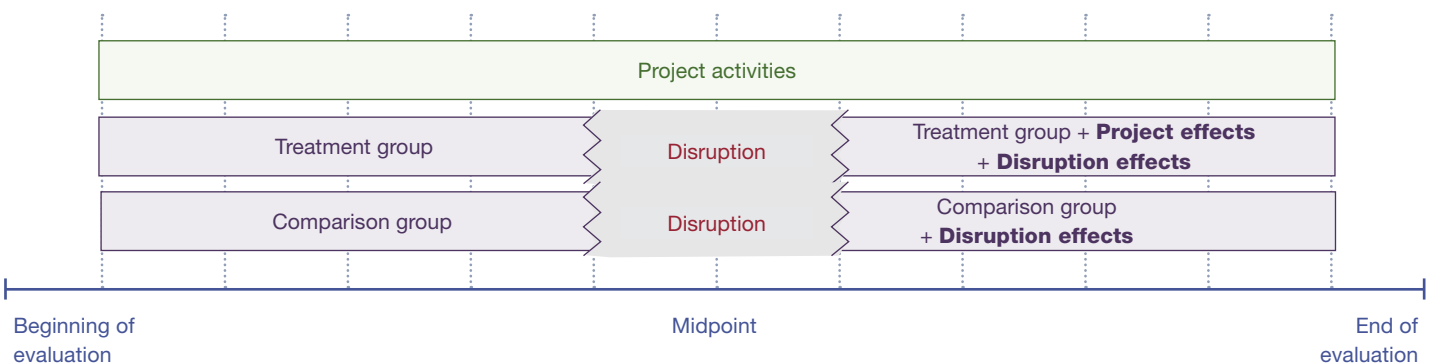
Partial Disruption (treatment)

A disruption affected the treatment group but not the comparison group. To the extent possible, the team will need to consider the project effects and the disruption effects separately in order to use the comparison group as a counterfactual.

Project activities		
Treatment group	Disruption	Treatment group + **Project effects** + **Disruption effects**
Comparison group		Comparison group

Beginning of evaluation   Midpoint   End of evaluation

Partial Disruption (comparison)

A disruption affected the comparison group but not the treatment group. The team will need to consider and potentially measure how the disruption affected or altered the comparison group and in what ways.
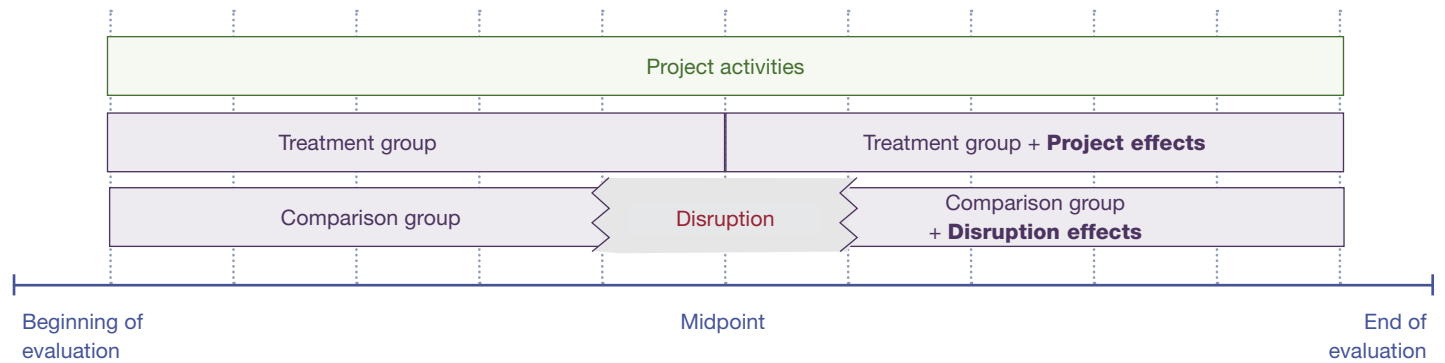
Project activities		
Treatment group		Treatment group + **Project effects**
Comparison group	Disruption	Comparison group + **Disruption effects**

Beginning of evaluation   Midpoint   End of evaluation

Exhibit 9. *Ways to address disruption in a hypothetical non-equivalent pre-post control group design*

Original Design: Faculty rated PDP scholars and matched non-PDP scholars at the beginning and end of their practicum using the TRAIT-SD to assess effects of the PDP project on scholar gains.

Who is Affected	Severity of Disruption	Possible Solutions Based on Timing of the Disruption		
		Early Disruption	**Midpoint Disruption**	**Late Disruption**
Both the treatment and comparison groups	**Mild**	Pre-assessment delayed, but PDP project started as planned ▸ continue with the original evaluation plan with slight delay to pre-assessment data collection	Project activities briefly interrupted after the collection of pre-assessments ▸ collect evaluation data as expected Project activities delayed (and time added to the project calendar) ▸ delay evaluation activities	Project activities completed ▸ determine if enough post-assessments can still be administered to have a sufficient sample for both groups
	Severe	Project activities delayed ▸ delay evaluation	Project activities delayed ▸ delay evaluation ▸ add virtual interviews or focus groups to assess how starting, stopping, and resuming the PDP project affected TRAIT-SD scores	Project activities completed by one, but not both, groups ▸ collect some, but not all, of the post TRAIT-SD data ▸ eliminate observation component, conduct teacher interview component ▸ document incomplete evaluation
Either the treatment or comparison group, but not both	**Mild**	Pre-assessment delayed, but PDP project started as planned ▸ continue with the original evaluation plan with slight delay to pre-assessment data collection	Project activities interrupted after the collection of pre-assessments for some, but not all, participants ▸ shift schedule to ensure full implementation prior to data collection or collect data on the non-affected group first	Project activities completed ▸ shift data collection schedule to collect data on the non-affected group first
	Severe	Project activities delayed for one, but not both, groups ▸ delay evaluation ▸ use original matching variables to assess whether the groups remain comparable	Project activities delayed for one, but not both, groups ▸ delay evaluation ▸ use original matching variables to assess whether the groups remained comparable ▸ assess completeness of practicum experience and include it as a variable in analysis	Project activities completed by one, but not both, groups ▸ collect some, but not all, of the post TRAIT-SD data ▸ eliminate observation component, conduct teacher interview component ▸ document incomplete evaluation

Who is Affected	Severity of Disruption	Possible Solutions Based on Timing of the Disruption		
		Early Disruption	**Midpoint Disruption**	**Late Disruption**

Some members of either the treatment or comparison group	**Mild**	Pre-assessment delayed
		▸ create new study populations:
		  o Treatment and no disruption
		  o Treatment and disruption
		  o Comparison and no disruption
		  o Comparison and disruption
		▸ add variables to the pre and post data collection to estimate the impact of the disruption
		  o verify sufficient sample size across four groups for statistical comparisons
		  o use visual observations for interpreting findings with or without statistical comparisons
			Practicum interrupted after pre-assessment
			▸ create new study populations:
			  o Treatment and no disruption
			  o Treatment and disruption
			  o Comparison and no disruption
			  o Comparison and disruption
			▸ add variables to the pre and post data collection to estimate the impact of the disruption
			  o verify sufficient sample size across four groups for statistical comparisons
			  o use visual observations for interpreting findings with or without statistical comparisons
				Practicum completed before disruption
				▸ revise schedule to collect data on the non-affected participants first

	Severe	Project activities delayed for one, but not both, groups
		▸ delay evaluation
		▸ create new study populations:
		  o Treatment and no disruption
		  o Treatment and disruption
		  o Comparison and no disruption
		  o Comparison and disruption
		▸ modify the evaluation design to study the disruption's effects on the treatment by comparing outcomes for those who experienced the disruption and those who did not
			Project activities delayed for one, but not both, groups
			▸ delay evaluation
			▸ create new study populations:
			  o Treatment and no disruption
			  o Treatment and disruption
			  o Comparison and no disruption
			  o Comparison and disruption
			▸ collect post TRAIT-SD data as planned with the non-disrupted treatment and comparison groups
			▸ suspend the PDP project and evaluation for the group whose program was disrupted
			▸ use visual observations of trends if sample sizes do not allow for statistical comparisons
				Project activities completed by some treatment or comparison scholars
				▸ collect some, but not all, of the post TRAIT-SD data
				▸ suspend evaluation for some scholars
				▸ use visual observations of trends if sample sizes do not allow for statistical comparisons

CONCLUSION

The COVID-19 pandemic highlighted the complexity of disruptions to projects and their evaluations across the country. However, disruptions can and do occur for other reasons, and evaluators should have a strategy for dealing with such disruptions when they occur. When the worst happens, we recommend moving sequentially through the four steps that we outlined in this document:

1. Find your bearings
2. Assess the available options
3. Select the best options for your evaluation
4. Interpret and report results accurately

As our hypothetical examples highlight, much will depend on the specifics of your situation: the implementation activities and the original evaluation plan, the timing and severity of the disruption, and the groups of participants affected. You will not always be able to determine immediately how long the disruption will last or how severe it will be, further complicating your efforts to respond appropriately. With all that in mind, take a deep breath, tap into your creativity, and engage with other professional and community resources. Consider opportunities to collect new data and answer new evaluation questions that were not previously possible. Talk to the project's OSEP Project Officer/funder to get their ideas. Reach out to other project teams and their evaluators to hear what they are planning to do and to share your own ideas. In addition, throughout your planning around the disruption, remember to engage families, project participants, and even other community partners to determine what is reasonable and feasible given the situation.

CONTACT INFORMATION

https://www.cippsite.org | 1-888-843-4101

Thomas Fiore | CIPP Principal Investigator
Westat
ThomasFiore@westat.com

Jill Lammert | CIPP Co-Project Director
Westat
JillLammert@westat.com

Elaine Carlson | CIPP Co-Project Director
Westat
ElaineCarlson@westat.com