vroom

Evaluation Guidebook

A Guide to Evaluating Vroom in your Community

Created in partnership with

Child Trends

vroom

Vroom Evaluation Guidebook:

A Guide to Evaluating Vroom in your Community

About this Guidebook

The Bezos Family Foundation partnered with Child Trends, a nonprofit research organization focused exclusively on improving the lives and prospects of children, youth, and their families, to develop this new resource, the Vroom Evaluation Guidebook: *A Guide to Evaluating Vroom in your Community*. We designed this evaluation guidebook specifically for organizations that are leading the activation of Vroom in a community. Although it is not required, we strongly encourage all communities to understand the impact of Vroom, locally in your region. A starting place is to read through this guidebook and the tools we have developed. Many of you may already have strong evaluation practices underway, and we acknowledge the diversity in those practices, as well as the capacity for evaluation. There is an opportunity to braid Vroom evaluation into the larger piece of work your community might be addressing around early literacy, school readiness, parent engagement and parent success, especially in the early years of life. It is our hope that this guidebook is helpful for *all* organizations, regardless of the resources you have set aside for evaluation.

The purpose of this guidebook is to provide clear, accessible information about evaluation; even if you are not a professional evaluator, you can learn the basic steps of evaluation and benefit from applying it to your work. We intend for this guidebook to support you in an evaluation process that informs continuous improvement of your efforts and documents the effects of Vroom in your community. While we expect that you will individualize your evaluation for your community and your activation, we hope this guidebook will give you useful background, suggestions, and options to consider as you plan and conduct your evaluation. We have also developed a set Vroom Evaluation Tools & Resources that are complementary to this guidebook. These tools include survey and interview protocols designed for parents and Vroom messengers, as well as resources to guide evaluation planning.

Before you read this guidebook, it is important that you first review the Vroom Playbook and draft a Vroom Activation Plan for your community. This plan should outline how you will bring Vroom to your community as well as describe the different strategies you will use to increase awareness and share the Vroom messages and materials with parents. It should also describe your goals for activation and how Vroom can be integrated into existing initiatives or programs in your community. Along with evaluation, we encourage you to document your Vroom activation using stories, photos, and activation metrics. Documenting the activities of your Vroom activation is a precursor to determining the purpose and scope of your evaluation.

We believe there is more to bringing Vroom to a community than can be captured in stories, quantities, or surveys alone. We hope that this evaluation guidebook, along with the developed tools and resources, will support a better understanding of Vroom activation in your community.

A Note about Terminology

We would like to provide definitions for terms used throughout this guidebook:

Activation is the umbrella term for all of the strategies being used to share the Vroom messages and materials. These strategies may be direct or indirect, for example, building awareness in a community by posting flyers or connecting one-on-one with a parent by discussing Vroom tips.

Parents include biological or adoptive parents and guardians. *Messengers* are the individuals directly sharing Vroom with parents. For example, Vroom messengers may include staff at community-based organizations, health care professionals, parent educators, early care and education providers, social service workers, home visitors, or other messengers who work closely with parents and families.

Partners, the main audience for this guidebook, are representatives from the entities coordinating and overseeing Vroom activation. Partners may be community-based organizations, national organizations, state departments, county departments, and more. Depending on your activation, you may have an “anchor partner” who is overseeing “local partners.” Throughout the guidebook we distinguish between these terms when appropriate.

Additional Resources

As we were developing the Vroom Evaluation Guidebook, we were informed by many excellent resources that are available online. Since they were helpful to us, we wanted to share them with you in case you want to read more.

The Program Manager's Guide to Evaluation: Second Edition
Administration for Children and Families, Office of Planning Research and Evaluation

Program Evaluation Guide
The Robert R. McCormick Foundation

W.K. Kellogg Foundation Evaluation Handbook
The W.K. Kellogg Foundation

W.K. Kellogg Foundation Logic Model Development Guide
The W.K. Kellogg Foundation

Evaluation Toolkit
The Pell Institute and Pathways to College Network

The Office for Human Research Protections (OHRP)
U.S. Department of Health and Human Services

Community Tool Box: Evaluating an Initiative
Work Group for Community Health and Development at the University of Kansas

Performance Management and Evaluation: What's the difference?
Child Trends

American Evaluation Association Website
American Evaluation Association

Measuring Success Toolkit: Using Data for Health Program Planning, Monitoring and Evaluation
Measurement, Learning & Evaluation Project for the Urban Reproductive Health Initiative

I. Introduction

Why is evaluation important?

Vroom is being integrated into work happening across the nation with families. From pediatricians sharing Vroom messages about early brain development, to preschool teachers incorporating Vroom into their lesson plans, to Vroom tips on products at grocery stores—Vroom works to highlight the brain building moments in every day interactions, empowering parents to promote brain development in their children during the time it matters most. We expect the integration of Vroom into each community will be adapted to the local context, even while the Vroom messages and materials are standardized. How Vroom Tools and Resources are used and promoted is specific to the organization leading activation efforts and to the community where activation is happening. Since Vroom activation was designed to be flexible, it is especially valuable to understand how Vroom is being shared and to gauge the resulting effects on parents and communities. Evaluating Vroom activation in your community is important for the following reasons:

- To understand what is working—Use evaluation to find out what activation strategies and materials are working well and what is not working well in your community.
- To continuously improve—Use what you learn from evaluation to identify the strengths and weaknesses of your approach and to continually improve what you are doing.
- To show the benefits—Evaluation will help you document your approach and investigate the benefits associated with Vroom in your community.
- To build knowledge—Evaluating these efforts will allow you to contribute to the knowledge base on how Vroom is working in different communities across the nation.
- To inform sustainability—Contribute to the evidence base for Vroom by sharing your evaluation with Vroom stakeholders and using what you learn to broaden and build momentum.

As you embark on this work, we hope that you will value and reflect upon these principles.

- *Flexibility:* Vroom activation is intentionally flexible so it may be adapted to the needs and opportunities present in different communities, and be integrated with existing programs and practices. Vroom partners are encouraged to innovate, find alternative approaches, and create new connections to meet the complex needs of families and their communities. To match such a flexible approach, the evaluation design must be dynamic and responsive.
- *Continuous improvement:* We believe that evaluation should be an ongoing process to inform improvement. The purpose of evaluating Vroom activation is to strengthen what is working well and change what is not, and to the extent possible document the effects of what you are doing.
- *Diverse strategies:* We think using multiple strategies makes an evaluation better. Using a range of techniques that reach various audiences will provide more comprehensive information to answer your evaluation questions. To capture the effects of Vroom, we encourage you to use a mixed methods approach, prioritizing both quantitative and qualitative information.
- *Participatory:* We hope that you will include other Vroom partners and stakeholders throughout the course of your evaluation. With various activations of Vroom happening across the nation, hearing your lessons learned will help bolster other efforts and allow us to gauge how Vroom is working on a larger scale.

- *Sustaining change:* Evaluation should focus on understanding strategies, structures, and innovations that promote the sustainability of Vroom. Sustainability is a layered concept that concerns both the process and outcomes of Vroom activation. Sustainability can be integrating Vroom into ongoing work to ensure that the messages are shared over time in adaptive ways to meet the evolving needs of communities. It is also about the Vroom messages becoming internalized and as a result parents maximize their everyday interactions to promote children's development. It will be complex to assess the sustainability of Vroom, therefore strong evaluations over time are needed to help understand what is working and why it is working.

What do we mean by evaluation?

Evaluation is using systematic methods to collect, analyze, and use information to address questions about the implementation and effectiveness of Vroom activation. There are two sets of objectives for evaluating Vroom activation—the *process* of activation and the resulting *outcomes.* Both types of evaluation are critical in order to fully understand Vroom activation.

Process evaluation addresses questions like:

- How is Vroom activation working in your community?
- What activation strategies are more or less effective?
- What Vroom tools are more or less effective?

Outcomes evaluation will answer questions like:

- How are parents affected by Vroom?
- How is the work of messengers affected by Vroom?
- Is community awareness of Vroom increasing over time?

Process and outcomes evaluations are most effective when they are conducted simultaneously because it allows you to link outcomes back to the activities of Vroom activation. Process and outcomes evaluations should be conducted while Vroom activation is happening, not before or after activation has ended. This will allow you to use the information gathered to continuously improve your efforts while ensuring people are available to participate in the evaluation.

The design of your evaluation is also important to consider. This guidebook assumes that most Vroom evaluations will use non-experimental designs, meaning there will be no comparison group (i.e., a group of people that are strategically not exposed to Vroom in your community). You might choose a non-experimental design that describes your activation at one time point or at pre/post time points. **We believe that a non-experimental design, whether you examine one or two points, is an appropriate design for most Vroom partners.** Please note that other evaluation designs, like randomized and quasi-experimental designs which require adequate comparison groups, will not be covered in this guidebook. If you have a specific audience in mind for your evaluation, please make sure the intent and limitations of the evaluation are considered.

II. Getting started with your evaluation

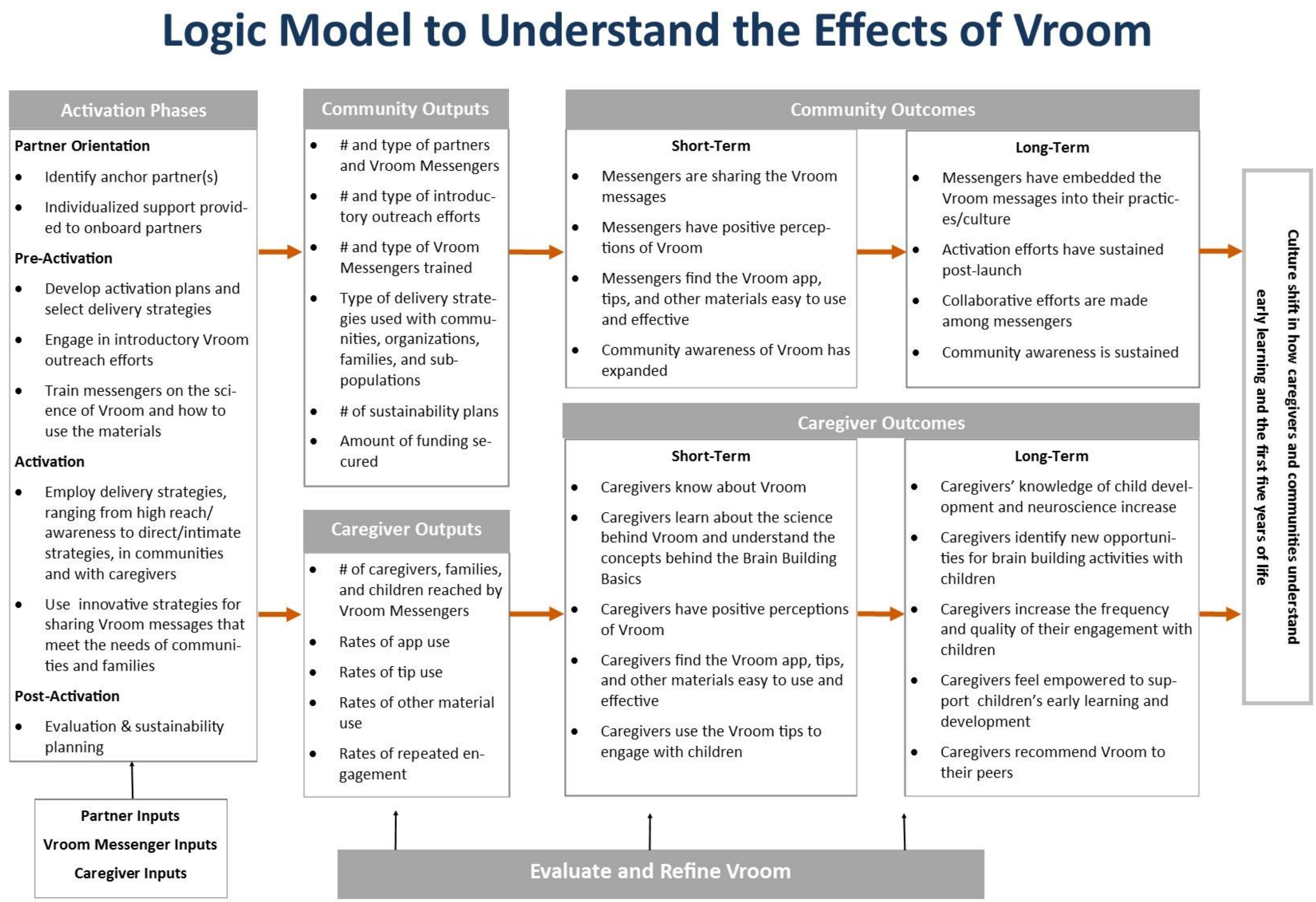
What do you want to evaluate?

The first step of evaluation is planning. One way to build the framework for your evaluation is by creating a logic model or theory of change. A logic model provides a picture of how and why you think Vroom activation will work by showing the relationships and pathways for how activities lead to desired outcomes. Logic model development challenges you to think through your assumptions about the population you are working with and how best to reach them.

We have developed an example logic model for Vroom activation. Our logic model shows how the different Vroom activation phases link to outputs and short- and long-term outcomes at the community and parent levels (Figure 1). We encourage you to use this logic model as a starting point for developing your own. There are many different kinds of logic models, but most include the same core components: inputs, activities, outputs, outcomes, and impacts. Here is how we thought about each of these components while we created the example logic model:

- ***Inputs*** - The resources available for Vroom activation. Inputs may include the messages and materials you are using or other supports that are in place to help you activate Vroom. Examples may include: the Daily Vroom app, tips cards, posters and flyers, and the Vroom Playbook.
- ***Activities*** - What you are doing with your inputs. Activities represent the "action" of Vroom activation. Activities are the strategies and tools you use in communities and with parents to bring about expected changes. Examples may include: training Vroom messengers on how to talk about early brain development, handing out Vroom flyers at community events, using Vroom tips during home visits.
- ***Outputs*** - The direct and measurable results of your activities on the people you reach. Outputs are typically expressed in units, like numbers of people or types of strategies used. Examples may include: the number of parents who hear about Vroom, the types of Vroom outreach efforts you use, the number of parents who download the Daily Vroom app.
- ***Outcomes*** - The specific changes in awareness, knowledge, attitudes, and behaviors you expect to see as a result of your activities. Outcomes are often categorized as short-term or long-term, meaning that you expect certain outcomes to happen in a short time frame (e.g., 1 to 3 months) and others to take longer (e.g., 6 months to 1 year). Examples may include: parents know about Vroom (short-term), parents are using Vroom tips to engage with children (short-term), parents feel empowered to support their children's brain development (long-term).
- ***Impact*** - The fundamental change occurring at the organizational, community, or systems level. The overall impact that Vroom intends to have is a culture shift around early learning. This is complex to track and measure, and involves layering outcomes over time and across communities. We do not expect individual evaluations to assess the overall impact of Vroom. Instead, your evaluation should focus on sustaining Vroom in your community.

Figure 1. Example Logic Model to Understand the Effects of Vroom in your Community



Through your logic model development, you will identify the *indicators* that can be used to track whether and how your goals for activation have been achieved. Using evaluation methods, you can measure these indicators to determine whether your Vroom activation efforts have been successful. For example, all of the bullet points under Activation in the example logic model are process indicators you can measure in a process evaluation. Bulleted community and parent outcomes are the outcome indicators you can measure in an outcomes evaluation. In order to get a complete picture of your activation efforts, each indicator in your logic model should be linked to a measurement strategy in your evaluation plan (see Section 4 for more information on choosing your measurement strategies). It is not expected that you measure all of the indicators in our example logic model; however, we think they provide a good starting place for developing your evaluation questions.

What are your evaluation questions?

Next, it is important to define the questions driving your evaluation—what questions do you need and want answered? When developing evaluation questions, it is important to tie them back to your logic model and the indicators that you want to measure to track progress. We recommend that you group your questions into categories that relate to either process or outcomes evaluation (see Section 1 for definitions), and then develop your primary or overarching evaluation questions under each category. Once you have determined your primary questions you can draft your secondary evaluation questions, which will further operationalize your primary questions. Engaging in this top-down approach will help you decide how to use the information *before* you determine how to collect it.

Exhibit 1 contains a set of example evaluation questions. While we encourage you to consider incorporating these questions into your evaluation, we also hope that you will develop your own questions that fit the unique context of your community using the questions in Exhibit 2.

Exhibit 1. Example Evaluation Questions

Process Evaluation Question Examples

- Primary Evaluation Question:
    - How is Vroom activation working in the community?
- Secondary Evaluation Questions:
    - How are messengers being trained and supported in their work?
    - What Vroom materials are being used? Which are more or less effective and why?
    - What activation strategies are being used? Which strategies are more or less effective and why?
    - What innovations are happening to meet the needs and opportunities of the community?
    - How can Vroom activation be improved and sustained?

Community Outcomes Evaluation Question Examples

- Primary Evaluation Question:
    - How is Vroom being integrated into the community?

- Secondary Evaluation Questions:
    - What are messengers' overall perceptions of Vroom?
    - What are messengers' perceptions of the Vroom materials? Which do they find, more or less, easy to use and effective?
    - What strategies are messengers using to share the Vroom with parents and families? Which strategies are do they think are more or less effective and why?
    - How have messengers incorporated Vroom into their ongoing work?
    - Is community awareness of Vroom increasing over time?

Parent Outcomes Evaluation Question Example

- Primary Evaluation Question:
    - How is Vroom affecting parents?
- Secondary Evaluation Questions:
    - Are parents in the community aware of Vroom?
    - Are parents knowledgeable about Vroom, the background science, and the concepts behind the brain building basics? Does this knowledge increase over time?
    - What are parents' overall perceptions of Vroom?
    - What are parents' perceptions of the Vroom materials? Which do they find, more or less, easy to use and effective?
    - How are parents using the Vroom tips to engage with their children?
    - Over time do parents increase the frequency and quality of their interactions with children?
    - Over time do parents identify new opportunities to interact with their children to promote their development?
    - Do parents feel empowered to support their children's development?
    - Do parents recommend Vroom to their peers?

Exhibit 2. Creating your own Evaluation Questions

It is important to generate your own questions to ensure your evaluation fits your community context. You can refer to your logic model, the examples above, and use the following questions to help guide your thinking:

- What are the unique opportunities for Vroom in your community? What are the unique needs for Vroom in your community? What strategies, tools, and partners are you using to address these opportunities and needs?
- Are there subgroups of people in your community that are a priority for early childhood supports? How has Vroom targeted these subgroups?
- Does Vroom relate to other initiatives or programs in your community? How can these connections be made?

It is unlikely that you will have the time and resources to include every question in your evaluation. After drafting all of your evaluation questions, reflect on your list and decide which ones take highest priority. Use the following probing questions to help you think through what to prioritize:

- *Which questions will be simple to answer and which are more complex?* This is not to say that complex questions are not worth answering, but they may require different kinds of measurement strategies that may be resource intensive.
- *Which questions will provide the most valuable information for understanding your Vroom activation efforts?* At this stage, do you need to know more about how your process is working? Or is it the right time to explore short-term or long-term outcomes?
- *Which questions directly link to the indicators in your logic model and are there any gaps?* It is important to continually revisit your logic model while you are developing your evaluation questions to ensure that there are clear linkages between each question and the indicators in your logic model, and to also check for gaps.

What information is needed to answer your questions?

When you have finalized your evaluation questions, it is time to determine how best to measure them. We recommend that you use our Evaluation Planning Discussion Guide (see Vroom Evaluation Tools & Resources). This resource will guide you through the process of identifying the appropriate respondents, methods, tools, and timeline. Section 4 of this guidebook will provide you with the details on how to get the information you will need to answer your questions.

III. Conducting a responsible and high quality evaluation

How do you conduct a responsible evaluation?

Before you start your evaluation, you must review the federal, institution, and ethical guidelines for research and evaluation. The purpose of Institutional Review Boards (IRBs) is to weigh the risks and benefits of participation in research and evaluation to protect the rights and welfare of the participants. An IRB must review your evaluation study if it matches the Federal Policy for the Protection of Human Subjects (Common Rule) definition of research—"a systematic investigation, including research development, testing and evaluation, designed to develop or contribute to generalizable knowledge..." Here is a decision tree to help determine whether your efforts are considered research and require IRB review or not. IRB review is not required if your project does not match the definition of research**. In many cases, an evaluation of Vroom activation will not necessitate IRB review if it has the primary goal of informing continuous quality improvement and the information gathered will not be generalized to the broader population.** However, if you plan to share your evaluation findings to external audiences, it is critical that you examine the federal policy definition for research and determine if you must submit an IRB application.

It is inevitable that evaluations of Vroom activation, whether or not they require IRB approval, will involve human subjects. The Belmont report summarizes the basic ethical principles and guidelines for research and evaluation involving human subjects. As you develop your evaluation plan, there are **three basic ethical principles** you must follow to ensure your evaluation methods and procedures are responsible for human subjects: respect for persons, beneficence, and justice. Next, we describe each principle and how you can uphold it in your methods and procedures throughout your evaluation. We also recommend that all individuals who will be involved in conducting your evaluation (e.g., individuals collecting and analyzing data) take a basic course in human subjects research training.

Respect for persons means that people are free to make their own decisions and that special protections are in place for vulnerable populations (e.g., children, those with mental illness). For your evaluation, this means that all participation is voluntary. In order to enable someone to voluntarily decide whether or not to participate in your evaluation, you must educate him or her about the evaluation study in terms that are easily understood; this is referred to as an informed consent process because the focus is on understanding, not just information-sharing. The Office for Human Research Protection provides an informed consent checklist that covers the main topics to include in an informed consent process. In addition, the interview and survey tools we provide in the Vroom Evaluation Tools & Resources include sample introductions that cover the important points to be adapted for your purposes.

Beneficence refers to a "do no harm" approach in which people are treated with kindness by respecting their decisions, protecting them from harm, and making efforts to secure their well-being. *Beneficence* is especially important both when you are considering the risks and benefits of evaluation procedures and when you are deciding what to do with the information that is collected. To uphold this principle, you must have a system for protecting the privacy and confidentiality of the people who are participating. Much of this responsibility relates to the procedures you use to protect the data you collect. Whenever possible, try to limit the amount of identifying information you collect, which includes but is not limited to: names, addresses, birthdates, phone numbers, email addresses, social security numbers. One way to protect the confidentiality participants is by assigning each participant an evaluation ID and using those ID numbers instead of names on all evaluation documents and records. Keep a separate document linking the evaluation ID and identifying information in a secure place with limited access (e.g., a locked file cabinet; secure electronic servers). If you are tracking participants over time, use caution to track their IDs and protect their personal information. You should also protect confidentiality by keeping all data files, participant lists, notes, recordings, etc. in a secure place that is only accessible to the evaluation team. And, remember to never use identifying information in your reporting of results.

Justice means that benefits and burden of the research are fairly distributed among people. This principle especially applies to how you select people to be part of your evaluation. There should be no undue burden on any people. This means that while you recruit people to participate in your evaluation, you must provide equal opportunity to all eligible people. And, if you decide to provide incentives for participation, whether it is food and snacks or cash, all participants receive the same benefits.

How do you get high quality information?

Evaluations take planning and resources, so it is important to ensure that you collect high quality information. High quality information is information that is valid, reliable, accurate, complete, and consistent across participants.

Having *valid* information means that what you are measuring and the conclusions you draw from the information you collect are true and consistent with the topic and the real world. A key element of valid information is related to the tools you use for the evaluation. You want, to the extent possible, to use tools that have been developed and tested. *Reliable* information means that the information is measured consistently across time and across respondents (note this is different from validity which focuses on whether your measures capture the content and process that you expect).

Accuracy or ensuring that the information you collect is free from error is a part of producing high quality information that can be checked through quality assurance procedures. Random quality checks at different stages of the evaluation can help ensure quality information or uncover issues that should be addressed. In order to draw conclusions about what your information means, it is important to have *complete* information. This means that participants answer most questions and that the number of people participating is adequate for drawing conclusions.

IV. Collecting data

Who has the information you need?

There are many roles involved in Vroom activation. When deciding what information you need to answer your evaluation questions, ask yourself, *who are the best respondents?* The answer to this question depends on the type of evaluation you are conducting and your main evaluation questions.

The relevant respondent group for *process* evaluation questions may be different than those needed for *outcome* evaluation questions. Most likely, the respondents for a process evaluation will be the individuals closely involved with overseeing or carrying out the activation of Vroom (e.g., messengers). Note that administrative data and outputs are often important data sources for process evaluations. Since many of the Vroom outcomes are at the parent level, parents and caregivers will likely comprise the primary respondent group for an outcome evaluation.

When selecting the best respondents you should consider: Who is available and accessible to me? How can I decrease the burden on that person (for example, the time needed to complete a survey, or transportation and child care needed to participate in a focus group)? What will my resources available for evaluation allow? You should also consider the resources required for survey administration and organizing your collected data. Could these tasks be completed by a contractor, volunteer or postdoctoral researcher? After you decide on appropriate respondents you should consider your approaches to *sampling* and *recruitment*.

Sampling is the process of selecting a subset of people, or a "sample," to participate in the evaluation that represent the larger community from which they were chosen. In order to make an inference from your sample to the community, the evaluation sample needs to be big enough to ensure adequate representativeness. Choosing the right *sample size* or number of people in your sample should be based on the methods for collecting information and type of analyses you would like to conduct. Advanced statistical methods (which we describe in Section 5) typically require large sample sizes (e.g., more than 100). If you plan to do these types of analyses we encourage you to seek training and resources to conduct a power analysis, which is an analysis to determine the adequate sample size for a particular statistical method. However, if you plan to do the descriptive analyses that we recommend in Section 5, a smaller sample size is generally fine (e.g., less than 100, more than 20). You can make sampling decisions based on the population you have access to, the tool you are using, and your capacity. For example, you may decide to send a web-based survey (which is a lower-cost option) to all eligible people to get a higher response rate. Conversely, you may only invite a subset of the eligible population to participate in interviews (which require more time). *Response rate* is another important factor to consider. In general, it is preferable to get a high response rate (80% or higher), but this is often not realistic in evaluation. **A general rule of thumb is that a higher response rate is best if your sample is small and a lower response rate is acceptable if your sample is large.** It is important to consider your audience too—you should expect a high response rate (80% or higher) from people working internally on Vroom activation (e.g., messengers) and you can expect a lower response rate (50% or less) from people external to your evaluation (e.g., parents).

Recruiting people to participate in your evaluation can happen in a number of ways. You might recruit participants through advertisements (e.g., posters, flyers, postcards), social media (e.g., Facebook, Twitter), email blasts, or through messengers in your community. Recruitment materials should include a short description of the purpose of your evaluation and how you will use the information; what participation entails, including a description of the time commitment; and whether an incentive is provided (e.g., money, food, gifts, etc.). Recruitment efforts tend to be most successful when the messages are short and clear, appropriate reminders are given, and adequate incentives are provided.

Which method will provide the type information you need?

To answer your evaluation questions, you can use one of three methods or approaches: quantitative, qualitative, or mixed methods. A *quantitative method* is an objective approach for describing and testing theories about relations between variables. Variables are factors or characteristics that can change over time or take on different values in different situations. To measure variables, quantitative methods use instruments (e.g., surveys or administrative data) with responses that can be numbered so that those numbers (representing differences in responses) can be analyzed using statistical procedures. A *qualitative method* is a more subjective approach for exploring and understanding a topic more openly (without pre-determined responses) so that a topic is understood from the perspective of the respondents, using their own words. A *mixed methods* approach involves both quantitative and qualitative information that are integrated to capture a more rounded understanding of the evaluation questions (Exhibit 3).

Exhibit 3. Quantitative, mixed, and qualitative methods

	Quantitative Methods	**Qualitative methods**	**Mixed methods**
Type of responses	Pre-determined or close-ended responses	Open-ended responses	Mix of pre-determined and open-ended responses
Examples of data	Ranking and multiple choice survey items, administrative data and outputs (e.g., number of attendees at an event)	Interviews, focus groups, stories	Multiple forms of data
Approach to analyses	Descriptive and inferential statistics	Text analyses	Statistical and text analyses
Approach to interpreting the data	Statistical interpretation (e.g., discuss how strong two variables are related)	Themes or patterns interpretation	Mix of statistical analysis and interpretation of themes or patterns

When deciding which method to use, you should ask yourself the following:

- *Which method is the best for answering my evaluation question?* For example, when asking parents about their use of Vroom materials, you may want to use mixed methods and ask both pre-determined and open-ended questions.
- *Given the amount of resources available, which method will provide the most value?* For example, if resources are limited you should consider using surveys that can be distributed and completed during community meetings to get the most use of your resources. While observations of parent-child interactions would provide rich, in-depth information, this may not be feasible with limited resources.
- *Given our capacity, what is feasible and what is not?* As you identify the members of your evaluation team, keep in mind the various steps of an evaluation: planning, collecting information, monitoring the collection of information, analyzing information, reporting information. While these steps can be done at the same time, the work load increases as the number of participants increase, so make sure you have the capacity to implement the evaluation plan you create. Evaluation can be conducted effectively even when resources are limited by select methods that require fewer adaptations and that can be launched using low-cost strategies, such as web-based surveys or a small number of interviews.

What tools will capture the information you need?

After you decide which respondents can answer your questions and determine what method will collect the best information, it is time to select your tools. The instruments or measurement tools you choose should be guided by the level of detail and type of information needed to answer your evaluation questions. For example, if you want to measure change over time, you can use a tool(s) designed to be administered at multiple time points. The tools should also align with the methods; for example, if you have chosen quantitative methods, you can use a survey. Common evaluation tools include document review, surveys, interviews, questionnaires, focus groups, observations, tests and assessments. We encourage you to use a variety of tools and, whenever possible, use tools that have already been developed and tested. **It is important that the tools you select are standardized, so that your respondents all answer the same questions in the same order.** When selecting the tools to use, you should ask yourself:

- *Will the tool(s) give me the information I need to answer my evaluation questions?* For example, if you are interested in changes in parent attitudes, practices, and knowledge, you will want to use a tool that measures these constructs. We would recommend that you consider using or adapting our Vroom Parent Survey or Interview. On the other hand, if your question is related to implementation you may want to consider other types of tools. For example, if your question is about attendance at trainings or events, tracking outputs like attendance records will provide the most useful information.
- *Which tool(s) are appropriate for my sample (e.g., reading level, language, ease of use, not too long)?* In general, it is best practice in evaluation to ask simple, straightforward questions of your respondents, especially when the reading ability of the individuals may vary. If your community has a dominant language other than English, you may also consider having evaluation tools translated into those languages. It is also important to consider formats that will work best for your respondents. For example, surveys may be completed on paper or through an online survey tool like SurveyMonkey.
- *Will the tool(s) respect and reflect my sample (e.g., culture, everyday experiences)?* Make it a priority to learn about the everyday experiences and values of the people being included in your evaluation. Before you can decide on the best tool to use for your evaluation, you want to make sure it respects and reflects the community. A good way to determine whether your tools work well for the community context is to do a test-run with a willing member(s) of the community. For example, invite someone to complete a survey and afterward have a conversation with them about the survey questions—did they make sense, did they seem relevant, was anything missing?

As previously stated, we have developed tools specifically for evaluations of Vroom, including surveys and an interview protocol for administering with parents, and a survey and interview protocol designed for Vroom messengers. Each tool comes has a “User’s Guide” with information on the purpose of the tool, the intended audience, and guidelines for administration. While we encourage you to use the tools we have developed, we understand that in some cases you may need to revise or add items to these tools to specifically answer your evaluation questions. If you do chose to develop your own tools, it is important to consider the type of information needed to answer your evaluation questions, as well as your potential sources, and conduct pilot testing to ensure the tools are designed to give you the right information.

How will you monitor your data collection?

Before you begin collecting information, you should outline your evaluation procedures and track your activities. There are several strategies you can use to monitor your evaluation. It is a good idea to set a timeline for your information-collecting efforts and decide where, how, and who will collect the information. You can use the “Evaluation Plan Discussion Guide” (see Vroom Evaluation Tools & Resources) to develop this information.

After data collection has begun, there are a few ways you can monitor the data that is being collected. First you want to track your response rates (which are described earlier in this section). If you find that response rates are low, you could try new recruitment strategies. You may also want follow up with people you initially contacted and remind them to participate if they have not yet done so. You can remind people over email, the phone, or by sending flyers or mailings. Just remember that participation is voluntary, so you should only contact people an appropriate number of times (for example, 3 times or less). Second, it is important to review initial data during the early stages of your data collection period to check for signs that people are misinterpreting questions, look at the completeness of the data, identify glitches, etc. Data checking can reveal whether your tools are working as intended and if the information you are gathering is high quality (see Section 3 for a description). If you determine there are any issues, you can address them before too much data has been collected.

V. Analyzing and interpreting data

Now that you have collected data, it’s time to analyze and interpret your results. This process can be complicated and technical—there are many textbooks and college courses dedicated to this topic. It is not possible for this guidebook to adequately prepare you to do advanced analysis and interpretation; however, we can describe some basic techniques you may use to organize and analyze quantitative and qualitative data. We focus this section on survey and interview methodology to align with the tools and resources that accompany this guidebook. We understand that you may have collected data using other tools, and if this is the case, we encourage you to seek resources outside of this guidebook to inform your analysis.

At this stage of your evaluation, it is a good idea to revisit your logic model and evaluation questions. Since several months may have passed while you were collecting data, reminding yourself of the goals for your evaluation will help you identify your top priorities for data analysis and select the best analysis strategies.

How will you analyze your quantitative data?

First, let's talk about quantitative data you may have gathered through a survey. These data are information in the form of numbers and other quantifiable items in the surveys such as counts, ranking questions, and multiple choice items.

How will you organize the data?

When you collect data using surveys, you may have collected this data on paper or through an online survey. Regardless of how you collected the data, you will want to pull together all of the data into an electronic spreadsheet for analysis using software such as Microsoft Excel.

If you have surveys completed on paper, you will need to enter the data into a spreadsheet. Because it can be easy to make mistakes during data entry, evaluators often use double-entry; that is, they have two different people enter the data independently, then compare their datasets and correct any errors. After you have corrected any discrepancies between the two versions of your dataset, use one of the versions as the final master file for your data.

It is important to organize the spreadsheet in a certain way for your analysis. Your spreadsheet should be organized so each row of the table contains the data from a participant in your evaluation and each column is a variable or question from your survey. Short but clear variable names are helpful to keep you organized. Each participant who completed your survey should have a unique identification number. Typically this ID number is entered into the first cell of each participant's row of data. If you used an online survey tool that prepopulated data into a downloadable spreadsheet, you should check to ensure the spreadsheet is organized in this way.

For quantitative analysis you will need to code some of your data into numerical values. This is done so that the analysis software can read and analyze the data. Sometimes your data will be meaningful numbers such as age of participant or number of children living in the participant's home. However, more often the data will be given in words or categories such as male/female or yes/no. During the coding process, pay close attention to survey items that were worded negatively (e.g., disagreement with a statement is a positive answer) and reserve your codes as appropriate. Here are a few examples of how you want to transform text responses into numeric values:

Q1: Do you interact regularly with a child or children between the ages of 0 and 5?

0 = No
1 = Yes

Q2: What is your relationship to the child(ren)?

1 = Parent or guardian
2 = Relative
3 = Neighbor
4 = Community member
5 = Early care and education provider

Q3: In a typical day, how often do you respond to your child's sounds, actions, and words?

1 = Never
2 = Rarely
3 = Sometimes
4 = Often

Here's a portion of what your electronic spreadsheet might look like (Exhibit 4):

Exhibit 4. Example data entry spreadsheet

	A	B	C	D
1	ID	Q1_Regular Interact	Q2_Relati onship	Q3_Respond Sounds
2	101	1	5	3
3	102	1	5	2
4	103	1	1	3
5	104	1	1	4
6	105	1	1	4
7	106	1	2	4
8	107	1	3	4
9	108	1	4	1
10	109	0	99	99
11	110	1	1	4
12	111	1	1	4
13	112	1	1	4

You may also want to assign a code for missing data. Having missing data is very common—a respondent may have accidentally skipped a question, have chosen not to respond, or there was a technical glitch. Regardless, it is important to have a consistent way to refer to missing data, whether it is left blank, noted with a period, or an unrealistic value like "99." We encourage you to *develop a codebook* that lists each variable name, the answer options for each question, and the definitions of your codes. This will be important for you to keep track of what values mean and keep your data organized.

Conducting double data entry or relying on an online survey tool will safeguard again most errors in your data. However, before you start analysis, we recommend that you do some final data "cleaning" to check for any remaining errors. A useful approach in data cleaning is to search for data values that are not valid by spot checking the data or running descriptive statistics (i.e., means, ranges, frequencies). For example, if you coded a yes/no question with 1=yes and 0=no, finding a value of 7 in your data would reveal an error in data entry. Examining frequencies for this variable would quickly reveal this type of error. If a question asked for the respondent's age, a response of '360' indicates an error. Finding the range of values in your data will quickly expose errors like this that are higher or lower than is reasonable.

How will you analyze the data?

Now that your data set is cleaned and organized in a spreadsheet, you need to choose what software you would like to use for analysis. Microsoft Excel and Survey Monkey have several analytics features, so you may want to continue using these software. If you would like to do more complex quantitative analysis, we recommend using a statistical program like SAS, SPSS, STATA, or R.

The first step in data analysis is to describe your data. *Descriptive statistics* such as frequency counts and means (that is, averages) will provide an overall sense of how your participants responded to your survey questions. The descriptive statistic you choose will depend upon the type of variable you want to analyze. Variables that are categories (e.g., gender; racial or ethnic background) can be described using frequencies. Variables that are numeric (age; number of children living in the participant's home) can be described using means.

- Tables of *frequency counts* are a helpful way of getting an overall feel for your data. Using them can give you a sense of what percentage of respondents fit various demographic categories, for example, or what percentage agrees with various statements in your survey. Most data analysis software packages have a straightforward way of asking for frequency counts on variables. The output you will get from your software package will likely include the frequency counts for each of the categories (e.g., male, female), the total number of responses, and the number of missing responses. The output will also give you the percentages of the total number of responses for each category. Some software packages will include "valid percentage" in the output – this is the percentage of responses out of the total of non-missing responses rather than the total number of participants who took the survey.
- For variables that are meaningful numbers and not just categories, the *mean*, or average, is the most common descriptive statistic. The *range* is the difference between the highest and lowest value in your data for a particular variable. From these statistics you can learn the range of age of your participants, for example, or the number of children they interact with.
- You may want to continue to explore your data by disaggregating it across different categories, which is commonly called *sub-group analysis*. For example, if you are interested to learn which respondents report using a particular Vroom material, you could examine the responses to this question sorted by various demographics, like gender, ethnicity, age, and more.
- If you were able to collect pre and post data for your Vroom evaluation (i.e., before and after your activation), you can compare the data to look for trends over time. For example, you can examine whether the number of parents who know about Vroom increased from the first time point to the second time point.

In addition to the basic methods we have described, there are a variety of more complicated analyses you may want to use. These techniques often require a large sample size and statistical software (i.e., SAS, STATA). They also require sophisticated interpretation skills. We provide basic descriptions of these analyses, but encourage you to seek additional information and training before using these procedures.

- Correlational statistics describe possible relationships between two variables. They answer the question, *"Does one variable tend to change (increase or decrease) as the other changes?"* A positive correlation between two variables means that as the value of one variable increases, the value of the second variable increases as well (a positive correlation) and conversely, as the first decreases, the second decreases (a negative correlation). In addition to telling you whether a relationship is positive or negative, correlations will tell you if the relationship is weak or strong.
- There are statistics you can use to test whether groups significantly vary on a variable of interest. To know whether two groups differ on some numeric variable, you would use a statistic called a *t*-test. For example, a *t*-test can be used to reveal whether male and female participants significantly differ in their perceptions of Vroom. If you would like to test the statistical significance between two or more groups, you can use an analysis of variance test (ANOVA). For example, an ANOVA could reveal whether perceptions of Vroom significantly differ depending on the age of the parents' child (e.g., infant, toddler, or preschooler).
- Regressions can be used to determine whether a variable significantly predicts another variable. For example, a regression could help you determine if using the Daily Vroom app on a regular basis predicts higher rates of engagement with children. Similar to correlations, regression will tell you if the relationship is positive or negative, and strong or weak.

How will you analyze your qualitative data?

Now let's talk about qualitative data you may have gathered through an interview or survey items. Qualitative data cannot be expressed in terms of numbers, only by words and descriptions. Open-ended survey items and interview questions are common sources of qualitative data.

How will you organize the data?

How you organize qualitative data from interviews or open-ended survey questions depends on your approach to gathering it. If qualitative data were collected through interviews, we recommend that the person or people who conducted the interviews take time soon after the interview was completed to clean their notes or transcribe the interview (if it was recorded). Notes should be organized by interview question or topic, and interviewers should document their reflections on the highlights and key points from the interview. All interview notes, whether they were collected on paper or typed in a Word document, should be stored together so all of the data is in the same location. If qualitative data were captured by open-ended questions in a survey, we recommend that you enter the information into a Microsoft Word or Excel document (online survey tools typically do this for you) so that you have all of the data in one place. Again, the responses should be organized by question or topic.

How will you analyze the data?
Regardless of how you gathered and organized the qualitative data, the analysis process looks similar. *Thematic analysis* is a process to identify categories and patterns in qualitative data. Themes are clusters of related information that can be used to describe what is learned from the qualitative data. The practical approach to this type of analysis varies. You can do thematic analysis in a group setting where the individuals who conducted interviews or reviewed the open-ended survey responses debrief together and derive themes from their conversations. Conversely, these individuals could review notes and responses independently to determine themes. If the latter method is used, there still should be a process for the themes to be reviewed and vetted with other team members.

When you offer open-ended questions to participants, it usually results in a plethora of qualitative data, but not all of it is necessarily important for your evaluation. Thematic analysis will help you drill down to the most relevant, meaningful, and useful findings for the purposes of your evaluation. It is important to review all of the data as a whole first and then start reducing it to reveal the themes that are naturally emerging. Themes should have codes (i.e., labels) that categorize the idea behind the data. For example, if the open-ended question you are analyzing was "Why do you think the Daily Vroom app is easy to use?" your codes might be "Easy to understand", "Easy to access", or "Fits into my schedule". One code or multiple codes should be assigned to each meaningful response from the interview or open-ended survey question.

To analyze for themes, whether it is done as a group or individually, it is important to keep guiding questions in mind:

- What interesting stories are emerging from the data?
- What are the common themes emerging related to the topics in the interview or survey?
- Where are there deviations from common themes?
- How do the emerging themes support or undermine other analysis being done?
- Do these themes (or absence of themes) help address any evaluation questions?

You may start with general themes first and then get more specific with your codes. As your themes evolve, you will want to continuous review and double check your codes to make sure they properly fit. After all qualitative data are coded we suggest having a clear description about what each code means and count number or percentage of participants where the code was applied. You may also decide to track quotes from participants that are illustrative of the themes.

How will you interpret your data?
Interpreting your data is a continuous process throughout your analysis. At points throughout the process and at the end, it is important to take a step back to think about the context of the data and the story it tells you. Here are some common issues evaluators encounter when interpreting quantitative and qualitative data:

- Remember the limitations of your evaluation methods and self-reported information. People may not always answer your questions truthfully and accurately. People are prone to providing socially desired responses.
- Be careful not to assume that Vroom is the cause behind your findings. The quantitative associations and relationships you find in your data do not imply causation (unless you have an experimental design). For example, if you find that parents who use the Daily Vroom app report doing more brain building activities with their child, this does not necessarily mean that using the app *caused* this behavior. It simply means that these behaviors are associated—these variables move together but you do not necessarily know why. There are many other factors that you are not measuring that could be influencing parent's use of brain building activities with their child. This is where qualitative data can be helpful to understand the nuances of the quantitative data.
- Reflect on the diversity in your community when you are attempting comparisons between groups of people. Demographic characteristics can help you draw appropriate comparisons between similar groups of people and avoid comparing groups that are too different from one another.
- Take care to only reflect on the data you have analyzed when you are drawing conclusions. Avoid making any claims that your evaluation data represents a wider audience or geographic area than what you studied.
- Look at the nuances in your data and remember not to place too much weight on singular data points. Take time to group your data in meaningful ways and consider each piece of data in the context of other data. For example, let's say that five of your survey questions asked about a certain topic and the responses to one of these questions were negative, while responses to other 4 questions were positive. You would not want to weight the negative response more than the 4 positive responses.
- Be open to unexpected findings. It is important to look for themes or patterns in the data that may be different from your expectations. Exploring the data that do not confirm your expectations can sometimes provide more valuable information than the data that does.

Remember that quantitative data is not just a number and qualitative data is not just anecdotal. Careful and coordinated interpretation of both types of data can develop a complete picture of what your respondents are reporting. This is why we believe that mixed methods approaches provide the best opportunity to answer your evaluation questions, to understand how your Vroom activation is working, and to start gauging the effects in your community.

VI. Using and sharing your findings

Now that you have analyzed and interpreted your data, how are you going to use and share what you have learned? This question brings us back to the five main purposes for evaluating Vroom activation in Section 1: to understand what's working, to continuously improve, to show the benefits, to build knowledge, and to inform sustainability. The way you use and share the findings from your evaluation will bring these purposes to life. This section provides suggestions about how to use your findings to promote continuous quality improvement of your Vroom activation, as well as ideas for sharing your results with external audiences.

How will you use your findings to inform continuous quality improvement?

As we stated at the beginning of this guidebook, the purpose of evaluating Vroom activation is to strengthen what is working well and to change what is not. We believe that evaluation should be an ongoing process to inform improvement and to document the effects of what you are doing when possible. There is always something to learn from evaluation. Sometimes, findings suggest making major changes, but more often, findings can point to incremental improvements that can be made, while keeping the goals for your Vroom activation and the needs of your community in mind.

What you learn from your evaluation may have different implications for your work, depending on the role you play in activation. As an anchor or local partner, you might use evaluation results to inform how to better prepare your messengers to talk about brain development or how to channel your resources to have more direct engagement opportunities with families. Your findings may also reinforce some of the strategies you are already doing, for example, your presence at community events being linked with increased awareness of Vroom.

For Vroom messengers, evaluation findings might influence their interactions with parents. For example, if parents are reporting that they do not understand the Vroom messages about brain development, messengers may need to spend more time on this topic before sharing the Daily Vroom app and tips. Evaluation findings may also guide you to which Vroom materials are working better than others in your community. For example, if parents are reporting that the Daily Vroom app is the most useful to them, a messenger may want to reconsider distributing the hard copy tip cards and prioritize helping parents to download the app instead.

With the benefits of evaluation, it is possible that your findings or lack thereof will leave you confused. Evaluation findings not only provide you with feedback on your activation efforts, they also might reflect limitations of your evaluation and indicate how future evaluations could be improved. Take time to reflect on your overall evaluation study and ask yourself these guiding questions:

- As you re-examine your logic model, does your theory of change still make sense? Did your evaluation adequately assess key process components, outputs and desired outcomes? Which pieces of your logic model were not assessed?
- As you reassess your evaluation questions, which ones were addressed and which ones were not? What do you think prevented them from being answered? Have your questions evolved over time?
- As you think about your sample of participants, was your sample too small or not representative enough of your community? If so, how can recruitment efforts be improved?

- As you reflect on the methods and tools you chose for your evaluation, did the tools provide you with the right kinds of data? Was the data high quality? Where were there gaps in your data? Would other tools or methods work better?
- As you contemplate your analysis and interpretations, are there other analysis strategies you should try? What are some other ways to interpret your data? Are there other limitations to your data that you haven't considered?

Overall, as you reflect on how to use your evaluation findings, consider how the results inform sustainability planning. Recall that the goal of Vroom is to layer and braid with the ongoing work with families happening in your community. Identify findings that speak to how Vroom is being integrated into the daily practices of messengers and parents, and look for indications of how behaviors may be changing over time. All results can be used to inform strategic planning and to make course corrections necessary to promote sustainability.

How will you share your findings with others?

First and foremost, it is important to share your evaluation findings with your activation team to discuss lessons learned and the implications for your work. Outside of your internal team, sharing your findings may be beneficial for you, your community, and the early childhood field. While you are considering ways to share your evaluation findings with external audiences, remember to check whether or not they the match the Federal Policy rule for research or deviate from your IRB submission (if applicable). And, always remember to protect the privacy of your participants in your reporting.

To share your evaluation findings, you need to summarize and present your results in a meaningful way. It is important to consider your audience when determining what information is important to share and how it should be presented. Evaluation findings can be effectively described in text, tables, charts and graphs, infographics, and other visuals. You might decide that an evaluation report, executive summary, or PowerPoint presentation will work best to meet the needs of your audience. You may also want to generate high level summary documents, like a one-page handout, flyer, or poster that highlights your key findings. You might consult with communications professionals to help you identify and develop a strong set of key messages, and to ensure that your language and outreach strategies are effective for the audiences you want to reach.

After you have summarized your key findings and decided how to present them, here are some ways you might share your results externally:

- Include evaluation findings in your organization's annual report.
- Cite your evaluation results in your organization's proposals to promote Vroom to potential funders.
- Disseminate key findings with the broader public through press and media promotions, workshops, or forums.

- Share positive findings with your community through different avenues, like attending community events, posting flyers, and more. Use the channels you already have in place to spread the word.
- Present your results with the broader early learning community at conferences or other professional gatherings.
- Learn together with us and share your evaluation experience and newfound knowledge with Vroom Learning Community!