SAS-HV

Supporting and Strengthening
the Home Visiting Workforce

MEASURE MANUAL | VERSION 1, SEPTEMBER 2025

Supervisor Practices in Reflective Supervision–Home Visiting (SuPRS–HV)

OPRE Report 2025-126

MEASURE MANUAL | Version 1, September 2025

Supervisor Practices in Reflective Supervision–Home Visiting (SuPRS–HV)

Supporting and Strengthening the Home Visiting Workforce

OPRE Report 2025-126

Authors

Mariel Sparr, Eleanor Fisk, and Jessica Bruning, James Bell Associates
Allison West, Johns Hopkins University
Nancy Asdigian, Colorado School of Public Health
Jessica Fulford

Submitted to

Nicole Denmark, Project Officer
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services
Contract Number: HHSP-2332-01-5001-33I

Prepared by

James Bell Associates
2000 15th Street North, Suite 100
Arlington, VA 22201
(703) 528-3230
www.jbassoc.com
Mariel Sparr, Project Director

Disclaimer

Acknowledgments

We are grateful to our Measure Testing Workgroup members for their collaboration and thought partnership in the development and testing of the SuPRS–HV: Erica Bluford Ellerbe, David Comer, Justin Fogt, Jaclyn Gray, Crystal Munoz, Dawn Nixon, Bijoux Pighee, David Schultz, Sherryl Scott Heller, Heather Smith, Angela M. Tomlin, and C. Edward Watkins, Jr. The views expressed in this publication do not necessarily reflect the views of these members.

We are also grateful to the following models that provided feedback on the SuPRS–HV throughout measure development and testing: Child First, Healthy Families America, Nurse-Family Partnership, Nurses for Newborns, Parents as Teachers, ParentChild+, and Promoting First Relationships.

We would also like to note the valuable contributions of Nicole Denmark, Shirley Adelstein, Farha Marfani, and Kerry Cassidy Norton throughout the measure development and testing process. Additionally, we appreciate Alex Joraanstad, Nancy Whitesell, and Emily Heberlein for their support of the project, Loren Masters for data preparation and analysis, and Kris Neurauter for formatting. We would also like to acknowledge Marissa Strassberger and Patrice Cachat for their facilitation and analysis of focus groups with home visiting supervisors. Finally, the SuPRS–HV would not be possible without the many contributions of the home visiting workforce in sharing their experiences, perspectives, and time.

Contents

Exhibits

Overview

Introduction

Home visitors enter the homes and lives of families to provide valuable support as families navigate pregnancy and parenting. Being a home visitor brings unique stressors, challenges, and emotions that may be mitigated through workplace support, such as reflective supervision. Reflective supervision is a widely recommended strategy, however, there is limited research on its impacts in home visiting or other contexts. This lack of research may be due in part to limitations in available measures of reflective supervision. This project sought to fill this measurement gap by developing and validating a measure of reflective supervision practices in home visiting.

Purpose

This manual introduces the Supervisor Practices in Reflective Supervision—Home Visiting (SuPRS–HV), a measure of reflective supervision for the home visiting workforce. Measure development was part of a larger project, Supporting and Strengthening the Home Visiting (SAS-HV) Workforce, which explored how reflective supervision is defined, measured, supported, and associated with program outcomes in home visiting and related fields.

Based on gaps identified in earlier project phases, the SAS-HV team set out to create a feasible, acceptable, relevant, and reliable measure to support research and evidence building on reflective supervision in the home visiting context. To do so, they engaged individuals with technical and practice expertise across four measure development phases.

This manual details the measure development and testing process. It includes findings related to the structure of the SuPRS–HV, preliminary evidence of concurrent validity, supervisor experiences using the measure, and guidance on using the SuPRS–HV for research and practice.

Key Findings and Highlights

The SuPRS–HV is a short, self-report measure of specific practices used and content discussed in a single reflective supervision session. Through an iterative process of development with practitioners, researchers, and home visiting model representatives and large-scale testing, the SuPRS–HV demonstrates strong content validity. Specifically:

- Large-scale testing of the SuPRS–HV with over 500 home visiting supervisors identified five subscales—Responsiveness, Collaborative Capacity Building, Unconditional Positive Regard, Promoting Awareness of Self and Others, and Supportive Feedback. Eleven items document topics and content discussed during a reflective supervision session.

- The five subscales showed adequate variability, (i.e., scores vary across sessions and supervisors), with Responsiveness and Unconditional Positive Regard having slightly higher averages.
- We found preliminary evidence of concurrent validity (i.e., agreement with a validated measure used at the same time).
- Supervisors reported positive experiences using the SuPRS–HV, suggesting the measure is feasible and relevant for the home visiting context.

The SuPRS–HV is a publicly available measure and shows promise for use in both home visiting research and practice.

Methods

The research team developed the SuPRS–HV across four phases:

1. During the **Conceptualize** phase, the team developed and refined a list of key elements of reflective supervision.
2. As part of the **Operationalize** phase, the team decided the measure would assess a single supervision session, focus on specific behaviors and techniques, and function as a self-report for supervisors to complete. The researchers also developed preliminary items for measuring key elements of reflective supervision.
3. Activities conducted during the **Preliminary Test** phase included pilot testing (n = 41) and focus groups with supervisors to assess clarity, acceptability, and perceived value of the measure.
4. For the **Large-Scale Test** phase, the team surveyed 502 home visiting supervisors and conducted a repeated measures survey with a subsample of respondents (n = 39).

The SAS-HV team actively engaged people with practical experience and expertise throughout each phase. The group providing input included home visiting model representatives, home visitors, supervisors, reflective supervision trainers, reflective supervision researchers, and measure developers.

Introduction

Supervisor Practices in Reflective Supervision—Home Visiting (SuPRS–HV) is a measure of reflective supervision for the home visiting context. SuPRS–HV measures specific reflective supervision practices and content addressed in a single supervision session, as reported by supervisors. The Supporting and Strengthening the Home Visiting Workforce (SAS-HV) project team[1] (we) engaged in a multiyear process to develop, collect preliminary data, and examine the structure and validity of the measure through large-scale testing. We engaged individuals with technical and practice expertise throughout all phases of development and testing. Throughout the manual, we summarize active engagement touchpoints for each phase in sidebars and share supervisors' experiences and perceptions of the measure.

This manual provides information about—

1. SuPRS–HV subscales and items
2. Measure development process
3. Measure performance
4. Recommendations for using the measure in research and practice

The appendices present additional technical information and details about the SuPRS–HV.

Active Engagement

Throughout measure development and testing, we engaged multiple groups to help us better understand key elements and functions of reflective supervision in the home visiting context.

- **Technical workgroup**: Seven individuals with expertise in reflective supervision and measure development, including researchers, evaluators, trainers, and psychometricians.
- **Practitioner workgroup**: Nine individuals working in local home visiting programs, including home visitors, supervisors, and program managers.
- **Home visiting model representatives**: Representatives from seven evidence-based home visiting models.
- **Measure testing workgroup**: Seven local home visiting program staff (home visitors and supervisors) and five individuals with expertise in reflective supervision (researchers, trainers).

[1] The SAS-HV project team is comprised of James Bell Associates, Johns Hopkins University, and the University of Colorado Denver.

Rationale for the SuPRS–HV

Reflective supervision is a workforce support strategy that can help mitigate workplace stressors and challenges, support home visitor well-being and reflective capacity, and improve the quality and nature of work with families (Gilkerson, 2004; Shamoon-Shanok, 2009; Van Horn, 2018). Reflective supervision commonly occurs between supervisors and home visitors in individual or group settings and provides space and support for home visitors to reflect on their thoughts and feelings about their work with families. Reflective supervision is widely recommended and endorsed across early childhood home visiting models and programs. However, there are no valid measures of reflective supervision for the home visiting context.

Existing measures of reflective supervision, within and outside of the home visiting context, capture practices used and content addressed across supervision sessions rather than practices used within a single supervision session. This makes it difficult to link specific practices with outcomes and introduces recall biases. Reflective supervision practices may vary substantially between sessions and home visitors. Multiple reports may be needed to accurately capture this variability.

We developed the SuPRS–HV to address these gaps in measurement. SuPRS–HV assesses how supervisors deliver a single reflective supervision session to home visitors, including specific reflective supervision practices and topics discussed. Our goal is to create a feasible, acceptable, relevant, and reliable measure to support the home visiting field. We discuss the four phases of SuPRS–HV measure development and testing in the Measure Development Process section below.

See the SAS-HV Reflective Supervision report for more background information on reflective supervision in home visiting.

About the SuPRS–HV

The SuPRS–HV is a self-report measure to be completed by any supervisor, consultant, or manager who provides direct, one-on-one reflective supervision to home visitors. It measures content discussed and specific supervision practices during a single reflective supervision session but does not currently measure supervision quality. It includes a total of 37 items. Eleven items document the topics discussed, and 26 measure specific reflective supervision practices. The SuPRS–HV takes approximately 10–15 minutes to complete. See appendix B for a full version of the measure.

Topics Discussed

To fill gaps in knowledge surrounding the content discussed during reflective supervision sessions in home visiting and the potential impact on hypothesized outcomes, the SuPRS–HV includes 11 items to document topics discussed including:

- Impact of a situation on the baby or child
- Caregiver and child interactions
- Strengths of the families the home visitor is working with
- Needs of the families the home visitor is working with
- Setting and maintaining healthy boundaries with the families the home visitor is working with
- Impact of program requirements, logistics, or other aspects of the work environment on the home visitor's work with families
- Impact of community characteristics (e.g., availability of safe housing, employment, or affordable childcare) on the home visitor's work with families
- Impact of working with families on the home visitor's well-being
- What is going well in the supervisory relationship
- Encouraging the home visitor to consider how work-related stressors may affect their personal life
- An opportunity to discuss the home visitor's professional development goals

Reflective Supervision Practices

The SuPRS–HV includes five subscales to document specific practices that support key elements of reflective supervision in the home visiting context (see exhibit 1). Each subscale includes three to seven items. We referenced existing conceptual and theoretical models of reflective supervision followed by rigorous testing to identify subscales. See exhibit 3 and the Measure Development section for more detailed information on this process. Items by subscale are outlined in exhibit 2.

Exhibit 1. SuPRS–HV Subscales

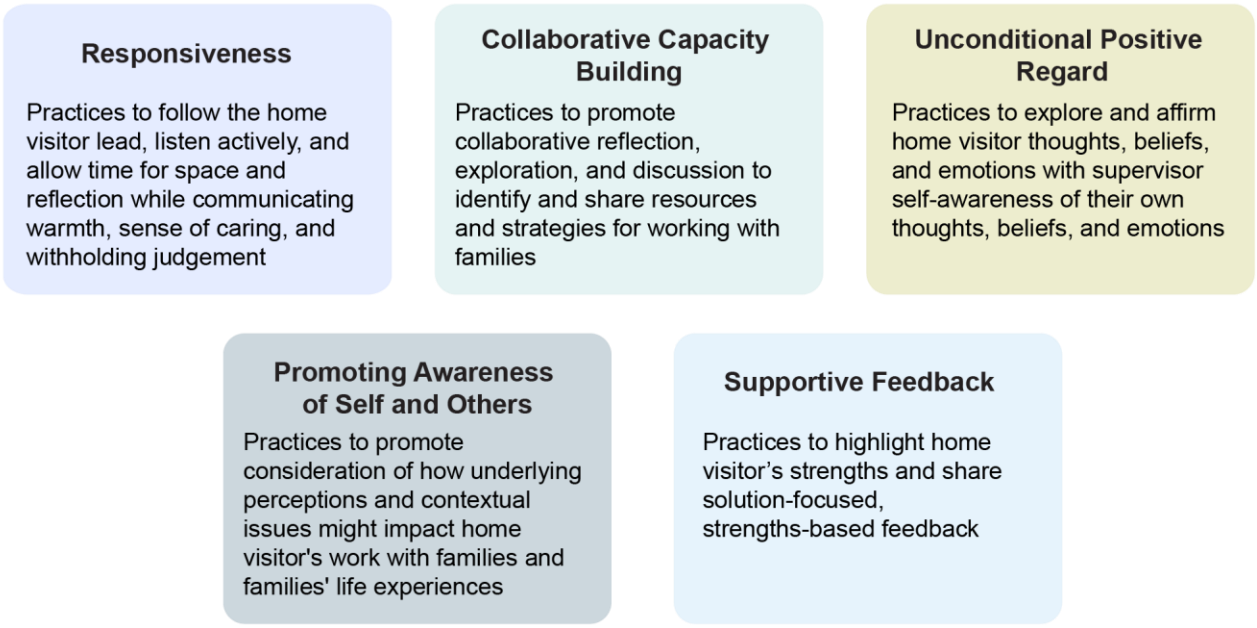


Exhibit 2. SuPRS–HV Subscales and Items

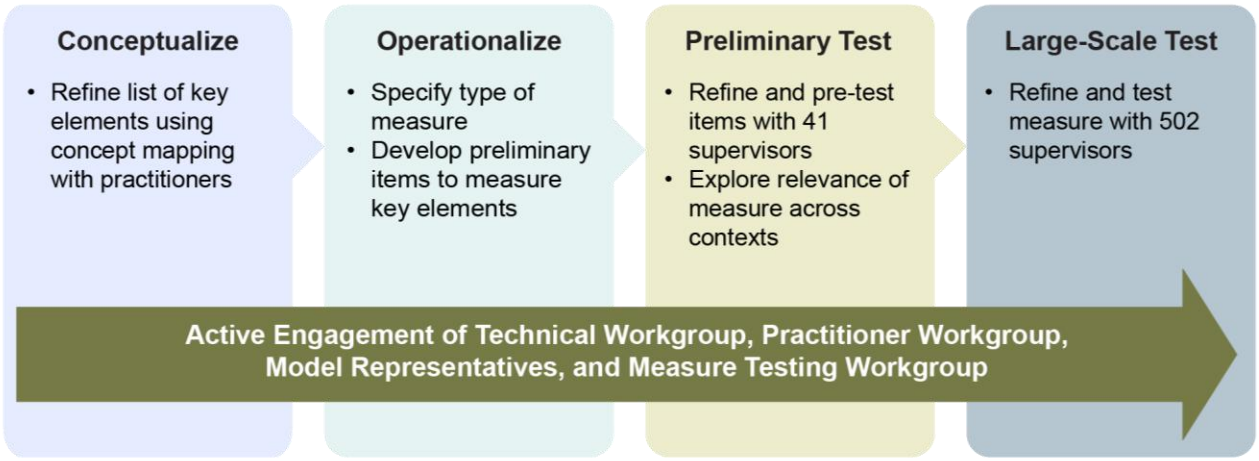
SuPRS–HV subscale	SuPRS–HV item
Responsiveness	How often, if at all, did you: • Wait for the home visitor to gather their thoughts • Let the home visitor know you heard and understood their feelings • Respond to what the home visitor shared without judgement • Offer support to promote the home visitor's well-being
Collaborative Capacity Building	How often, if at all, did you: • Explore solutions together • Share specific skills with the home visitor • Demonstrate a potential strategy to use with a family • Help the home visitor process what is going on with families they find challenging to work with • Provide concrete resources the home visitor can use in their work with families • Provide resources to support the home visitor's professional development

SuPRS–HV subscale	SuPRS–HV item
Unconditional Positive Regard	How often, if at all, did you show: • Interest in the home visitor's views • Appreciation for the home visitor's ideas • Acceptance of the home visitor as a person regardless of their behaviors or ideas How often, if at all, were you: • Authentic in your interactions with the home visitor • Comfortable discussing sensitive topics • Aware of how your emotions impacted the session
Promoting Awareness of Self and Others	How often, if at all, did you: • Ask the home visitor to describe and explore their feelings about experiences with families How often, if at all, did you encourage the home visitor to consider how the following might impact their work with families? • Home visitor's personal identities (e.g., race, ethnicity, or culture) • Home visitor's assumptions or beliefs about a family based on family characteristics (e.g., race, ethnicity, or culture) • Home visitor's experiences How often, if at all, did you encourage the home visitor to consider: • How a family's culture might shape a family's life experience • How a family's racial or ethnic identity might shape a family's life experience • The perspectives of people involved in a situation (e.g., family members, coworkers, people important to a family)
Supportive Feedback	How often, if at all, did you share: • A specific strength of the home visitor • Specific, positive feedback on something the home visitor did • Feedback on what the home visitor was doing well before offering a suggestion for improvement

Measure Development Process

The measure development process consisted of four phases (see exhibit 3). Each phase incorporated active engagement through ongoing collaboration with home visiting model representatives, home visitors, supervisors, reflective supervision trainers, researchers, and measure developers. Below, we describe each phase of measure development.

Exhibit 3. Measure Development Phases



Phase One: Conceptualize

The goal of Phase One was to identify the key elements of reflective supervision to be measured. We developed a list of key elements that aligned with the SAS-HV conceptual model developed in an earlier phase of the project (exhibit 4). The conceptual model was developed to address gaps in existing literature and advance understandings of reflective supervision in the home visiting context by illustrating how key elements contribute to staff, program, and family level outcomes. The conceptual model was informed by a literature review, environmental scan, and active engagement. To ensure a comprehensive list of elements, we also reviewed and extracted elements from existing measures. Additional details on key elements of reflective supervision can be found in the SAS-HV Reflective Supervision report.

Phase One Active Engagement

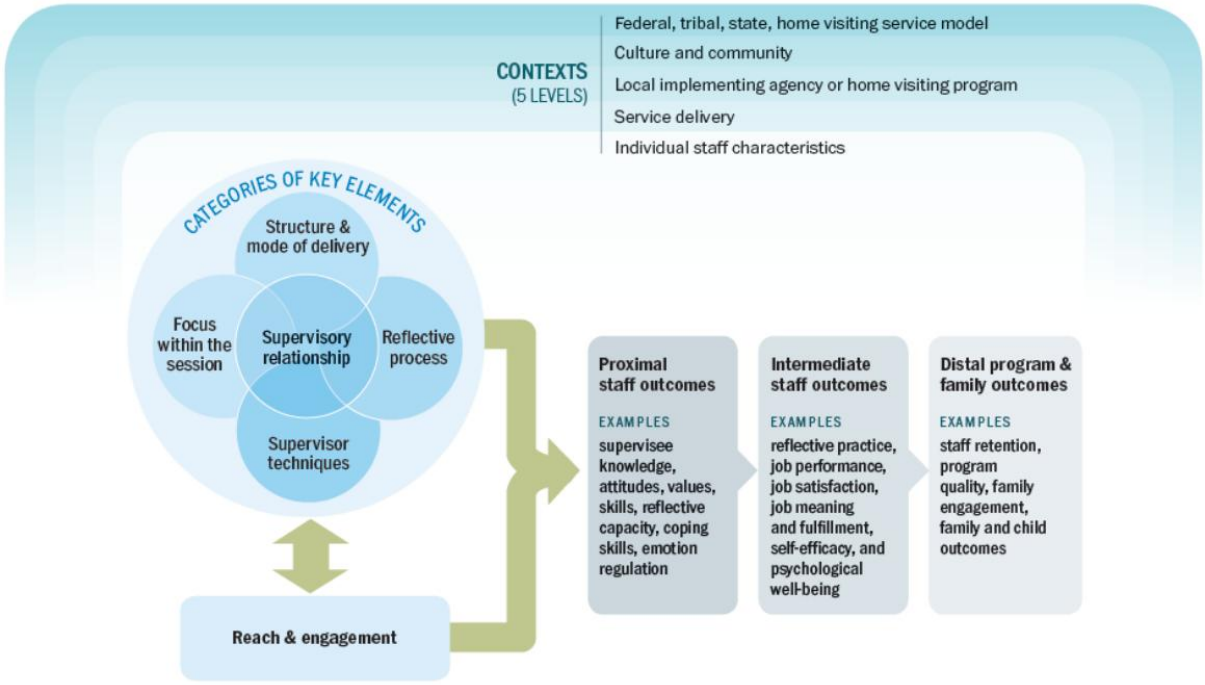
Discussions with technical and practitioner workgroup members about—

- Elements of reflective supervision to include in the measure
- Clarity of definitions for key elements
- Feedback on ratings, element groupings, and names of groupings from concept mapping activities

Discussions with model representatives about—

- Model perspectives on key elements of reflective supervision
- Clarity of definitions for key elements

Exhibit 4. SAS-HV Conceptual Model of Reflective Supervision



A note on terminology

Element: Refers to a broader goal or concept of reflective supervision in home visiting. Multiple items may be required to assess a single element.

Item: Refers to a statement that operationalizes an element in specific, measurable terms.

We sought feedback from a technical workgroup of researchers and evaluators, a practitioner workgroup of home visitors and supervisors, and model representatives to refine the list of elements (see call out box). In general, technical and practitioner workgroup members thought the list of key elements was inclusive and thorough. Both groups suggested ways to clarify some elements, such as those related to trust and safety, and responsive techniques. In response to feedback, we revised some element definitions, removed some, and added new ones. For example, we added elements focused on racial and ethnic awareness and responsiveness to signs of home visitor trauma.

We also sought feedback on the elements from nine home visiting models that require reflective supervision, based on model descriptions on the Home Visiting Evidence of Effectiveness (HomVEE) website.[2] We asked the model representatives whether their model requires, recommends, or has no position on each key element and whether any important elements were

[2] HomVEE is a federally sponsored initiative that reviews research to provide an assessment of the evidence of effectiveness for early childhood home visiting models that serve families with pregnant women and children from birth to kindergarten entry.

missing. Responses were received from eight models; each key element was required by at least one model.

Finally, we obtained the perspectives of a larger group of supervisors, home visitors, reflective supervision researchers, and measure developers using concept mapping methodology to gain deeper insights into how they understand and conceptualize reflective supervision. Concept mapping offered a structured approach for collecting, analyzing, and visually organizing views on the list of key elements. Concept mapping participants completed three activities: (1) sorting the key elements into groupings that made sense to them, (2) rating the key elements based on how essential they are for reflective supervision in home visiting, and (3) discussing the results together in group interpretation sessions. We used findings from concept mapping activities to further refine, define, and organize elements in meaningful groups. Detailed methods and findings from concept mapping activities are provided in the SAS-HV concept mapping brief.

Phase Two: Operationalize

The goal of Phase Two was to select the type of measure to develop and create items to measure the key elements of reflective supervision identified in Phase One. We used knowledge gained from earlier project phases to make three key decisions about the measure: The measure will (1) assess a single supervision session, (2) focus on specific behaviors and techniques, and (3) be a self-report measure completed by the supervisor. The decision to assess a single supervision session and specific behaviors and techniques was made to minimize recall bias; offer the level of precision needed to capture variability in reflective supervisory practices; and promote evidence building by linking use of specific practices with anticipated outcomes. We chose to develop a supervisor self-report measure because supervisors' perspectives regarding specific practices are likely more accurate than home visitors, who may not be aware of the full range of practices a supervisor uses within a session. Additionally, reflective supervision training and professional development activities most commonly target supervisors. While also having home visitor reports on practices would be ideal, we decided this could best be accomplished in a subsequent project.

To develop items, we—

- Drafted an item pool.
- Developed scale of measurement.
- Gathered input from technical and practitioner workgroup members.
- Gathered input from model representatives.
- Explored differences in perspectives.

***Drafted item pool*.**

1. We drafted an initial list of items that aligned conceptually with each element identified in Phase One. Each item was selected with the goal of describing specific, discrete behavioral indicators of a key element. We also examined items within existing measures and consulted with experts in measure development.
2. We drafted multiple items for each element. For example, items for the element Responsiveness assessed practices include allowing space for the home visitor to gather their thoughts, responding to signs of home visitor trauma or burnout, and acknowledging and affirming the home visitor's feelings. At least two project team members reviewed and rated each draft item independently for clarity and alignment with best practices in item development (DeVellis & Thorpe, 2021).
3. The team then used our ratings and open comments to either revise or remove poorly worded, unclear, or unaligned items.

Developed scale of measurement. We then determined the most useful scale of measurement and response options for each item. As a first step, we reviewed response options in similar measures of supervisor and therapist strategies, techniques, and behaviors. We considered binary (e.g., yes, no), ordinal (e.g., item is not in place/occurs infrequently to item is in place/occurs consistently) and continuous rating scales. Measure development and content experts as well as practitioners helped to refine item response options.

***Gathered input from technical and practitioner workgroup*.** Technical and practitioner workgroup members helped to identify unclear or overlapping items and to indicate which items are most central to a given element via written feedback and small group discussions. We used this feedback, along with findings from concept mapping activities, to refine the item pool. We considered the extent to which each item was clear, did not overlap with other items, and measured an element deemed essential. Four team members independently reviewed this information and feedback for each item to make a recommendation of whether to revise, remove, or retain items. The project team met to review individual team member recommendations, resolve areas of disagreement, and flag items with outstanding questions to be addressed in Phase Three.

Phase Two Active Engagement

Discussions with technical and practitioner workgroup members about—

- Item revisions to improve clarity, reduce duplication, and capture essence of key elements

Written feedback and follow-up discussion with six home visiting model representatives about—

- Model position on items
- Item revisions to improve clarity, reduce duplication, and capture essence of key elements.

Gathered input from home visiting models. We then asked model representatives engaged in Phase One for their input on the clarity of the items and whether they capture aspects of reflective supervision required or recommended by their models. Seven models provided feedback on the draft items. Most models reported that their model either requires or recommends most of the items.

Explored differences in perspectives. Across activities in Phase One and Phase Two, we noted some differing views between engagement groups. For example, whereas home visiting staff endorsed the use of reflective supervision practices related to capacity building (e.g., suggesting specific strategies, offering concrete resources and materials to support home visitors' work with families, discussing home visitors' professional development needs), most researchers, evaluators, and model representatives thought of these practices as helpful but not central to reflective supervision. To enable empirical testing of these practices in relation to the intended outcomes of reflective supervision, we retained capacity building items.

Phase Three: Preliminary Test

The goal of Phase Three was to gather preliminary data on the clarity, feasibility, and appropriateness of the measure and examine whether items functioned as intended in preparation for large scale testing. To do this, we—

- Conducted cognitive interviews.
- Pilot tested the measure with a small sample of supervisors.
- Conducted focus groups to explore relevance of the measure across subgroups.
- Used findings to refine items.

Phase Three Active Engagement

Discussion with measure testing workgroup about the implications of the focus group findings and which findings most resonated with their experience.

***Conducted cognitive interviews*.** Using the item pool and response options developed in Phase Two, we completed cognitive interviews with 7 supervisors; preliminary pilot testing with a sample of 41 supervisors; and focus groups with supervisors to assess clarity, acceptability, and perceived value of the measure. The cognitive interviews helped identify items in need of refinement regarding interpretability and language.

Pilot testing with a small sample. Pilot testing helped ensure the items functioned as intended. We recruited 56 supervisors who were eligible to participate in pilot testing. Of those, 41 participated in pilot testing by completing the measure. We calculated basic descriptive statistics to examine central tendency and variability.

Conducted focus groups. Focus groups explored the relevance of the measure across supervisors with varying racial and ethnic identities, including supervisors who identify as Black or African American, Hispanic or Latino, or American Indian or Alaskan Native, to explore the relevance and importance of reflective supervision and perceptions of the draft measure.

Refined items. Results of pilot testing informed revisions to the measure prior to large scale testing. Cognitive interviews and pilot testing led to minor rewording and removal of items that were not clear or demonstrated limited variability. For example, we removed two items (How often, if at all, did you provide— "examples of potential strategies for working with a family" and "model techniques to use when working with families") that demonstrated limited variability and overlapped with other items. Focus group findings informed revision of several items, particularly those related to culture, unconscious assumptions, and social positioning, to use more precise words and terminology used by the supervisors in practice.

Phase Four: Large-Scale Test

The goal of Phase Four was to assess item performance, factor structure, internal consistency, concurrent validity, reliability, feasibility, and acceptability. We—

- Recruited over 500 supervisors to complete the SuPRS–HV and related measures.
- Invited a subgroup to complete the SuPRS–HV over multiple supervision sessions with 2–3 home visitors.
- Invited home visitors to complete validity measures of their experiences of reflective supervision.

Large-scale survey of supervisors. We recruited a large sample of supervisors to complete a survey immediately after a one-on-one reflective supervision session with one home visitor. The survey included the SuPRS–HV; items to assess supervisor demographic characteristics, roles, and experience; feedback on their experience completing the SuPRS–HV; and three measures to assess concurrent validity (discussed further in the Concurrent Validity Analyses section below). The survey took approximately 1 hour to complete. We recruited supervisors through home visiting conferences and listservs, informational webinars, and outreach to model representatives. A total of 502 home visiting supervisors representing 23 home visiting models; 51 states and territories; and urban,

Phase Four Active Engagement

Feedback meetings from measure testing workgroup

- Following large scale testing, discussed preliminary findings from factor analysis and the group's interpretation of dimensions.
- Following repeated measures testing, the group helped name subscales and revise item groupings.

rural, and suburban communities participated in the survey. See below and appendix A for details on the sample, analytic approach, and results.

Repeated surveys of supervisors. We recruited supervisors to participate in repeated measures testing from a pool of supervisors in the large-scale sample. We selected supervisors using stratified random sampling (based on supervisor race/ethnicity and model[s] implemented) to maintain sample characteristics similar to the large-scale sample. A total of 39 supervisors participated in the repeated measures survey. Supervisors completed the SuPRS–HV following three reflective supervision sessions with two or three home visitors (i.e., each supervisor completed the SuPRS–HV up to nine times, once after three reflective supervision sessions with up to three different home visitors).

A note on terminology

Content validity: The degree to which a measure captures all aspects of what it intends to measure.

Concurrent validity: The degree to which scores on the measure relate to scores on existing measures of the same or similar concepts, administered at the same time.

Home visitor surveys. Once supervisors were recruited to participate in the repeated measures survey, they were asked to identify up to three home visitors to whom they would provide weekly or biweekly, individual reflective supervision. Supervisors then obtained their consent to participate in the study. One hundred eleven home visitors were recruited and 110 completed surveys describing their experience with reflective supervision.

Supervisor Experiences Using the SuPRS-HV

Supervisors shared their experiences using the SuPRS-HV in the developmental survey and repeated measures. Example quotes from supervisors are below.

"My first concern was, this is going to be super repetitive and going to drive me crazy, but it didn't. It allowed me to see each one of my staff through different lenses."

"I just caught myself thinking about the questions regarding personal identity and beliefs because it was consistently, 'No we didn't talk about that, we didn't talk about that'. It's a good reminder for me to think more critically, 'are these things coming up and I'm just missing it', or did it honestly just not come up because it didn't pertain to the conversation?"

"The questions targeted all aspects of reflective supervision and highlighted all that is going on, intentionally and naturally during a supervision session"

"This is a very interesting measure, and the questions really allowed me to reflect on what I did or didn't touch on or do in my supervision session. I enjoyed this!"

"I think this is a great measure. It really made me think about some things to put into perspective. I want my approach to be authentic and be available for any situation for my parent educators."

"It was very thorough but also helped me identify or think about areas I may work on using more during supervision. I could see it being helpful in planning for future reflective supervision sessions."

"It helped remind me of topics to go over with them as time went on… Just having that reminder as far as the relationship… I didn't really have the same growth with the other home visitors that I wasn't doing the measure with."

We gratefully acknowledge the supervisors who assisted in the development and testing of the SuPRS-HV and who shared their experiences.

Measure Performance

Below, we provide results from the large-scale test of the measure, starting with a description of the large-scale and repeated measures samples. This is followed by a summary of findings, including results of exploratory factor analysis (EFA), confirmatory factor analysis (CFA), descriptive statistics, internal consistency, variability in scores across sessions and home visitors, and concurrent validity. See appendix A for more details, including descriptive statistics on questions assessing participant experiences with the SuPRS–HV.

Sample

As described below, the sample includes 502 supervisors who participated in the large-scale test and a subgroup of 39 supervisors who also participated in repeated administration of SuPRS–HV.

Large-Scale Sample

The large-scale sample (sample) included a total of 502 supervisors. Supervisors represented home visiting programs in 51 states and territories. The highest number of respondents were from Illinois (11 percent), California (6 percent), and Alabama (6 percent). As reported by supervisors, participants worked at home visiting programs supporting families in urban (50 percent), suburban (49 percent), rural (66 percent), and frontier (3 percent) settings. Sixteen percent of supervisors reported they worked for programs serving Indigenous communities. Home visiting programs implemented a total of 24 different home visiting models.[3] The most common models were—

- Parents as Teachers (35 percent)
- Healthy Families America (28 percent)
- Nurse-Family Partnership (11 percent)
- Early Head Start Home-Based Option (7 percent)

Home visiting programs varied in size, though over half (58 percent) served at least 75 families (exhibit 5).

[3] Models included Attachment and Biobehavioral Catch Up (ABC), Baby TALK, Child First, Early Head Start Home-Based Option, Early Steps to School Success (ESSS), Family Check Up, Family Connects, Family Spirit, Following Baby Back Home (FBBH), Health Access Nurturing Developmental Services (HANDS), Healthy Families America, High Risk Perinatal Assistance Program, Home Instruction for Parents of Preschool Youngsters (HIPPY), Maternal Early Childhood Sustained Home Visiting (MECSH), Maternal Infant Health Outreach Worker (MIHOW), Nurse Family Partnership, Parents as Teachers, Parent-Child Assistance Program (PCAP), ParentChild+, Play and Learning Strategies (PALS), Promoting First Relationships, SafeCare Augmented, Team for Infants Exposed to Substance Use (TIES), and Welcome Baby.

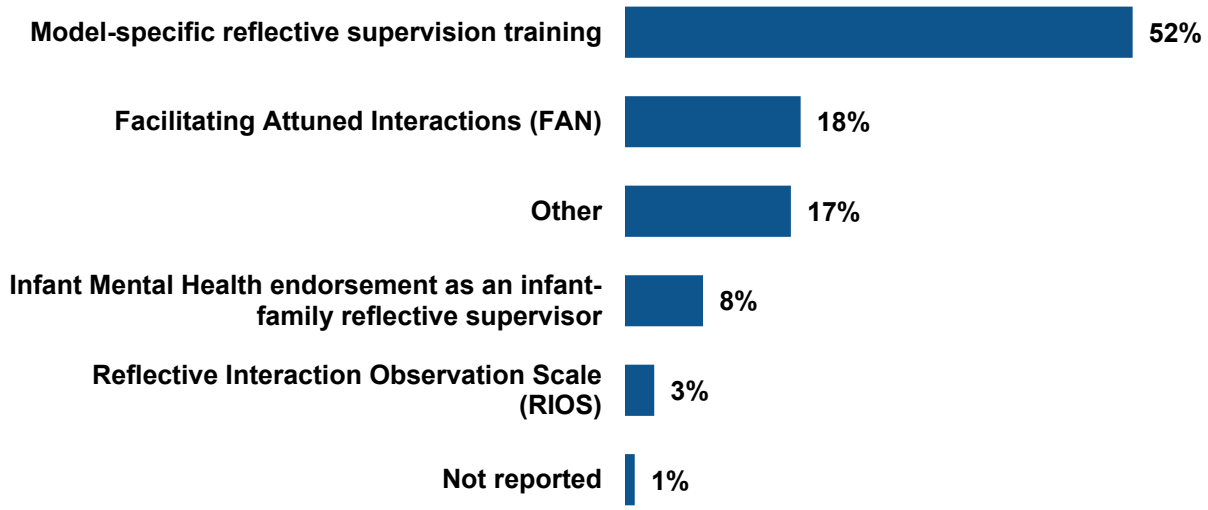
Exhibit 5. Number of Families Served by Home Visiting Programs of Large-Scale Sample Respondents

Number of families	*n*	%
Fewer than 25	34	7
Between 25 and 50	103	21
Between 51 and 74	74	15
Between 75 and 99	84	17
More than 100	207	41

Most supervisors (61 percent) received and provided reflective supervision, while 39 percent only provided it. They supervised varying numbers of home visitors: 5 percent only supervised one, 14 percent supervised two, 17 percent supervised three, and nearly two-thirds (64 percent) supervised four or more. Supervisors provided in-person, individual (95 percent); in-person, group (37 percent); virtual, individual (54 percent); and virtual, group (19 percent) supervision.

Most supervisors (88 percent) had received reflective supervision training or professional development. Of those who received training, model-specific training was the most common type of training received (52 percent; exhibit 6). Finally, 64 percent of supervisors also provided reflective supervision training, coaching, and/or support to others.

Exhibit 6. Type of Reflective Supervision Training Received by Large-Scale Sample Respondents



Note: *N* = 439 supervisors who reported having received reflective supervision training or professional development.

Roughly two-thirds of participating supervisors were White (exhibit 7). A majority (97 percent) of supervisors spoke English as their primary language. Supervisors reported high levels of education; nearly half (48 percent) had a bachelor's degree and 40 percent had a graduate degree (exhibit 8). The top three most common fields of study were—

- Education (early education/special education) (25 percent)
- Nursing (17 percent)
- Social work (16 percent)

The sample was also experienced in providing both home visiting services (36 percent with more than 10 years of experience) and reflective supervision (36 percent with more than 5 years of experience). Most (86 percent) reported at least 1 year of experience providing reflective supervision (exhibit 10).

Exhibit 7. Large-Scale Sample Supervisor Race/Ethnicity

Race/ethnicity	*n*	%
American Indian or Alaska Native	8	2
Asian	7	1
Black or African American	63	13
Latino or Hispanic	74	15
Native Hawaiian or Pacific Islander	2	<1
White	317	63
Multiple races/ethnicities	24	5
Other	1	<1
Prefer not to answer	6	1

Note: Percentages exceed 100 because multiple selections were allowed.

Exhibit 8. Large-Scale Sample Supervisor Education

Highest level of education	*n*	%
High school diploma or equivalent	5	1
Some college or technical/training program	22	4
Associate's degree	32	6
Bachelor's degree	241	48
Graduate degree	202	40

Exhibit 9. Large-Scale Sample Supervisor Experience in Home Visiting

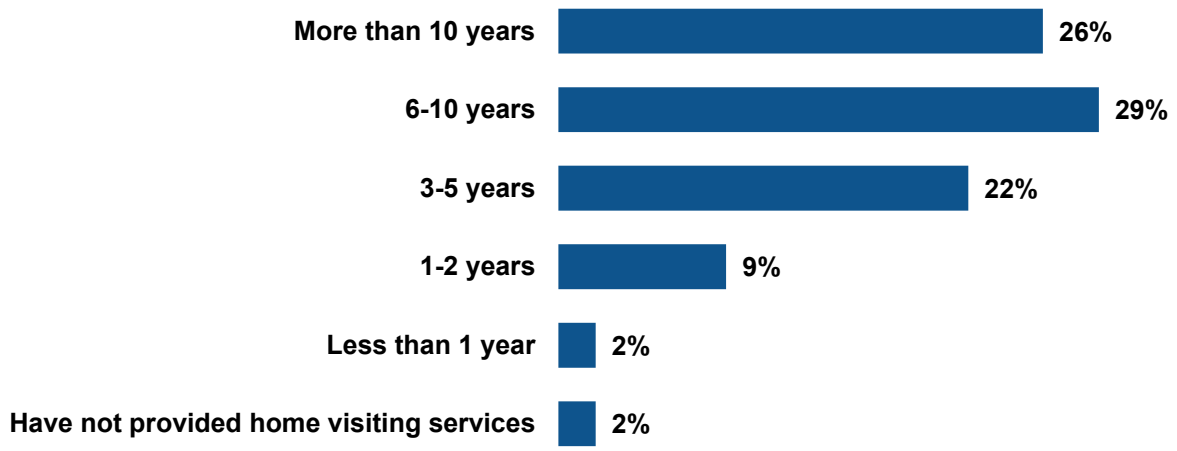
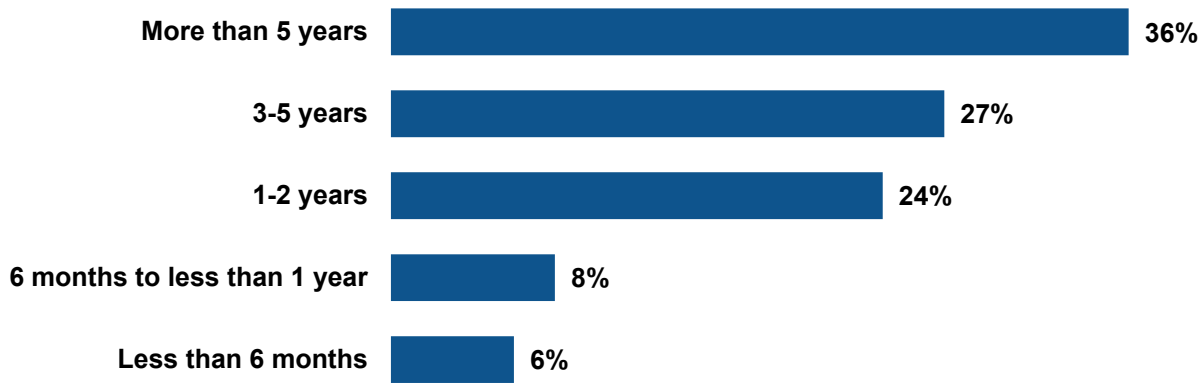


Exhibit 10. Large-Scale Sample Supervisor Experience Providing Reflective Supervision



Repeated Measures Sample

We collected repeated measures data from a subsample (*N* = 39) of supervisors to examine variability in SuPRS–HV scores across supervision sessions and home visitors. Supervisors completed the SuPRS–HV after each of three reflective supervision sessions with up to three home visitors (*N* = 111 home visitors). This resulted in a total of 333 observations over time. See appendix A for details on the characteristics of this subsample.

Findings

Below we provide findings from the large-scale test and repeated measures administration. First, we describe the EFA and CFA, then descriptive statistics, information about internal consistency, variability in SuPRS–HV scores across supervision sessions and home visitors, followed by results from concurrent validity testing.

Exploratory Factor Analysis

We used a random 60 percent of the large-scale sample (*N* = 293) for EFA that examined whether items on the SuPRS–HV represent distinct subscales of reflective supervision techniques and practices. We began analyses using all 60 items on the draft measure and iteratively eliminated 25 that did not meet established criteria for inclusion (i.e., failed to align with any subscale; aligned with more than one subscale; and/or were deemed redundant, irrelevant, or otherwise not useful by the study team and expert advisors). The final EFA showed that the SuPRS–HV was organized into five subscales of items. Of note, the 11 items that document topics discussed during a reflective supervision session were among those eliminated based on EFA results. Although those 11 items were not included as part of the five SuPRS–HV subscales, they were retained on the SuPRS–HV because of their importance to research on the content of reflective supervision sessions in the home visiting context and their value to supervisors in tracking the topics they address with home visitors over time.

Confirmatory Factor Analysis

The SuPRS–HV contains 11 items to assess topics discussed and 26 items to assess supervisory practices, organized in five subscales. These five subscales were derived using CFA, following the EFA, which is summarized above and described in more detail in appendix A. We conducted the CFA using an independent random sample of approximately 40 percent of the 502 cases in the large-scale sample (*N* = 209). The factor model tested in the CFA was based on the five-factor EFA, in line with best practice for integrating expert judgement and knowledge with data-driven solutions (Brownstein et al., 2019). The following four nonloading items identified as critical to reflective

supervision practice by the study team and individuals consulted through active engagement were added:

1. How often, if at all, did you encourage the home visitor to consider how the home visitor's experiences might impact their work with families?
2. How often, if at all, did you ask the home visitor to describe and explore their feelings about experiences with families?
3. How often, if at all, did you help the home visitor process what is going on with families they find challenging to work with?
4. How often, if at all, did you show appreciation for the home visitor's ideas?

The modified five-factor solution demonstrated adequate fit ($\chi 2(289) = 680.93$, $p < .001$; RMSEA = 0.08; CFI = 0.94; and SRMR = 0.08) in the independent random CFA sample, and no modification indices suggested alternative specifications. Appendix A contains additional details about the CFA.

Descriptive Statistics and Internal Consistency

Exhibit 11 provides descriptive statistics, internal consistency coefficients, and the number of items for each of the five SuPRS–HV subscales measuring supervisory practices, using data from the full large-scale sample (N = 502). Scores on the five subscales were largely normally distributed and showed adequate variability. Exceptions included *Responsiveness* and *Unconditional Positive Regard,* which had somewhat higher averages (Ms = 3.2 and 3.4, respectively) and were slightly negatively skewed (sk = -0.21 and -0.76, respectively) relative to the other three subscales. Analysis of internal consistency reliability showed that Cronbach's alpha coefficients (0.75–0.87) of the five subscales specified in the CFA (exhibit 11) were in the good to excellent range (DeVellis & Thorpe, 2021; Xiao & Hau, 2023). Inspection of alpha-if-item-deleted statistics identified no items that adversely impacted internal consistency or were redundant or unnecessary (Tavakol & Dennick, 2011).

Exhibit 11. Descriptive Statistics and Internal Consistency Coefficients of Subscales from Confirmatory Factor Analysis

Subscale	Mean	Standard deviation	Cronbach's alpha	Number of items
Responsiveness	3.21	0.53	0.75	4
Collaborative Capacity Building	2.74	0.60	0.83	6
Unconditional Positive Regard	3.41	0.50	0.83	6
Promoting Awareness of Self and Others	2.46	0.64	0.82	7
Supportive Feedback	2.96	0.74	0.87	3

Note: N = 502. Scores range from 1–4. Scores on *Promoting Awareness of Self and Others* were calculated after converting the three items measured on a 3-point scale to a 4-point scale that was used for all other items on the subscale (IBM, 2020).

Variability Across Sessions and Home Visitors

Using the repeated measures subsample of 39 supervisors, we examined descriptive statistics on each SuPRS–HV subscale at each session across home visitors. Means and standard deviations on each subscale at each time point are shown in exhibit 12. The pattern of means suggests each type of reflective practice may have been used more frequently at each subsequent session. Examination of standard deviations suggests variability among home visiting supervisors in the use of those practices across home visitor supervisees.

Exhibit 12. Variability in SuPRS–HV Scores Across Time

Subscale	Session one	Session two	Session three
Responsiveness	3.18 (0.57)	3.24 (0.58)	3.34 (0.55)
Collaborative Capacity Building	2.51 (0.65)	2.56 (0.69)	2.61 (0.73)
Unconditional Positive Regard	3.49 (0.49)	3.54 (0.48)	3.63 (0.41)
Promoting Awareness of Self and Others	2.20 (0.62)	2.21 (0.64)	2.34 (0.71)
Supportive Feedback	2.70 (0.85)	2.76 (0.80)	2.86 (0.82)

Note: Mean (standard deviation) are displayed. Possible scores range from 1–4.

We also used box and whisker plots to visually explore variability in scores across sessions and home visitors (all plots are shown in appendix A). These plots showed score variability across sessions and home visitors as evident for nearly all supervisors on all subscales. Some supervisors demonstrated greater variability in scores than others. Although variable, scores on the *Responsiveness* and *Unconditional Positive Regard* subscales clustered toward the higher end of the response scale. In contrast, scores on the *Promoting Awareness of Self and Others* subscale clustered toward the lower end of the scale.

Concurrent Validity Analyses

We conducted a series of correlational analyses with data from the full large-scale sample (N = 502) to assess concurrent validity. A post hoc power analysis conducted in GPower v.3.1 (Faul et al., 2009) showed that a sample size of 502 has greater than 80 percent power to detect correlations of a magnitude of 0.30 or greater with confidence level of 95 percent. These analyses examined associations between scores on the five SuPRS–HV subscales and scores on three existing measures that assess constructs related to reflective supervision. The Supervisory Styles Inventory (SSI) is a 25-item measure that assesses different approaches to supervision: attractive, interpersonally sensitive, and task-oriented (Friedlander & Ward, 1984). The Supervisory Working Alliance Inventory-Supervisor Form (SWAI) is a 23-item measure of the working alliance in supervision from the perspective of a supervisor, including subscales for client focus, rapport, and

identification (Efstation et al., 1990). The 18-item Reflective Supervision Self-Assessment Scale (RSS) (Shea et al., 2012; 2016) is a unidimensional measure used to assess supervisor confidence in reflective supervision capacity.

Scores on each subscale of the SuPRS–HV (exhibits 13–15) showed small to moderate positive correlations (Papageorgiou, 2022) with scores on the SSI, SWAI, and RSS. This pattern of findings provides preliminary evidence of the concurrent validity of the SuPRS–HV.

Exhibit 13. Pearson Product Moment Correlations Between SuPRS–HV and the Supervisory Styles Inventory

Subscale	SSI attractive	SSI interpersonally sensitive	SSI task oriented
Responsiveness	0.34	0.39	0.26
Collaborative Capacity Building	0.19	0.34	0.47
Unconditional Positive Regard	0.33	0.35	0.19
Promoting Awareness of Self and Others	0.18	0.36	0.33
Supportive Feedback	0.31	0.37	0.28

Note: All correlations are significant at the $p < .001$ level.

Exhibit 14. Pearson Product Moment Correlations Between SuPRS–HV and the Supervisory Working Alliance Inventory

Subscale	SWAI rapport	SWAI client focused	SWAI identification
Responsiveness	0.26	0.35	0.32
Collaborative Capacity Building	0.13	0.56	0.32
Unconditional Positive Regard	0.34	0.20	0.35
Promoting Awareness of Self and Others	0.17	0.40	0.22
Supportive Feedback	0.18	0.36	0.29

Note: All correlations are significant at the $p < .001$ level.

Exhibit 15. Pearson Product Moment Correlations Between SuPRS–HV and the Reflective Supervision Self-Assessment Scale

Subscale	Reflective supervision self-assessment scale
Responsiveness	0.41
Collaborative Capacity Building	0.38
Unconditional Positive Regard	0.39
Promoting Awareness of Self and Others	0.40
Supportive Feedback	0.37

Note: All correlations are significant at the $p < .001$ level.

Using the repeated measures sample of 39 supervisors, we examined associations between the SuPRS–HV and home visitor reports of the reflective supervision they receive. We used a series of bivariate nested regression analyses predicting scores on each of the home visitor measures from SuPRS–HV subscale scores. Home visitor reports included the Reflective Supervision Rating Scale (RSRS) (Ash, 2010; Gallen et al., 2016), the Supervisory Styles Inventory (SSI) (Friedlander & Ward, 1984), the Short Supervisory Relationship Questionnaire (S-SRQ) (Cliffe et al., 2016), and the Supervisory Satisfaction Questionnaire (Ladany et al., 1996). Results of these analyses are summarized in exhibit 16, with complete results in appendix A. The pattern of findings provides preliminary evidence of the concurrent validity of the SuPRS–HV.

Exhibit 16. Statistically Significant Associations Between SuPRS–HV Subscales and Home Visitor Reports of the Reflective Supervision they Receive

	RSRS mentoring	RSRS supervision structure	SSI task oriented	S-SRQ reflective education	S-SRQ structure
Responsiveness	**0.230**	0.167	0.380	0.313	**0.455**
Collaborative Capacity Building	0.062	0.040	**0.508**	0.136	0.225
Unconditional Positive Regard	**0.317**	**0.320**	0.101	**0.457**	0.331
Supportive Feedback	0.029	0.045	**0.367**	0.062	0.132
Promoting Awareness of Self and Others	0.036	-0.039	0.062	0.072	0.088

Note: Bolded unstandardized regression coefficients for associations statistically significant at $p < .05$ are shown. RSRS = Reflective Supervision Rating Scale, SSI = Supervisory Styles Inventory, S-SRQ = Short Supervisory Relationship Questionnaire.

Using the SuPRS–HV

In this section, considerations for using the SuPRS–HV and instructions for how to access, complete, and score the measure are described. The measure was developed and tested for use by any supervisor, consultant, or manager who provides direct, one-on-one reflective supervision to home visitors. Prior to using the SuPRS–HV—

- Decide whether to use a paper version of the measure or a version programmed into survey software.
- Ensure 10-15 minutes immediately following a reflective supervision session to complete the SuPRS–HV.

The SuPRS–HV takes approximately 10–15 minutes to complete. Supervisors should complete the measure as soon as possible after a reflective supervision session, as recall is most accurate then. This is especially important when supervisors are completing the measure following multiple reflective supervision sessions.

A free PDF version of the SuPRS–HV can be downloaded from the OPRE website. This version can be printed, or supervisors may complete the SuPRS–HV digitally by marking each item score in a PDF editor on the computer. If opting to program the SuPRS–HV into a survey administration software, see appendix B for a version of the measure formatted for electronic surveys (e.g., Qualtrics, REDCap, Survey Monkey) and optimization for mobile devices (Dillman et al., 2014). This version contains three questions per page, repeats instructions before each new section, and contains transition statements. While this is one recommendation for how the SuPRS–HV might look programmed into survey software, the final programming may vary depending on which software is being used, the length of the survey, and the audience.

Scoring Guidance

The SuPRS–HV contains two sections: (1) Topics Discussed and (2) Reflective Supervision Practices. Scoring guidance for each section is below. Scores for the two sections should remain separate and are not meant to be combined as a total score.

Topics Discussed

To address gaps in knowledge surrounding how reflective supervision is implemented in practice and common focal topics, the SuPRS–HV includes 11 items documenting topics discussed during reflective supervision sessions. This allows for a better understanding of topics addressed and an examination of potential associations between topics discussed and anticipated outcomes. However, we currently do not have guidance surrounding how frequently topics should be discussed or at what

depth. Additional testing of these items is necessary before providing specific scoring and interpretation guidance.

To document topics discussed, the items can be scored in two ways, either quantifying the total number of topics discussed or the depth of topics discussed.

1. To explore the total number of topics discussed:
    - Score “not [discussed] in the session” as 0.
    - Score all other response options as 1.
    - Sum the items to create a count of the number of topics discussed.

This scoring method provides the quantity of topics discussed, not the depth of topics discussed. Note: The more topics discussed does not necessarily equate to a better score; the user decides how to interpret the total score.

2. To document the depth of topics discussed:
    - All items must be scored on the same scale (i.e., a scale of 1–3 instead of 1–4). Recode items 10 and 11 by converting scores of 3 to 2.5 and scores of 4 to 3.
    - Calculate averages for items 1–11, which will range from 1–3.

This scoring method allows a comparison of the depth of conversation across all topics discussed, with higher averages indicating more discussion about a topic.

Reflective Supervision Practices

This section contains five subscales that are scored in the same way, as shown below.

Responsiveness

Responsiveness score (Item 12 + Item 13 + Item 14 + Item 15) / 4 =

Collaborative Capacity Building

Collaborative Capacity Building score (Item 16 + Item 17 + Item 18 + Item 19 + Item 20 + Item 21) / 6 =

Unconditional Positive Regard

Unconditional Positive Regard score (Item 22 + Item 23 + Item 24 + Item 25 + Item 26 + Item 27) / 6 =

Promoting Awareness of Self and Others

Recode items 29, 30, and 31 by converting scores of 2 to 2.5 and scores of 3 to 4.

Promoting Awareness of Self and Others score ____ (Item 28) + ____ (Item 29) + ____ (Item 30) + ____ (Item 31) + ____ (Item 32) + ____ (Item 33) + ____ (Item 34) / 7 = ____

Supportive Feedback

Supportive Feedback score ____ (Item 35) + ____ (Item 36) + ____ (Item 37) / 3 = ____

This process will yield five subscale scores ranging from 1–4. A score of 1 indicates the reflective supervision practices in a subscale were not present in the reflective supervision session, while a score of 4 indicates the practices in a subscale were largely present in the session. If desired, a SuPRS–HV total score can be calculated by averaging the five subscale scores.

Please note that higher scores do not necessarily indicate higher quality reflective supervision. Whereas items in the measure were widely endorsed as being essential to reflective supervision, further testing of the measure is needed to understand whether scores are associated with positive outcomes for home visiting programs, home visiting staff, and families. The user decides how to interpret the scores.

Potential Uses

Below are suggestions for using the SuPRS–HV for research and practice.

For Researchers and Evaluators

The SuPRS–HV can be used to describe reflective supervision in home visiting, examine variation in, and explore associations with anticipated outcomes. Examining associations between practices with outcomes such as home visitor professional well-being or retention, job performance and program quality would advance understandings of effective components of reflective supervision to support the home visiting workforce. This information could inform efforts to improve the quality of reflective supervision and train home visiting supervisors. The SuPRS–HV can also be used to examine variation in the use of practices across different home visiting contexts, settings, and samples. Subgroup analysis exploring how training, years of experience, or supervision setting may influence the use of reflective supervision practices would be useful in advancing our understanding of reflective supervision for home visiting.

Specifically, the SuPRS–HV could be used to—

- Assess whether supervisors are using practices that demonstrate associations with hypothesized outcomes of reflective supervision, such as home visitor well-being, job satisfaction, or self-efficacy.
- Explore if the frequency and depth of discussing topics during reflective supervision sessions relates to home visitor confidence or self-efficacy in discussing and providing support to families on the same or similar topics.
- Examine whether current practices are effective in achieving intended short- and long-term outcomes for the supervisor and home visitor.
- Identify key elements of effective reflective supervision in specific contexts and populations.
- Assess the relative contribution of reflective supervision in supporting home visiting program outcomes.
- Evaluate changes in reflective supervision practices in response to training or other professional development activities (e.g., coaching, community of practice) focused on supporting reflective supervision.

For Supervisors

The SuPRS–HV can be used to plan for and review reflective supervision sessions. The items in the measure capture key topics and reflective supervision practices for the home visiting context. Reviewing the items in advance provides an opportunity for a supervisor to consider incorporating specific practices into the session. Examining the measure completed over time could also provide an opportunity for supervisors and managers to gain insight into the content of ongoing reflective supervision sessions. Whether collected over a short period of time (e.g., measure is completed after every weekly session for a month) or a longer period of time (e.g., measure is completed once a month every month for a year), this information could allow supervisors to see which key topics are discussed more and less frequently and which practices supervisors may want to incorporate more often. Comparing the information collected after sessions with different home visitors could also provide insight into which practices are used more often with individual home visitors and provide an opportunity for a supervisor to reflect on an individual's supervision needs.

Specifically, the SuPRS–HV can be used to—

- Provide guidance for preparing for and documenting reflective supervision sessions.
- Provide a menu of specific reflective supervision practices.
- Gain insight into the practices currently used and topics discussed during reflective supervision.
- Develop guidelines, training, professional development (e.g., coaching, community of practice), and practice improvement grounded in effective reflective supervision practices.
- Support supervisors' individual reflection and practice improvement.

Strengths and Limitations

Below, we discuss strengths and limitations of the measure.

Strengths

The SuPRS–HV achieved strong content validity through ongoing, in-depth engagement with interested parties and large-scale testing with supervisors across home visiting models. As previously discussed, we included multiple rounds of review and input from interested parties, including groups with experience and expertise in reflective supervision. Participants included practitioners, researchers, evaluators, and reflective supervision consultants. Additionally, our large-scale testing included a national sample of supervisors from a wide range of locales, program types, and home visiting models. Another strength of the SuPRS–HV is the ease with which it can be used in practice and research.

Limitations

The following limitations should be considered before using the SuPRS–HV, several of which arise from its design as a self-report measure intended to fill a current gap in the field. First, the SuPRS–HV is subject to the bias of supervisors rating their own performance; its results have not yet been compared to those of observational measures completed or home visitor reports of individual sessions. With respect to concurrent validity analysis using home visitor reports, we identified concerns among the limited home visitor reports available for comparison. These included limited psychometric evidence and a tendency to demonstrate ceiling effects (i.e., inflate results). Available measures also looked at general supervisory practices and styles across multiple sessions, not a single session, as measured by the SuPRS–HV. Relatedly, we did not explore associations between the SuPRS–HV and home visitor, program, or family outcomes to assess predictive validity.

Other limitations stem from our relatively early stage in the measure development and testing process. For example, because we used a convenience sample, rather than a nationally representative sample, the data reported in the technical appendix should not be viewed as national norms. In addition, we have not yet tested the SuPRS–HV as a tool to inform practice. More work is needed to understand how supervisors and programs can use the SuPRS–HV in practice.

References

Ash, J. (2010). *Reflective supervision rating scale*. Unpublished assessment tool.

Bernard, J. M., & Goodyear, R. K. (1998). *Fundamentals of clinical supervision*. Allyn & Bacon.

Bradley, W. J., & Becker, K. D. (2021). Clinical supervision of mental health services: A systematic review of supervision characteristics and practices associated with formative and restorative outcomes. *The Clinical Supervisor*, *40*(1), 88–111. https://doi.org/10.1080/07325223.2021.1904312

Cliffe, T., Beinart, H., & Cooper, M. (2016). Development and validation of a short version of the supervisory relationship questionnaire. *Clinical Psychology & Psychotherapy*, *23*(1), 77–86. https://doi.org/10.1002/cpp.1935

DeVellis, R. F. (2017). *Scale development: Theory and applications*. Sage Publications.

DeVellis, R. F., & Thorpe, C. T. (2021). *Scale development: Theory and applications*. Sage Publications.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail and mixed-mode surveys: The tailored design method*, 4th edition. John Wiley.

Efstation, J. F., Patton, M. J., & Kardash, C. M. (1990). Measuring the working alliance in counselor supervision. *Journal of Counseling Psychology, 37*(3), 322–329. https://doi.org/10.1037/0022-0167.37.3.322

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Friedlander, M. L., & Ward, L. G. (1984). Development and validation of the Supervisory Styles Inventory. *Journal of Counseling Psychology, 31*(4), 541–557. https://doi.org/10.1037/0022-0167.31.4.541

Gallen, R. T., Ash, J., Smith, C., Franco, A., & Willford, J. A. (2016). How do I know that my supervision is reflective?. *Zero to Three*, *37*(2).

Ghasemi, A., & Zahediasl, S. (2012). Normality tests for statistical analysis: A guide for non-statisticians. *International Journal of Endocrinology and Metabolism*, *10*(2), 486–489. https://doi.org/10.5812/ijem.3505

Gilkerson, L. (2004). Irving B. Harris Distinguished Lecture: Reflective supervision in infant-family programs: Adding clinical process to nonclinical settings. *Infant Mental Health Journal, 25*(5), 424–439.

Goodyear, R. K. (2014). Supervision as pedagogy: Attending to its essential instructional and learning processes. *The Clinical Supervisor*, *33*(1), 82–99. https://doi.org/10.1080/07325223.2014.918914

Hair, J., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis* (7th Edition). Prentice Hall.

Harrington, D. (2009). *Confirmatory Factor Analysis*. Oxford University Press.

Home Visiting Evidence of Effectiveness (HomVEE). (2024). *Models eligible for Maternal, Infant, and Early Childhood Home Visiting (MIECHV) funding.*

Hornby Zeller Associates Inc. (2014). *West Virginia home visitation program evaluation of professional development and community collaboration*.

IBM Corp. (2020). *IBM SPSS Statistics for Windows (Version 27.0)* [Computer software]. IBM Corp.

Ladany, N., Hill, C. E., Corbett, M. M., & Nutt, E. A. (1996). Nature, extent, and importance of what psychotherapy trainees do not disclose to their supervisors. *Journal of Counseling Psychology, 43*(1), 10–24. https://doi.org/10.1037/0022-0167.43.1.10

Mishra, P., Pandey, C. M., Singh, U., Gupta, A., Sahu, C., & Keshri, A. (2019). Descriptive statistics and normality tests for statistical data. *Annals of Cardiac Anaesthesia*, *22*(1), 67–72. https://doi.org/10.4103/aca.ACA_157_18

Muthén, B., & Muthén, L. (2017). Mplus. In *Handbook of item response theory* (507–518). Chapman and Hall/CRC.

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, *9*(4), 599–620. https://doi.org/10.1207/S15328007SEM0904_8

Brownstein, N. C., Louis, T. A., O'Hagan, A., & Pendergast, J. (2019). The role of expert judgment in statistical inference and evidence-based decision-making. *The American Statistician*, *73*(1), 56–68. DOI:10.1080/00031305.2018.1529623

Nevedal, A. L., Reardon, C. M., Opra Widerquist, M. A., Jackson, G. L., Cutrona, S. L., White, B. S., & Damschroder, L. J. (2021). Rapid versus traditional qualitative analysis using the Consolidated Framework for Implementation Research (CFIR). *Implementation Science*, *16*(1), 67. https://doi.org/10.1186/s13012-021-01111-5

Papageorgiou, S. N. On correlation coefficients and their interpretation. *Journal of Orthodontics* 49(3), 359–361. doi: 10.1177/14653125221076142

Sandstrom, H. S., Benatar, S., Peters, R., Genua, D., Coffey, A., Lou, C., Adelstein, S., & Greenberg, E. (2020). *Home visiting career trajectories: Final report* (OPRE Report #2020-11). Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. https://www.urban.org/sites/default/files/publication/101641/home_visiting_career_trajectories_0.pdf.

Shamoon-Shanok, R. (2009). What is reflective supervision? In S. S. Heller & L. Gilkerson (Eds.), *A practical guide to reflective supervision* (pp. 7–23). ZERO TO THREE.

Shea, S. E., Goldberg, S., & Weatherston, D. J. (2016). A community mental health professional development model for the expansion of reflective practice and supervision: Evaluation of a pilot training series for infant mental health professionals. *Infant Mental Health Journal*, *37*(6), 653–669. https://doi.org/10.1002/imhj.21611

Shea, S., Goldberg, S., & Weatherston, D. J. (2012). *Reflective supervision self-efficacy scale for supervisees.* Unpublished.

Shea, S. E., Jester, J. M., Huth-Bocks, A. C., Weatherston, D. J., Muzik, M., Rosenblum, K. L., & Michigan Collaborative for Infant Mental Health Research. (2020). Infant mental health home visiting therapists' reflective supervision self-efficacy in community practice settings. *Infant Mental Health Journal*, *41*(2), 191–205. https://doi.org/10.1002/imhj.21834

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education*, *2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Van Horn, J. (2018). *Reflective supervision: A guide from region X to enhance reflective practice among home visiting programs*. Reflective Supervision Collaborative in Region X, WA-AIMH.

Virginia Department of Health. (2017). *Virginia MIECHV expansion grant final report*.

Watkins, C. E., Jr. (2017). Convergence in psychotherapy supervision: A common factors, common processes, common practices perspective. *Journal of Psychotherapy Integration*, *27*(2), 140–152. http://dx.doi.org/10.1037/int0000040

Watson, C. L., Bailey, A. E., & Storm, K. J. (2016). Building capacity in reflective practice: A tiered model of statewide supports for local home-visiting programs. *Infant Mental Health Journal, 37*(6), 640–652. https://doi.org/10.1002/imhj.21609

Xiao, L., & Hau, K. T. (2023). Performance of coefficient alpha and its alternatives: Effects of different types of non-normality. *Educational and Psychological Measurement*, *83*(1), 5–27. https://doi.org/10.1177/00131644221088240

Appendix A. Additional Technical Information

This appendix provides item-level descriptive statistics for the SuPRS–HV and results of exploratory and confirmatory factor analysis completed as part of the large-scale test. We also provide information about participants and results from repeated measures administration of the SuPRS–HV.

Large Scale Testing

Descriptive Analyses

We examined measures of central tendency (mean, mode, and median) and dispersion (minimum, maximum, range, standard deviation, univariate normality, distributional skew, floor, and ceiling effects) for all SuPRS–HV items. Descriptive data are shown in exhibits A1–A8. Inspection of descriptive statistics and item distributions revealed most items on the SuPRS–HV were not highly skewed and demonstrated sufficient variability. Items with skewness values greater than -1 or 1 are shown with an asterisk in the tables below.

Exhibit A1. Descriptive Statistics on Items in the Responsiveness Subscale

Item	Mean	Standard deviation	Range
How often, if it all, did you:			
Wait for the home visitor to gather their thoughts	3.09	0.71	1–4
Let the home visitor know you heard and understood their feelings	3.25	0.70	1–4
Respond to what the home visitor shared without judgment	3.51	0.59	1–4
Offer support to promote the home visitor's well-being	2.99	0.79	1–4

Note: Higher scores indicate the practice happened more often in the session.

Exhibit A2. Descriptive Statistics on Items in the Collaborative Capacity Building Subscale

Item	Mean	Standard deviation	Range
How often, if at all, did you:			
Explore solutions together	3.07	0.63	1–4
Share specific knowledge or skills with the home visitor	2.73	0.78	1–4

Item	Mean	Standard deviation	Range
Demonstrate a potential strategy to use with the family	2.64	0.82	1–4
Help the home visitor process what is going on with families they find challenging to work with	3.03	0.80	1–4
Provide concrete resources the home visitor can use in their work with families	2.66	0.87	1–4
Provide resources to support the home visitor's professional development	2.30	1.00	1–4

Note: Higher scores indicate the practice happened more often in the session.

Exhibit A3. Descriptive Statistics on Items in the Unconditional Positive Regard Subscale

Item	Mean	Standard deviation	Range
How often, if at all, did you show:			
Interest in the home visitor's views	3.31	0.69	1–4
Appreciation for the home visitor's ideas	3.38	0.66	1–4
Acceptance of the home visitor as a person regardless of their behaviors or ideas*	3.48	0.64	1–4
How often, if at all, were you:			
Authentic in your interactions with the home visitor*	3.63	0.58	1–4
Comfortable discussing sensitive topics*	3.46	0.70	1–4
Aware of how your emotions impacted the session	3.21	0.85	1–4

Note. Higher scores indicate that the practice happened more often in the session.
*Skewness value greater than -1.

Exhibit A4. Descriptive Statistics on Items in the Promoting Awareness of Self and Others Subscale

Item	Mean	Standard deviation	Range
How often, if at all, did you:			
Ask the home visitor to describe and explore their feelings about experiences with families	2.91	0.86	1–4

Item	Mean	Standard deviation	Range
How often, if at all, did you encourage the home visitor to consider how the following might impact their work with families?			
Home visitor's personal identities (e.g., race, ethnicity, or culture)	1.59	0.63	1–3
Home visitor's assumptions or beliefs about a family based on family characteristics (e.g., race, ethnicity, or culture)	1.85	0.63	1–3
Home visitor's experiences	2.31	0.60	1–3
How often, if at all, did you encourage the home visitor to consider:			
How a family's **culture** might shape a family's life experience	2.46	0.92	1–4
How a family's **racial or ethnic identity** might shape a family's life experience	2.16	0.98	1–4
The perspectives of people involved in a situation (e.g., family members, coworkers, people important to a family)	2.53	0.84	1–4

Note: Higher scores indicate that the practice happened more often in the session.

Exhibit A5. Descriptive Statistics on Items on the Supportive Feedback Subscale

Item	Mean	Standard deviation	Range
How often, if at all, did you share:			
A specific strength of the home visitor	2.95	0.80	1–4
Specific, positive feedback on something the home visitor did	3.08	0.74	1–4
Feedback on what the home visitor was doing well before offering a suggestion for improvement	2.84	0.92	1–4

Note: Higher scores indicate the practice happened more often in the session.

Exhibit A6. Descriptive Statistics on Topics Discussed

Item	Mean	Standard deviation	Range
How often, if at all, did you discuss:			
Impact of a situation on the baby or child	2.28	0.61	1–3
Caregiver and child interactions	2.34	0.60	1–3
Strengths of the families the home visitor is working with	2.33	0.52	1–3
Needs of the families the home visitor is working with	2.63	0.50	1–3

Item	Mean	Standard deviation	Range
Setting and maintaining healthy boundaries with the families the home visitor is working with	2.11	0.67	1–3
Impact of program requirements, logistics, or other aspects of the work environment on the home visitor's work with families	2.11	0.65	1–3
Impact of community characteristics (e.g., availability of safe housing, employment, or affordable childcare) on the home visitor's work with families	2.16	0.64	1–3
Impact of working with families on the home visitor's well-being	2.34	0.61	1–3
What is going well with your supervisory relationship	1.75	0.68	1–3
How often, if at all, did you:			
Encourage the home visitor to consider how work-related stressors may affect their personal life	2.56	1.04	1–4
Provide an opportunity to discuss their professional development goals	2.15	1.02	1–4

Note: Higher scores indicate the practice happened more often in the session.

Item Intercorrelations

We examined bivariate correlations among items. Most correlations exceeded $r = 0.30$, a standard threshold for the appropriateness of exploratory factor analysis (EFA) (Hair et al., 2010) but were less than 0.80, which can suggest redundancy.

Participant Experiences With the SuPRS–HV

Participants responded to 10 questions where they rated their agreement (1 = strongly disagree to 4 = strongly agree) with statements about their experience taking the SuPRS–HV. Across these questions, respondents reported positive experiences with the reflective supervision measure (exhibit A7).

Exhibit A7. Participant Experiences Completing the SuPRS–HV

Question	Mean	Standard deviation
It was easy for me to complete this measure about my most recent reflective supervision session.	3.48	0.60
I was able to complete this measure all at one time, without having to stop and finish later.	3.52	0.65

Question	Mean	Standard deviation
The instructions for completing this measure were clear.	3.61	0.55
The questions on this measure were clear.	3.59	0.53
It would be easy for me to complete this measure after a reflective supervision session.	3.47	0.64
These are good questions to ask about what happens in reflective supervision sessions.	3.60	0.54
I would prefer to complete this kind of measure on paper rather than online.*	1.58	0.67
The length of this measure is just about right—not too long or burdensome to complete.	3.23	0.65
The answer options for this measure make sense and worked well for most or all of the items.	3.36	0.53
The questions seemed overly redundant or repetitive.*	1.88	0.51

Note: *For two of the questions, lower scores indicate a more positive perception of the experience of completing the SuPRS–HV.

Exploratory Factor Analysis

We used EFA to examine the dimensionality of the SuPRS–HV and identify potential subscales for further examination. We conducted a series of EFAs using a random 60 percent of the large-scale sample (*N* = 293), reserving the remaining cases (*N* = 209) for a confirmatory factor analysis (CFA) to test the stability of the final factor solution derived through EFA. A sample size of approximately 300 is sufficient to obtain stable, replicable, and generalizable factor structures with the 60 items on the draft measure (DeVellis, 2017). We used the Mplus software version 8.7 for the EFAs, employing Full Information Maximum Likelihood (FIML) to handle missing data (Muthén & Muthén, 2017). All analyses specified between one and seven factors and used a mean- and variable-adjusted weighted least squares (WLSMV) estimator for categorical responses along with oblique rotation (i.e., geomin), which allows factors to correlate with one another. We used several quantitative and qualitative criteria to evaluate model fit across the different factor solutions (exhibit A8). The former included examination of eigenvalues (> = 1.0), scree plots, and goodness of fit statistics (which provide information on how well each factor model reproduces the observed data)—such as (a) the chi-square test of whether the reproduced covariance matrix matches the data covariance matrix (nonsignificant chi-square values indicate good fit; but because this test is very sensitive to sample size, it will be interpreted with caution and in conjunction with the remaining indices); (b) Standardized Root Mean Square Residual (SRMR), a measure of the average residuals between reproduced and data matrices (values < .08 will be considered good fit); (c) Root Mean Square Error

of Approximation (RMSEA), a measure of average residuals that adjusts for model parsimony (RMSEA values < .05 will be considered good fit); and (d) the Confirmatory Fit Index (CFI), a measure of incremental fit compared to the baseline (no relationships) model (CFI values > .90 will be considered adequate fit and > .95 good fit). We also examined patterns of factor loadings and cross-loadings in alternative factor solutions. Criteria for item retention included a factor loading of 0.50 or higher with no cross-loadings of that magnitude on other factors. Qualitative criteria included the substantive interpretability of factors by the study team and its expert advisors, based on our underlying conceptual model of RS, review of the RS literature, concept mapping results, and existing frameworks and theories that have been used to understand and explain effective supervision in the helping professions (e.g., Bernard & Goodyear, 1998; Bradley & Becker, 2021; Goodyear, 2014; Watkins, 2017).

Exhibit A8. Fit Indices for Exploratory Factor Analysis Solutions

Fit indices	One-factor solution	Two-factor solution	Three-factor solution	Four-factor solution	Five-factor solution	Six-factor solution	Seven-factor solution
Chi square	2767.23	1749.84	1416.76	1119.26	859.28	684.35	565.20
Degrees of freedom	560	526	493	461	430	400	371
P-value	< .001	< .001	< .001	< .001	< .001	< .001	< .001
RMSEA	0.12	0.09	0.08	0.07	0.06	0.05	0.04
CFI	0.79	0.88	0.91	0.94	0.96	0.97	0.98
SRMR	0.13	0.09	0.07	0.06	0.05	0.04	0.04

Note: *N* = 293. Standards for fit indices are non-statistically significant chi square, RMSEA < .05, CFI > .95, and SRMR < .05.

The best-fitting factor solution was identified after iteratively eliminating 25 of the original 60 items that either failed to load at 0.50 or higher on any factor in the EFA; cross-loaded (0.50 or higher) on more than one factor; and/or were deemed redundant, irrelevant, or otherwise not useful by the study team and their expert advisors. The final EFA revealed that five factors provided the best fit to the data and provided the most conceptually meaningful solution.

Confirmatory Factor Analysis

Using Mplus software version 8.7 (Muthén & Muthén, 2017), we conducted CFA to assess the stability of the five-factor EFA solution described above. CFA tests whether a specified model fits the data (i.e., whether the covariance matrix estimated based on the model adequately approximates

the covariance matrix from the observed data [Harrington, 2009]). As noted above, we conducted the CFA using an independent random sample of approximately 40 percent of the 502 cases in the large-scale sample (N = 209). A sample of 200 is sufficient to provide a stable solution in the CFA model of the full sample (Muthén & Muthén, 2002). The factor model tested in the CFA was based on the five-factor EFA but added four nonloading items identified by the study team as critical to RS practice (see exhibit A9). The CFA used a mean- and variable-adjusted WLSMV estimator for categorical responses; it freely estimated the loading of each item on its respective factor from the EFA (or the most conceptually related factor for additional items) and constrained loadings on all other factors at zero. Model fit was assessed using the parameters described above for testing EFA model fit (i.e., CFI, SRMR, RMSEA, chi-square). The modified five-factor solution demonstrated adequate fit ($\chi 2(289) = 680.93$, $p < .001$; RMSEA = 0.08; CFI = 0.94; and SRMR = 0.08) in the independent random CFA sample and no modification indices suggested alternative specifications. Factor loadings are shown in exhibit A9.

Exhibit A9. Results of Confirmatory Factor Analysis: Five-Factor Solution

Subscale	Item	Factor loading
Responsiveness	Wait for the home visitor to gather their thoughts	0.60
Responsiveness	Let the home visitor know you heard and understood their feelings	0.85
Responsiveness	Respond to what the home visitor shared without judgment	0.76
Responsiveness	Offer support to promote the home visitor's well-being	0.78
Collaborative Capacity Building	Explore solutions together	0.69
Collaborative Capacity Building	Share specific knowledge or skills with the home visitor	0.77
Collaborative Capacity Building	Demonstrate a potential strategy to use with a family	0.82
Collaborative Capacity Building	Process what is going on with families they find challenging to work with	0.70
Collaborative Capacity Building	Concrete resources the home visitor can use in their work with families	0.88
Collaborative Capacity Building	Resources to support the home visitor's professional development	0.86
Unconditional Positive Regard	Interest in the home visitor's views	0.83

Subscale	Item	Factor loading
Unconditional Positive Regard	Appreciation for the home visitor's ideas	0.98
Unconditional Positive Regard	Acceptance of the home visitor as a person regardless of their behaviors or ideas	0.81
Unconditional Positive Regard	Authentic in your interactions with the home visitor	0.70
Unconditional Positive Regard	Comfortable discussing sensitive topics	0.71
Unconditional Positive Regard	Aware of how your emotions impacted the session	0.75
Promoting Awareness of Self and Others	To describe and explore their feelings about experiences with families	0.82
Promoting Awareness of Self and Others	Home visitor’s personal identities	0.69
Promoting Awareness of Self and Others	Home visitor’s assumptions or beliefs about a family based on family characteristics (e.g., race, ethnicity, culture)	0.65
Promoting Awareness of Self and Others	Home visitor's experiences	0.62
Promoting Awareness of Self and Others	How a family's culture might shape a family's life experience	0.84
Promoting Awareness of Self and Others	How a family's racial or ethnic identity might shape a family's life experience	0.86
Promoting Awareness of Self and Others	The perspectives of people involved in a situation (e.g., family members, coworkers, people important to a family)	0.80
Supportive Feedback	A specific strength of the home visitor	0.83
Supportive Feedback	Specific, positive feedback on something the home visitor did	0.97
Supportive Feedback	Feedback on what the home visitor was doing well before offering a suggestion for improvement	0.88

Note: *N* = 209. Only factor loadings > .5 are shown.

Repeated Measures Testing

Participating Supervisors

In the repeated measures testing, 39 supervisors completed the SuPRS–HV after three reflective supervision sessions with two or three home visitors. While supervisors were intentionally selected to participate in this phase of measure testing to achieve racial and ethnic diversity, the distribution of race/ethnicity mirrored the large-scale sample overall (exhibit A10).

Nearly all supervisors (97 percent) reported English as their primary language. The repeated measure supervisor sample also had high levels of education (exhibit A11). The fields of highest level of education were diverse (exhibit A16). The top three most common fields of education were—

- Other social science
- Education (early education/special education)
- Social work

The supervisors who participated in the repeated measures data collection were also experienced at providing home visiting services (exhibit A12) and reflective supervision (exhibit A13).

Exhibit A10. Repeated Measures Supervisor Race/Ethnicity

Race and Ethnicity	*N*	%
American Indian or Alaska Native	4	10.3
Black or African American	8	20.5
Latino or Hispanic	4	10.3
White	24	61.5

Exhibit A11. Repeated Measures Supervisor Highest Level of Education

Highest level of education	*N*	%
Some college or technical/training program	1	2.6
Associate's degree	1	2.6
Bachelor's degree	17	43.6
Graduate degree	20	51.3

Exhibit A12. Repeated Measures Supervisor Field of Study of Highest Level of Education

Field of study of highest level of education	*n*	%
Other social science	10	25.6
Education (early education/special education)	7	17.9
Social work	7	17.9
Nursing	6	15.4
Psychology	4	10.3
Child development/human development	2	5.1
Humanities/Liberal arts	2	5.1
Business administration/management/finance/accounting	1	2.6

Exhibit A13. Repeated Measures Supervisor Experience Providing Home Visiting Services

Experience	*N*	%
1-2 years	3	7.7
3-5 years	9	23.1
6-10 years	12	30.8
More than 10 years	15	38.6

Exhibit A14. Repeated Measures Supervisor Experience Providing Reflective Supervision

Experience	*n*	%
Less than 6 months	1	2.6
6 months to less than 1 year	1	2.6
1-2 years	13	33.3
3-5 years	10	25.6
More than 5 years	14	35.9

Program Characteristics of Participating Supervisors

Supervisors in the repeated measure sample worked in a wide variety of contexts at home visiting programs in a variety of locales across the country. Fifteen percent worked in programs that serve

Indigenous communities. Supervisors represented programs in 25 states and territories, with the highest percentages coming from the following states:

- Illinois (15 percent)
- South Carolina (8 percent)

Supervisors worked at home visiting programs that served all urbanicities (exhibit A15), though very few provided home visiting services in frontier areas (3 percent).

Supervisor programs used 12 different home visiting models.[4] The most common of these were—

- Parents as Teachers (36 percent)
- Healthy Families America (31 percent)
- Nurse-Family Partnership (15 percent)

Home visiting programs varied in size, though few (3 percent) served less than 25 families and almost half (44 percent) served more than 100. Exhibit A16 shows the complete breakdown of the size of programs. Families served by these programs were diverse in terms of race/ethnicity (exhibit A17) and language. They supervised varying numbers of home visitors (exhibit A18), though over three quarters supervised four or more. Supervisors provided in-person, individual (97 percent); in-person, group (28 percent); virtual, individual (62 percent); and virtual, group (15 percent) supervision. All the supervisors in the repeated measures sample received reflective supervision training or professional development. The specific trainings that supervisors participated in are shown in exhibit A19, with model-specific reflective supervision being the most common (46 percent of supervisors who reported receiving reflective supervision training).

Exhibit A14. Geographic Area(s) Served by Home Visiting Programs of Repeated Measure Supervisors

Urbanicity	*N*	%
Urban	27	69.2
Suburban	22	56.4
Rural	24	61.5
Frontier	1	2.6

Note: Percentages exceed 100 because multiple selections were allowed.

[4] Models represented include Attachment and Biobehavioral Catch Up (ABC), Baby TALK, Child First, Early Head Start Home-Based Option, Early Steps to School Success (ESSS), Family Spirit, Health Access Nurturing Developmental Services (HANDS), Healthy Families America, Nurse Family Partnership, Parents as Teachers, ParentChild+, and SafeCare Augmented.

Exhibit A15. Number of Families Served by Home Visiting Programs of Repeated Measure Supervisors

Number of families	*n*	%
Fewer than 25	1	2.6
Between 25 and 50	6	15.4
Between 51 and 74	7	17.9
Between 75 and 99	8	20.5
More than 100	17	43.6

Exhibit A16. Race/Ethnicity of Families Served by Home Visiting Programs of Repeated Measure Supervisors

Race/ethnicity	*n*	%
American Indian or Alaska Native	4	10.3
Asian	4	10.3
Black or African American	28	71.8
Latino or Hispanic	27	69.2
Native Hawaiian or Pacific Islander	2	5.1
White	23	59.0
Other	2	5.1

Note: Percentages exceed 100 because multiple selections were allowed.

Exhibit A17. Number of Home Visitors Supervised by Repeated Measure Supervisors

Number of home visitors supervised	*n*	%
2	3	7.7
3	3	7.7
4 or more	33	84.6

Exhibit A18. Reflective Supervision Training Repeated Measures Respondents Participated in

Training	*n*	%
Model-specific reflective supervision training	17	45.9
Facilitating Attuned Interactions (FAN)	6	16.2

Training	*n*	%
Infant Mental Health endorsement as an infant-family reflective supervisor	2	5.4
Reflective Interaction Observation Scale (RIOS)	2	5.4
Other	10	27.0

Participating Home Visitors

One-hundred ten home visitors, who were supervised by 39 supervisors participating in repeated measures, completed a survey about their experiences with their supervisors.[5] The racial and ethnic makeup of the home visitor sample was similar to supervisors, with 47 percent of home visitors reporting their race as White, 31 percent Latino, 18 percent Black, 6 percent American Indian or Alaskan Native being the largest response categories (exhibit A19).

Most home visitors reported their primary language as English (91 percent), and 9 percent indicated their primary language was Spanish. Just over half of home visitors reported having a bachelor's degree (55 percent). Others reported having a graduate degree (13 percent), associate's degree (16 percent), some college with no degree (14 percent), or a high school diploma (3 percent). The full breakdown of education levels is displayed in exhibit A20. The fields of highest level of education were diverse (exhibit A21). The top three most common fields of education were—

- Social work
- Nursing
- Education (early education/special education)

Home visitors also had a range of experience levels, varying from 14 percent reporting less than 1 year of home visiting experience to 7 percent with more than 10 years. Most of the sample (39 percent) reported 1–2 years of experience followed by 21 percent with 3–5 years and 19 percent with 6–10 years. All experience levels are shown in exhibit A22.

Exhibit A19. Home Visitor Race/Ethnicity

Race/ethnicity	*n*	%
American Indian or Alaska Native	7	6.4
Asian	1	0.0

[5] While the 39 supervisors who participated in repeated measures testing reported on reflective supervision sessions for 111 home visitors that they supervised, one home visitor left their position before completing the home visitor survey, yielding a home visitor sample size of 110.

Race/ethnicity	*n*	%
Black or African American	20	18.2
Latino or Hispanic	34	30.9
White	52	47.3
Other	1	0.0
Prefer not to answer	2	1.8

Note: Percentages may exceed 100 because multiple selections were allowed.

Exhibit A20. Home Visitor Highest Level of Education

Highest level of education	*n*	%
High school diploma or equivalent	3	2.7
Some college, no degree	15	13.6
Associate's degree	18	16.4
Bachelor's degree	60	54.5
Graduate degree	14	12.7

Exhibit A21. Home Visitor Field of Study of Highest Level of Education

Field of study of highest level of education	*n*	%
Social work	21	19.3
Nursing	20	18.3
Education (early education/special education)	18	16.5
Psychology	16	14.7
Other social science	11	10.1
Child development/human development	5	4.6
Business administration/management/finance/accounting	3	2.8
Public health	1	0.9
Humanities/liberal arts	1	0.9
Other	13	11.9

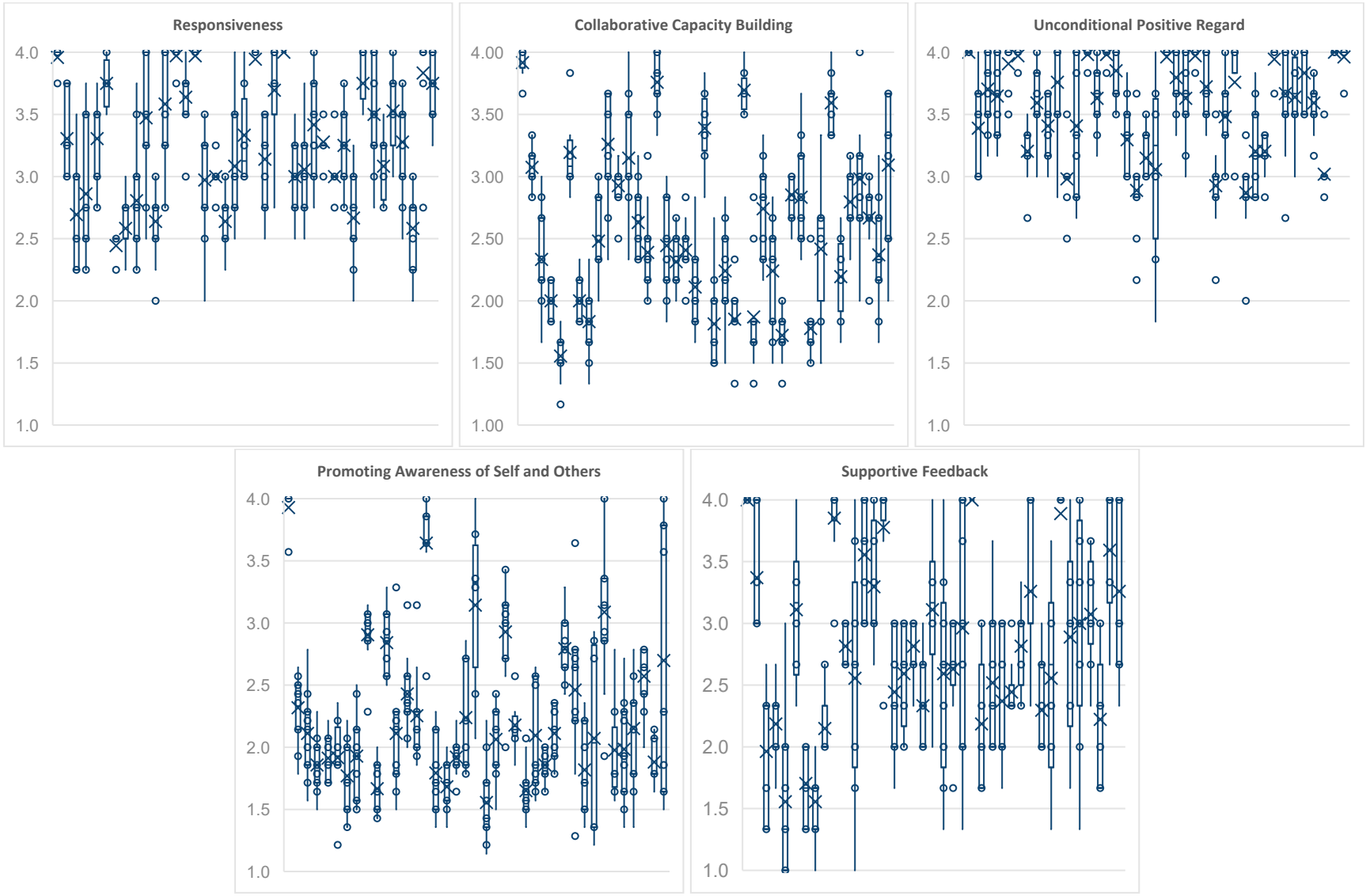
Exhibit A22. Home Visitor Experience Providing Home Visiting Services

Experience	*n*	%
Less than 1 year	15	13.6
1-2 years	43	39.1
3-5 years	23	20.9
6-10 years	21	19.1
More than 10 years	8	7.2

Variability in Responses Across Sessions and Home Visitors

Box-and whisker plots are shown in exhibit A23 for each SuPRS–HV subscale. Score variability across home visitors and sessions was evident for nearly all supervisors on subscales. Some supervisors demonstrated greater variability in scores than others. Although variable, scores on the *Responsiveness* and *Unconditional Positive Regard* subscales clustered toward the higher end of the response scale. In contrast, scores on the *Promoting Perspective Taking* subscale clustered toward the lower end of the scale.

Exhibit A23. Variability in SuPRS–HV Scores Across Sessions and Home Visitors



Note: Each box and whisker plot show the variability in SuPRS–HV subscale scores for each supervisor (*N* = 39) participating in repeated measures across all sessions and supervisees. Each subscale score ranges from 1–4, where higher scores indicate the practices occurred more often in the session.

Associations Between SuPRS–HV and Home Visitor Validity Measures

We examined the associations between the SuPRS–HV and home visitor reports of the reflective supervision they received. These reports included the Reflective Supervision Rating Scale (RSRS; Ash, 2010; Gallen et al., 2016), the Supervisory Styles Inventory (SSI; Friedlander & Ward, 1984), the Short Supervisory Relationship Questionnaire (S-SRQ; Cliffe et al., 2016), and the Supervisory Satisfaction Questionnaire (Ladany et al., 1996). We used a regression approach to account for the nested nature of the data (home visitor within supervisor). Specifically, we conducted a series of bivariate regression analyses predicting scores on each of the home visitor validity measures from SuPRS–HV subscale scores averaged across the three reflective supervision sessions. The magnitude and direction (positive or negative) of unstandardized regression coefficients were examined to determine the degree to which supervisor scores on the SuPRS–HV are associated with home visitor scores on each validity measure. Power analyses showed that a sample size of 111 nested supervisor-supervisee pairs is appropriate for this type of clustered regression analysis.

Scores on most home visitor validity measures were negatively skewed, with means and medians at the high end of each scale. Internal consistency was in the very good to excellent range for most home visitor validity measures except the Supervision Structure subscale of the RSRS and the Structure subscale of the S-SRQ, with alpha coefficients in the poor range (*alpha* = 0.65 and 0.57, respectively). Scores on the validity measures were highly intercorrelated, ranging from 0.48–0.78.

Regression results are reported in exhibits A24-A27. Scores on the Promoting Perspective Taking Subscale of the SuPRS–HV were not significantly associated with scores on any of the validity measures. Scores on the *Collaborative Capacity Building* and *Supportive Feedback* subscales of the SuPRS–HV showed significant associations with scores on the Task Orientation subscale of the SSI. *Unconditional Positive Regard* and *Responsiveness* were significantly associated with the Mentoring and Supervision Substructure subscales of the RSRS measure and with the Reflective Education and/or Structure Subscales of the S-SRQ.

Exhibit A24. Associations Between the SuPRS–HV and the Reflective Supervision Rating Scale

SuPRS–HV Subscale	Reflective process and skills	Mentoring	Supervision structure	Mentalization
Responsiveness	0.193	**0.230**	0.167	0.101
Collaborative Capacity Building	0.039	0.062	0.040	0.032
Unconditional Positive Regard	0.285	**0.317**	**0.320**	0.131

SuPRS–HV Subscale	Reflective process and skills	Mentoring	Supervision structure	Mentalization
Promoting Awareness of Self and Others	0.013	0.036	-0.039	0.011
Supportive Feedback	0.011	0.029	0.045	-0.012

Note: Bivariate nested regression unstandardized coefficients are displayed. **Bolded** coefficients were statistically significant, $p < .05$. SuPRS–HV subscale scores were averaged across three sessions.

Exhibit A25. Associations Between the SuPRS–HV and the Supervisory Styles Inventory

SuPRS–HV Subscale	Attractive	Interpersonally sensitive	Task oriented
Responsiveness	0.193	0.203	0.380
Collaborative Capacity Building	0.237	0.274	**0.508**
Unconditional Positive Regard	-0.033	0.092	0.101
Promoting Awareness of Self and Others	0.021	0.092	0.062
Supportive Feedback	0.179	0.175	**0.367**

Note: Bivariate nested regression unstandardized coefficients are displayed. **Bolded** coefficients were statistically significant, $p < .05$. SuPRS–HV subscale scores were averaged across three sessions.

Exhibit A26. Associations Between the SuPRS–HV and the Short Supervisory Relationship Questionnaire

SuPRS–HV Subscale	Safe Base	Reflective education	Structure
Responsiveness	0.264	0.313	**0.455**
Collaborative Capacity Building	-0.040	0.136	0.225
Unconditional Positive Regard	0.346	**0.457**	0.331
Promoting Awareness of Self and Others	-0.097	0.072	0.088
Supportive Feedback	-0.020	0.062	0.132

Note: Bivariate nested unstandardized regression coefficients are displayed. **Bolded** coefficients were statistically significant, $p < .05$. SuPRS–HV subscale scores were averaged across three sessions.

Exhibit A27. Associations Between the SuPRS–HV and the Supervisory Satisfaction Questionnaire

SuPRS–HV Subscale	Supervisory Satisfaction Questionnaire
Responsiveness	0.128
Collaborative Capacity Building	0.017
Unconditional Positive Regard	0.192
Promoting Awareness of Self and Others	-0.002
Supportive Feedback	-0.022

Note: Bivariate nested unstandardized regression coefficients are displayed. **Bolded** coefficients were statistically significant, $p < .05$. SuPRS–HV subscale scores were averaged across three sessions.

Appendix B. SuPRS–HV for Programming

In this appendix, we provide a suggested format for programming the SuPRS–HV that includes survey administration best practices and mirrors how the SuPRS–HV was programmed and administered during the large scale and repeated measures testing. Numbers in parentheses indicate the scoring options we used, but individual researchers may choose to use different values.

Supervisor Practices in Reflective Supervision-Home Visiting (SuPRS–HV)

Please answer the questions below based on the reflective supervision session you just finished. The questions are meant to capture what is happening during reflective supervision sessions. They are not an assessment of quality or an evaluation of your performance as a supervisor. We understand that every session is different, and that most supervisors do not have time to use all the practices asked about below in a single supervision session.

Thinking about the reflective supervision session you just completed—

How often, if at all, did you discuss:	Not discussed in the session (1)	Discussed in some of the session (2)	Major focus of the session (3)
1. Impact of a situation on the baby or child			
2. Caregiver and child interactions			
3. Strengths of the families the home visitor is working with			
4. Needs of the families the home visitor is working with			
5. Setting and maintaining healthy boundaries with families the home visitor is working with			

[page break]

How often, if at all, did you discuss:	Not discussed in the session (1)	Discussed in some of the session (2)	Major focus of the session (3)
6. Impact of program requirements, logistics, or other aspects of the work environment on the home visitor's work with families			
7. Impact of community characteristics (e.g., availability of safe housing, employment, or affordable childcare) on the home visitor's work with families			
8. Impact of working with families on the home visitor's well-being			
9. What is going well in your supervisory relationship			

[page break]

How often, if at all, did you:	Not in the session (1)	A little bit in the session (2)	Often in the session (3)	Most of the session (4)
10. Encourage the home visitor to consider how work-related stressors may affect their personal life				
11. Provide an opportunity for the home visitor to discuss their professional development goals				

[page break]

Thinking about the reflective supervision session you just completed—

How often, if at all, did you:	Not in the session (1)	A little bit in the session (2)	Often in the session (3)	Most of the session (4)
12. Wait for the home visitor to gather their thoughts				
13. Let the home visitor know you heard and understood their feelings				
14. Respond to what the home visitor shared without judgment				
15. Offer support to promote the home visitor's well-being				

[page break]

Thinking about the reflective supervision session you just completed—

How often, if at all, did you:	Not in the session (1)	A little bit in the session (2)	Often in the session (3)	Most of the session (4)
16. Explore solutions together				
17. Share specific skills with the home visitor				
18. Demonstrate a potential strategy to use with a family				

[page break]

How often, if at all, did you:	Not in the session (1)	A little bit in the session (2)	Often in the session (3)	Most of the session (4)
19. Help the home visitor process what is going on with families they find challenging to work with				
20. Provide concrete resources the home visitor can use in their work with families				
21. Provide resources to support the home visitor's professional development				

[page break]

Thinking about the reflective supervision session you just completed—

How often, if at all, did you show:	Not in the session (1)	A little bit in the session (2)	Often in the session (3)	Most of the session (4)
22. Interest in the home visitor's views				
23. Appreciation for the home visitor's ideas				
24. Acceptance of the home visitor as a person regardless of their behaviors or ideas				

[page break]

How often, if at all, were you:	Not in the session (1)	A little bit in the session (2)	Often in the session (3)	Most of the session (4)
25. Authentic in your interactions with the home visitor				
26. Comfortable discussing sensitive topics				
27. Aware of how your emotions impacted the session				

[page break]

Thinking about the reflective supervision session you just completed—

How often, if at all, did you:	Not in the session (1)	A little bit in the session (2)	Often in the session (3)	Most of the session (4)
28. Ask the home visitor to describe and explore their feelings about experiences with families				

[page break]

How often, if at all, did you encourage the home visitor to consider how the following might impact their work with families? *You may encourage reflection on these topics either directly or indirectly or use different words or terms. We are interested in how often, if at all, reflection on the general topics occurred.*	Not discussed in the session (1)	Discussed in some of the session (2)	Major focus of the session (3)
29. Home visitor's personal identities (e.g., race, ethnicity, or culture)			
30. Home visitor's assumptions or beliefs about a family based on family characteristics (e.g., race, ethnicity, or culture)			
31. Home visitor's experiences			

[page break]

How often, if at all, did you encourage the home visitor to consider:	Not in the session (1)	A little bit in the session (2)	Often in the session (3)	Most of the session (4)
32. How a family's **culture** might shape a family's life experience				
33. How a family's **racial or ethnic identity** might shape a family's life experience				
34. The perspectives of people involved in a situation (e.g., family members, coworkers, people important to a family)				

[page break]

Thinking about the reflective supervision session you just completed—

How often, if at all, did you share:	Not in the session (1)	A little bit in the session (2)	Often in the session (3)	Most of the session (4)
35. A specific strength of the home visitor				
36. Specific, positive feedback on something the home visitor did				
37. Feedback on what the home visitor was doing well before offering a suggestion for improvement				