Adapting psychophysiological data collection for COVID-19: The “Virtual Assessment” model

Alexandra R. Tabachnick[1], Tabitha Sellers[1], Emma Margolis[1], Madelyn Labella[1], Dylan Neff[2], Shelia Crowell[2], K. Lee Raby[2], Celine Saenz[2], Elisabeth Conradt[2], Mary Dozier[1]

[1] Department of Psychological and Brain Sciences, University of Delaware, Newark, Delaware, USA
[2] Department of Psychology, University of Utah, Salt Lake City, Utah, USA

Correspondence
Alexandra R. Tabachnick, Department of Psychological and Brain Sciences, Univer-sity of Delaware, 108 Wolf Hall, Newark, DE 19716, USA. Email: atabach@udel.edu

EC and MD share senior authorship on this paper.

Funding information
the National Institutes of Health,
Grant/Award Numbers: (F31DA050426 to AT, R01HD098525 to MD, R01MH119070 to SC; Anonymous Donor; The Univer-sity of Utah Consortium for Families and Health Research; Purdue Pharma, L. P.; The Interdisciplinary Research Pilot Pro-gram

Abstract

The COVID-19 pandemic has significantly disrupted research activities globally. Researchers need safe and creative procedures to resume data collection, particu-larly for projects evaluating infant mental health interventions. Remote research is uniquely challenging for psychophysiological data collection, which typically requires close contact between researchers and participants as well as techni-cal equipment frequently located in laboratory settings. In accordance with pub-lic health guidance, we adapted procedures and developed novel protocols for a “virtual assessment” in which women and infants provided behavioral and psychophysiological data from their own homes while researchers coordinated remotely. Data collected at virtual visits included video-recorded parent–child interactions and autonomic nervous system data. Adaptations were designed to optimize safety and data quality while minimizing participant burden. In the current paper, we describe these adaptations and present data evaluating their success across two sites in the United States (University of Delaware and Univer-sity of Utah), focusing specifically on autonomic nervous system data collected during the well-validated Still-Face Paradigm (SFP). We also discuss advantages and challenges of translating traditional lab procedures into the virtual assess-ment model. Ultimately, we hope that disseminating these procedures will help other researchers resume safe data collection related to infant mental health dur-ing the COVID-19 pandemic and beyond.

Keywords

autonomic nervous system, methodology, psychophysiology

Introduction

Infant mental health research typically requires repeated measurements at precise developmental time periods. However, the COVID-19 pandemic has posed significant challenges to such research due to the ongoing risks associated with in-person interactions and the unknown duration of the disruption to daily life. To collect time-sensitive data safely and ethically, innovative adaptations to traditional in-person assessment practices must be considered.

Although some data collection procedures may translate easily to online platforms (e.g., online survey collection, behavioral data collection via video conferencing), collecting physiological data safely during COVID-19 is more challenging. The Centers for Disease Control and Prevention (CDC) safety recommendations highlight the importance of mask wearing and social distancing (i.e., remaining at least six feet away from individuals outside of your household) as vital precautions to mitigating the spread of COVID-19 (Centers for Disease Control and Prevention, 2021). This CDC guidance poses unique challenges to psychophysiological data collection, which requires close contact, such as placement of electrodes for electrocardiogram (ECG) measurement.

In response to these challenges, we developed an innovative data collection procedure to safely resume psychophysiological infant mental health research during the COVID-19 pandemic. Specifically, we describe procedures to measure infant respiratory sinus arrhythmia (RSA) during the Still-Face Paradigm (SFP; Tronick et al., 1978), across two sites. The SFP is a validated social stress task, which is commonly used to measure infant behavioral and physiological regulation in a relational context (Jones-Mason et al., 2018; Mesman et al., 2009). RSA is an indicator of parasympathetic nervous system activity and has been associated with emotion regulation when measured at rest and in response to emotionally salient tasks, including the SFP (Beauchaine, 2015). Over time, researchers have become increasingly interested in measuring infant RSA during the SFP. The standard SFP consists of three 2-min episodes: Play, Still-Face, and Reunion (or second Play episode). A recent meta-analysis indicated that on average, infants show a pattern of decreasing RSA in response to the Still-Face episode (Jones-Mason et al., 2018). However, some infants fail to show RSA decreases or withdrawal, and still other infants experience RSA decreases to the Still-Face episode but fail to exhibit RSA recovery or increases to the Reunion or second Play episode (Jones-Mason et al., 2018; Mesman et al., 2009). More research is needed regarding the degree to which individual differences in physiological response to the SFP may be explained by infant-level factors, parent-level factors, parent–infant relationship factors, and sociodemographic context. In addition, as RSA regulation may have implications for the development of psychopathology (Beauchaine, 2001, 2015), it is critical to evaluate whether early intervention programs may promote healthy autonomic regulation for vulnerable infants. For example, the Attachment and Biobehavioral Catch-up intervention (ABC), a 10-session parenting program delivered in infancy, has been associated with improved autonomic regulation in middle childhood compared to a control intervention (Tabachnick et al., 2019). However, it is not yet known whether parenting programs in infancy may proximally influence infant RSA responses to the SFP.

Although home-based and remote data collection procedures are not new, there are significant limitations to these traditional methods in the context of COVID-19. Prior to the pandemic, infant researchers made use of technology to collect video data, including parent–child interaction paradigms and looking time procedures (e.g., Tran et al., 2017). Since the onset of the pandemic, developmental researchers have called for increased infrastructure to facilitate purely online data collection (Sheskin et al., 2020; Zaadnoordijk et al., 2021). Further, researchers interested in psychophysiology have developed innovative procedures involving online task administration and mailing materials to collect saliva (Gunnar et al., 2021). However, it is not currently possible to collect RSA data via purely online data collection or by mailing materials to participants. Therefore, there is a specific need for procedures to facilitate safe RSA data collection during the COVID-19 pandemic, especially for researchers interested in infant mental health, given the impacts of the COVID-19 pandemic on parents and families (Brown et al., 2020).

Relevance statement

Although COVID-19 caused a major disruption to infant mental health intervention research, the present study demonstrates the feasibility of resuming safe psychophysiological data collection using a "virtual assessment" model. Specifically, the majority of virtual assessment visits across two sites produced usable autonomic nervous system data, and infant autonomic reactivity appeared similar to data collected in the laboratory. Key adaptations as well as strengths and challenges of this model are summarized.

In this paper, we describe the adaptations we made across two sites (the University of Delaware and the University of Utah) in the United States and propose considerations for other researchers who are interested in virtual psychophysiological data collection. Next, we describe procedures related to our specific measures and present data regarding preliminary feasibility and validity of the virtual assessment model for collecting infant psychophysiological data, including comparing virtual assessment data to laboratory data collected at the University of Delaware. Finally, we review the strengths and limitations of this model and discuss its relevance to data collection after COVID-19.

2 | METHOD

2.1 | Project context

Data for the present study were collected at two universities in distinct regions of the United States. Although data collection at each of the sites was motivated by different sets of aims, infant autonomic nervous system data in the SFP were collected at both sites prior to the onset of the pandemic, and the two research teams collaborated to develop the virtual assessment model to continue autonomic data collection during the pandemic. Preliminary virtual assessment data presented in the current study are therefore collapsed across both contexts. The presented laboratory data were derived only from the University of Delaware.

2.1.1 | The University of Delaware – US Mid-Atlantic Region

Data were collected as part of a randomized clinical trial evaluating the efficacy of a modified version of the Attachment and Biobehavioral Catch-up (mABC) intervention for pregnant women in treatment for opioid dependence. Pregnant and peripartum women were eligible to participate if they endorsed receiving opioid treatment medication during pregnancy. ABC is a brief home-visiting program, which aims to increase parental sensitivity and nurturance, and decrease parental intrusiveness (Dozier & Bernard, 2019). Data collection for the trial was planned for the third trimester of pregnancy (pre-intervention), and when the infant was three, six, and 12 months old. Primary outcome measures included parenting observations and infant behavioral and autonomic nervous system regulation. Thus, continuing to collect these data at the appropriate ages was key to the overall aims of the trial and necessary for establishing an evidence base for a novel infant mental health intervention. For the present study, data collected in the laboratory prior to and during the pandemic and virtual assessment data from the University of Delaware were drawn from the same clinical trial to maximize the similarity between procedures for comparison. The data collection procedures were nearly identical, aside from the setting and the adaptations described below.

2.1.2 | The University of Utah – US Mountain West Region

Autonomic nervous system data from mothers and infants were collected as part of a prospective longitudinal study on the intergenerational transmission of emotion dysregulation (Lin et al., 2019; Ostlund et al., 2019). Participants were recruited from local OB/GYN clinics and were eligible to participate if they had a singleton pregnancy, no substance use during pregnancy, and no pregnancy complications. A key premise of this study is that emotion dysregulation can be measured at multiple levels of analysis: maternal and infant behavior, maternal report, and autonomic nervous system functioning. Data collection occurred prenatally and at 7 and 18 months, with autonomic assessments at each time point. It was therefore essential that we adapted our physiological assessment protocols to reduce missing data and maintain high retention.

2.2 | Participants

Participants were combined across the two sites because the same procedures for the virtual assessment were used across both sites, and because the present study is focused on preliminary feasibility and validity. Combining samples is a common practice to maximize sample size and generalizability (e.g., NIH Environmental influences on Child Health Outcomes (ECHO) Program and the Study of Early Child Care and Youth Development). Future studies with larger samples may control for potential site differences in statistical analyses.

Participants were 27 mothers and their infants. Twenty participants were part of data collection at the University of Utah, and seven participants were enrolled in data collection at the University of Delaware. One participant at the University of Delaware completed virtual assessments at two timepoints during the pandemic. Therefore, eight virtual assessments were conducted at the University of Delaware, and a total of 28 assessments were conducted across both sites. Notably, data collection is ongoing at both universities, so this group represents a subsample of the total virtual assessments that will ultimately be collected. Both projects were ongoing at the onset of COVID-19, and participants varied in how many timepoints they completed pre- vs. mid-pandemic. All procedures including research and consent protocols were approved by the University of Delaware and the University of Utah Institutional Review Boards.

All mothers identified as female. Most infants were described by their mothers as being female (n = 20), with six infants described as male and one infant missing data regarding biological sex. At the time of data collection for the virtual assessments, mothers were approximately 30 years old (M = 29.96, SD = 4.17) and infants were about 44 weeks old (M = 43.58, SD = 9.54). Mothers primarily identified as white and non-Latina (n = 21), with one identifying as white and Latina, one as Black or African American and non-Latina, and two as multiracial and non-Latina. Race/ethnicity data were not available for the remaining two mothers. Eighteen infants were described by their mothers as white and non-Latino/a/x, with three described as white and Latino/a/x, one as Black or African American and non-Latino/a/x, and four as multiracial and non-Latino/a/x. Race and ethnicity data were missing for one infant.

Race and ethnicity distributions were similar across university sites, with most participants across sites identifying as white and non-Latino/a/x. Sites also did not statistically differ with respect to participant age (infant $t = 1.07$, $p = .32$; mother $t = .02$, $p = .98$). The majority of infants across both sites were female (University of Delaware: 71%; University of Utah: 75%). However, participants at the University of Utah reported significantly greater educational attainment ($t = -7.80$, $p < .001$) and larger household incomes ($t = -6.80$, $p < .001$) than participants at the University of Delaware. Complete demographic data for participants from both sites are presented in Table 1.

2.3 | Summary of adaptations

When developing the virtual assessment model, our goals were to continue infant data collection on schedule while optimizing safety and data quality and minimizing participant burden. Table 2 provides a concise summary of the adaptations that were made to accomplish these goals.

2.3.1 | Goal #1: prioritizing safety

The safety of participants and researchers is of the utmost concern, particularly when attempting data collection during a pandemic. The day before the visit and the day of the visit, parents completed a screening tool assessing for self-reported COVID-19 symptoms and any recent COVID-19 exposures for themselves and their infants. Researchers delivering materials to the participants completed the same health screening, and all parties were required to pass the health screening for the visits to occur. To further minimize possible exposure, we employed a contactless delivery of assessment supplies. That is, the researcher left a data collection kit outside the participant's door, and the participant collected the kit after the researcher moved at least six feet away. Researchers and participants were instructed to wear high quality, well-fitting masks during this exchange. Surgical masks were also provided to researchers and participants upon request.

Although risk of fomite transmission is thought to be low relative to transmission by aerosol or respiratory droplets (Goldman, 2020), additional steps were taken to minimize possible transmission of COVID-19 through fomites on surfaces. First, researchers washed or sanitized their hands before and after handling equipment. Next, participants used appropriate disposable products (e.g., disposable electrodes, paper towels), and a trash bag was provided for participants to dispose of those single-use products in their home before returning the data collection kit to the researcher. Finally, researchers used CDC-recommended cleaning supplies (e.g., Sani-Cloth disposable wipes, disinfectant spray, soap, and water) to sanitize reusable products between uses.

2.3.2 | Goal #2: optimizing data quality

We took several steps during the planning stage to optimize the quality of the physiological data collected during the virtual assessments. First, psychophysiological measures were carefully selected based on robustness to artifact and administration challenges. In the laboratory, we typically collect both ECG and cardiac impedance data with seven spot electrodes. However, cardiac impedance is known to be highly sensitive to motion artifact and is typically collected with four spot electrodes on the chest and back. Such a measure places a relatively high burden on participants to collect by themselves, and there is a high risk for noisy and unusable data. Thus, we elected only to collect ECG data, which can be measured with three spot electrodes on the chest and is likely to yield usable data even with motion artifact and inexact sensor placement. In this way, evaluation and careful selection of measures are key steps to optimize data quality.

For all measures, pilot testing and practice administration were critical to optimize scripts and protocols, and to ensure researchers were able to clearly communicate the instructions to participants. For data collection at each of the two sites, it was important for mothers to be seated across from their infants, with one camera capturing the mother's behavior and a second camera capturing the infant's behavior. We developed simple, clear instructions as well as visual aids (see Figures 1 and 2) to help participants set up furniture, tablets for video recording, and physiological equipment. Researchers practiced delivering verbal instructions and providing visual aids using screen-sharing functions in Zoom. For example, "We will now put stickers on [BABY's NAME] ribs and collarbone. You can use the picture of the baby in the handout to help. Before you place each sticker, show me where you are going to put it so that I can check it is in the right spot." The instructions developed for the virtual assessment made it possible for parents to accurately place physiological recording sensors on their infants, correctly connect electrode leads to unfamiliar recording devices, and prepare recording devices for data acquisition.

In addition, we tested equipment intended for use during virtual assessment, including tablets and mobile hotspots for internet access, prior to finalizing data collection materials. Stable internet is essential for clear communication with participants and remote control of the

TABLE 1 Demographic data

Characteristic	Overall, $N = 27$[a]	University of Delaware, $N = 7$[a]	University of Utah, $N = 20$[a]
Family income			
Less than $10,000	3 (12%)	3 (43%)	0 (0%)
$10,000–$19,999	2 (7.7%)	2 (29%)	0 (0%)
$20,000–$35,999	2 (7.7%)	1 (14%)	1 (5.3%)
$35,000–$49,999	4 (15%)	1 (14%)	3 (16%)
$50,000–$74,999	5 (19%)	0 (0%)	5 (26%)
$75,000–$99,999	5 (19%)	0 (0%)	5 (26%)
$100,000 or more	5 (19%)	0 (0%)	5 (26%)
Mother education			
Less than high school diploma	2 (7.7%)	2 (29%)	0 (0%)
High school diploma or equivalent	6 (23%)	4 (57%)	2 (11%)
Some college	2 (7.7%)	1 (14%)	1 (5.3%)
Technical school	1 (3.8%)	0 (0%)	1 (5.3%)
Associate degree	2 (7.7%)	0 (0%)	2 (11%)
Bachelor's degree	8 (31%)	0 (0%)	8 (42%)
Master's degree	3 (12%)	0 (0%)	3 (16%)
Doctoral degree	2 (7.7%)	0 (0%)	2 (11%)
Mother race			
Black or African American	1 (3.8%)	1 (14%)	0 (0%)
Declined to answer	1 (3.8%)	0 (0%)	1 (5.3%)
Multiple races	2 (7.7%)	0 (0%)	2 (11%)
White	22 (85%)	6 (86%)	16 (84%)
Mother ethnicity			
Latina	2 (7.7%)	1 (14%)	1 (5.3%)
Not Latina	24 (92%)	6 (86%)	18 (95%)
Mother age (years)	29.96 (4.17)	30.00 (5.48)	29.95 (3.76)
Infant sex			
Female	20 (77%)	5 (71%)	15 (79%)
Male	6 (23%)	2 (29%)	4 (21%)
Infant race			
Black or African American	1 (3.8%)	1 (14%)	0 (0%)
Multiple races	4 (15%)	0 (0%)	4 (21%)
White	21 (81%)	6 (86%)	15 (79%)
Infant Latino/a/x			
Latino/a/x	3 (12%)	2 (29%)	1 (5.3%)
Not Latino/a/x	23 (88%)	5 (71%)	18 (95%)
Infant age (weeks)	43.15 (9.45)	47.10 (17.56)	41.77 (4.12)

[a] *n* (%); mean (SD).

physiological equipment. Local mobile hotspot options included the Verizon Jetpack MiFi 8800L, Ellipsis Jetpack, and AT&T Velocity. However, it is important that each research team assesses local mobile data coverage when selecting a hotspot provider. Tablet speaker volume is also important for participant communication, and tablet cameras must be able to capture videos of sufficient quality at varying levels of ambient light for later behavioral coding. Tablet options include the Microsoft Surface Pro, the iPad, and the Amazon Fire; tablet tripods are also required and may be set up in advance for participants. Existing models of home data collection may also use mobile hotspots

TABLE 2 Best practices for virtual psychophysiological assessments during the COVID-19 pandemic

Goal	Strategies
Prioritize safety of participants and researchers	• Screen symptoms prior to visit • Contactless delivery of materials • Wear high quality, well-fitting masks • Disinfect all supplies between visits
Optimize data quality	• Provide clear instructions for sensor placement and camera setup • Select robust physiological measures • Account for additional time needed for troubleshooting • Complete pilot testing and/or practice administration
Minimize participant burden	• Provide all necessary supplies • Set up equipment as much as possible in advance (e.g., plug in cables, connect tablet to video chat) • Use screen-sharing to display diagrams and administer tasks • Eliminate nonessential data collection

and tablets. However, remotely troubleshooting with participants to fix problems with the same devices used for communicating with participants, devices which are also not familiar to participants, may be a unique challenge to the virtual assessment without proper preparation.

Data quality may be optimized by building in sufficient time for in-the-moment troubleshooting with technology and other challenges. Helping researchers and participants anticipate that troubleshooting is to be expected decreases anxiety and frustration, increases the likelihood that problems will be solved, and decreases the frequency of mistakes. The most common challenges were hotspot or internet connectivity problems. If the tablets or recording computer disconnect from the internet, it may be necessary to call the participant by phone to talk through steps to reconnect, including restarting devices or software, turning Wi-Fi off and on again, and/or switching from the hotspot to the participant's home Wi-Fi network. Although participants are likely skilled at navigating these functions on their own devices, it may be difficult to adjust to an unfamiliar device, so it is important for the researcher to be prepared to provide simple but detailed instructions by phone. For example, "Ok, it looks like it is not connecting right now. Can you hit the back arrow in the top left of the screen to go back to the Wi-Fi page? Then, can you click the white circle next to Wi-Fi at the top to try turning off the Wi-Fi? We are going to wait a minute before turning it back on and then try to reconnect again." In general, we found that the virtual assessments took an average of 15 min longer than comparable procedures in the laboratory, but there were outlier cases that experienced unique difficulties (e.g., recurrent internet connection problems, chaotic home environment) and took up to an hour longer than comparable laboratory procedures.

2.3.3 | Goal #3: minimizing participant burden

The virtual assessment model described here requires participants to set up key components of technological and physiological equipment themselves. The prospect of managing unfamiliar electronics and conducting physiological data collection may be intimidating or overwhelming for participants. To minimize participant burden, we created a virtual assessment kit that included all necessary materials for behavioral and physiological data collection. Materials provided to the participant were prepared for use as much as possible. Specifically, the on-site research assistant completed the following steps: attached electrodes to color-coded leads, connected the two tablets to internet hotspots, logged into a Zoom meeting with the remote researcher on the tablets, placed the tablets on tripods set to the correct height, and connected the recording computer to the internet hotspot and enabled screen-sharing access for the remote researcher (via TeamViewer software). Thus, the remaining steps for participants to complete themselves (with coaching from the remote researcher) were minimal. The kit also included an infant seat for families in case they did not have their own highchairs. Several steps described above that were taken to optimize data quality also served to minimize participant burden (e.g., eliminating nonessential data collection and using screen-sharing features of Zoom to display visual aids and deliver instructions).

2.4 | Procedures

For each virtual assessment, there was an on-site research assistant and one or two remote researchers. The on-site research assistant was responsible for setting up the visit kit and contactless delivery of the kit to the participant, as described above. Typically, the research assistant who delivered the research equipment remained on site (e.g., in the car outside) during the visit to provide support if necessary. Prior to the research assistant's arrival, the research team determined the safest location for the assistant to wait while also being available for support.

The remote researcher(s) then conducted the virtual assessment via Zoom from any private location with a strong, stable internet connection (i.e., they did not need to be in an office building or the laboratory). First, the

FIGURE 1 Image from a video recording during the Still-Face Paradigm. The infant is clearly visible for later behavioral coding and monitoring the sensors, and the researcher can observe the mother's behavior to ensure she is following task instructions (e.g., making a "Still Face"). Image provided with permission

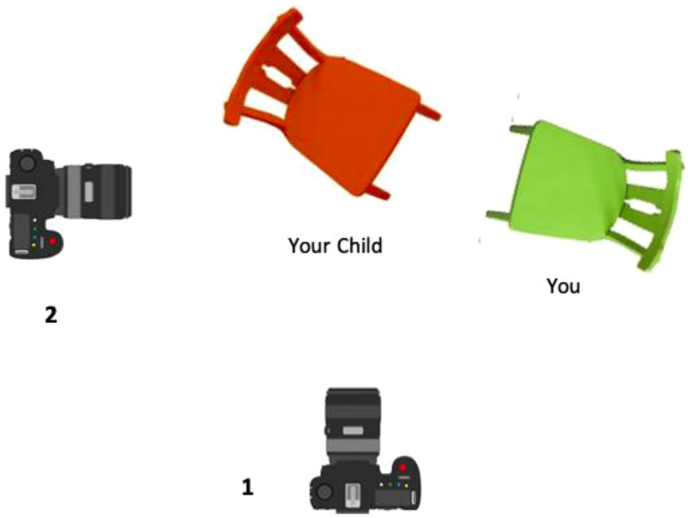

FIGURE 2 Diagram of furniture and camera arrangement for the Still-Face Paradigm, which can help participants understand how to set up the equipment

remote researcher(s) facilitated informed consent using an electronic consent form. Often it was easiest to text the link to the participant while discussing the contents of the consent form. Next, the remote researcher(s) coached the participant through setting up the physiological recording equipment, including placement of heart rate sensors, and used TeamViewer software to remotely control the physiological recording computer. Once physiological equipment was in place, the remote researcher(s) guided the participant through furniture and camera setup for the parent–child interactions. See Figure 1 for an example of optimal camera angles from the researcher's perspective, and Figure 2 for a diagram representing the furniture and camera arrangement. Both images are likely to help researchers work with participants to ensure adequate videos are captured for later coding. After interaction activities were completed, the remote researcher(s) coached the participant through sensor removal and packing up equipment. Finally, the on-site research assistant completed a contactless pick up of the visit kit and cleaned all equipment per CDC guidelines.

2.5 | Measures

2.5.1 | Still-Face Paradigm (SFP)

The SFP (Tronick et al., 1978) is a well-validated parent–child interaction task designed to assess infant behavioral reactivity and regulation during a social stressor. It has increasingly been used to assess infant autonomic reactivity and regulation as well (Jones-Mason et al., 2018). Several studies have found similar autonomic reactivity to the SFP in home and lab settings (Bush et al., 2017; Haley et al., 2006); however, these findings should be replicated to increase confidence that there are minimal setting effects on SFP outcomes.

The SFP consists of three 2-min episodes. In the first episode (Play 1), the parent is instructed to play with their infant as they normally would. In the second episode (Still-Face), the parent maintains a neutral face and avoids touching the infant or responding to infant signals. In the third episode (Play 2), the parent resumes playing or interacting with the infant as they normally would.

When developing virtual assessment procedures, we consulted with Dr. Edward Tronick, the creator of the task, to ensure it would be administered with fidelity. In the present study, the infant and the mother first completed a 2- or 3-min resting task in which they watched a Baby Einstein video to establish a physiological baseline. Next, the infant completed the SFP while seated in a highchair facing their mother. Both the infant's face and the mother's face were video-recorded for later behavioral coding. Laboratory data used for comparisons below followed the same procedure for the SFP as the virtual assessment.

2.5.2 | Autonomic nervous system data

RSA was derived from ECG data collected during the resting period and each episode of the SFP. MindWare hardware (Mobile Impedance Cardiograph) and software (BioLab Acquisition Software 3.1) were used for data collection[1]. Via Zoom, the remote researcher instructed the parent to place three disposable pediatric electrodes in a bipolar configuration on the infant's torso. The researcher then guided the parent in linking the mobile physiological recording device to the recording computer. Following sensor placement and hardware setup, the researcher remotely controlled the recording computer to monitor physiological data quality and record event markers in the continuous physiological data for later analysis. Mothers were asked to move the computer screen so that they could not view their incoming physiological signals.

Infant ECG was processed offline using MindWare's Heart Rate Variability Analysis Software (HRV). Infant physiological data were divided into 1-min segments for cleaning. ECG signals were cleaned via visual inspection and manual editing to remove artifacts resulting from infant movement or other interference (e.g., deleting erroneous beats, adding missing beats). Segments that required estimation of greater than 10% of total beats and segments in which less than 50% of beats were usable were excluded from analyses. After ECG data were cleaned, RSA levels during each segment were estimated as the natural log of the infant's average high-frequency heart rate variability in that segment, based on the frequency band recommended for use with infants (i.e., .24–1.04 Hz; Bar-Haim et al., 2000). Finally, average RSA was calculated for the baseline and each episode of the SFP, and an RSA reactivity change score was calculated by subtracting average RSA during the first play episode from average RSA during the Still-Face episode (SFP – Play 1). Laboratory data used for comparisons below were cleaned in the same way as data collected in the virtual assessments.

2.6 | Analytic plan

Descriptive data are presented to establish preliminary feasibility and validity of the virtual assessment procedures. First, rates of usable RSA data from the virtual assessment are described and compared to rates of usable data when the SFP was conducted in the laboratory. Next, patterns of RSA change across the SFP are examined and compared to what is typically found in the literature and to data collected in the laboratory. Based on prior work (Jones-Mason et al., 2018; Mesman et al., 2009), we expected that on average, infants would exhibit decreases in RSA from the initial PLAY episode to the Still-Face episode. Due to small sample sizes, statistical comparisons were not conducted.

3 | RESULTS

3.1 | Data usability

Autonomic nervous system data are often subject to technical difficulties, movement artifact, and other interference that can render the data unusable. To establish the utility of the virtual assessment model, we first calculated the proportion of participants with usable data. Across both sites, 21 infants had fully ($n = 9$) or partially ($n = 12$) usable RSA (i.e., at least one segment of the task usable), and seven participants did not have any usable data. Overall, infant physiological data acquired through virtual assessment were at least partially usable for both sites (University of Delaware: 62.5%; University of Utah: 80%; total: 75%). Those with partially usable data had 4.5 usable minutes on average (range: 3– 6 min; mode: 5 min) out of seven possible minutes at the

[1] An initial attempt to ease participant burden was the use of LED photoplethysmographs strapped around the mother's finger and the infant's ankle to record superficial blood flow for use as an approximation of heart rate; however, this method yielded data that were undiscernible due to movement artifact.

University of Delaware and six possible minutes at the University of Utah. The primary reason for unusable data was technical problems ($n = 9$), followed by infant fussiness ($n = 6$), noisy data ($n = 3$), and too-frequent arrhythmias ($n = 1$). Technical problems included episodes of signal interference, which disrupted data collection and required troubleshooting by the remote researcher, on-site research assistant, and participant to re-establish internet connection. For one case, the SFP was not completed as part of the virtual assessment due to repeated and persistent internet disruptions. For the remaining twenty-seven assessments, all videos were of sufficient quality for later behavioral coding.

Rates of unusable physiological data were somewhat higher for the virtual assessments than for similar data collected in the laboratory at the University of Delaware (either pre-pandemic or mid-pandemic with enhanced safety measures). Out of 43 assessments conducted in the lab at the University of Delaware, 39 had fully usable data, three had partially usable data (due to infant fussiness or excessive motion artifact), and only one did not have any usable data (due to infant fussiness). Thus, although most virtual assessments produced at least partially usable data (75%), lab assessments were less susceptible to technological problems and appeared to produce at least partially usable data more consistently (98%).

Notably, however, the proportion of usable data for the virtual assessments seemed to increase over time as the teams became more experienced and further improved procedures. Thus, the present estimate of usable data for virtual assessments is likely to be an underestimate of the proportion of usable data that can be achieved for the full sample.

3.2 | Preliminary validation of psychophysiological data

We tested the validity of our novel data collection procedures by examining the degree to which infants exhibited expected physiological reactivity patterns to the SFP. As expected (Jones-Mason et al., 2018), the majority of infants with usable reactivity data exhibited a decline in RSA from the first PLAY episode to the Still-Face episode ($n = 12$; 63%), with the remaining portion exhibiting an increase in RSA ($n = 7$; 37%). A somewhat greater proportion of infants at the University of Delaware exhibited RSA decreases (75%) compared to infants at the University of Utah (56.3%). However, this potential difference should be interpreted with caution because only four infants from the University of Delaware had sufficient usable data to calculate change in RSA, precluding statistical comparisons. Overall, this pattern of reactivity is comparable to data collected at the University of Delaware in the laboratory with similar procedures. In the 39 lab-based assessments with usable data, 27 infants exhibited RSA decreases to the SFP (69%), and 12 infants exhibited RSA increases (31%). Figure 3 presents a visual depiction of the patterns of infant RSA reactivity when the SFP is conducted in the home versus the laboratory.

4 | DISCUSSION

The present study aimed to establish the feasibility of the virtual assessment approach to collecting infant physiological data. Across two sites located in distinct regions of the United States, we demonstrated that most visits produced usable physiological and behavioral data, with most infants showing the expected pattern of physiological reactivity to a social stressor. We also described in detail the specific adaptations that were made to optimize safety during the COVID-19 pandemic, maximize data quality, and minimize participant burden.

4.1 | Strengths

This virtual assessment model has many notable strengths, including relatively efficient acquisition of usable psychophysiological data. Importantly, even at this initial validation stage where we expect higher rates of unusable data, the quality of the data collected by trained researchers is generally acceptable. In addition, physiological data collected in the home may be more ecologically valid than similar data collected in the lab because participants are not required to acclimate to an unfamiliar setting. Although collecting data in participants' homes are not new, the virtual assessment model enables autonomic nervous system data collection with infants without requiring a skilled technician on site. Although researchers may have perceived autonomic data acquisition as too complex for participants to collect themselves, the virtual assessment model is both feasible and acceptable for parents. Virtual assessment may be a key tool for infant mental health researchers to flexibly select appropriate data collection procedures as the COVID-19 pandemic persists and cases continue to wax and wane globally.

In addition to pragmatic advantages, the virtual assessment strategy has the benefit of ensuring equal access to opportunities for research while minimizing unintentional coercion. Offering a virtual data collection option helps to ensure the autonomy of under-resourced participants whose motivation to participate in paid assessments may otherwise outweigh hesitations about in-person interactions during the COVID-19 pandemic. Also, involving

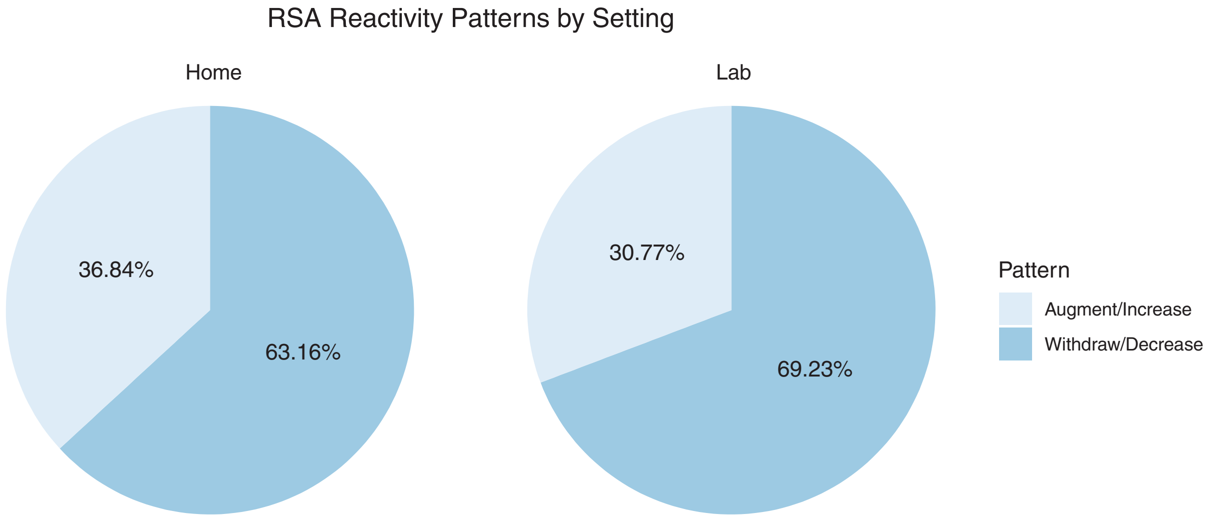

FIGURE 3 Patterns of infant respiratory sinus arrhythmia (RSA) reactivity to the Still-Face Paradigm in home and lab settings. Note: "Home" data were collected by teams at the University of Delaware and the University of Utah ($n = 19$). "Lab" data were collected by the team at the University of Delaware ($n = 39$)

participants directly in the data acquisition may help to increase their sense of agency and commitment to the research process. One mother at the University of Delaware commented proudly on her ability to complete the assessment: "I guess I was pretty savvy with the technology."

We believe that these flexible protocols for home-based psychophysiological assessment will remain invaluable as the current pandemic subsides. Acquiring physiological data in the home allows researchers to assume the burden of travel to meet participants where they are (including in motels, supportive housing, or Native American reservations) and facilitates the engagement of families with significant barriers to research participation. At the University of Delaware, participants cited multiple non-pandemic-related reasons for selecting virtual assessment, including difficulty traveling to campus with siblings, anxiety about traveling away from home, busy schedule preventing a commute to the laboratory, and infant's discomfort with car rides. Completing an assessment at home permitted these families to access research opportunities that would otherwise be unavailable to them. Thus, home-based psychophysiological assessments are critical for accessing remote, diverse, and underserved populations who are underrepresented in current studies of psychophysiology and infant mental health.

Although home-based assessments in general have many strengths as described above, the virtual assessment model offers several unique advantages. Researchers who use mobile data acquisition systems in their laboratories can adapt to virtual assessment at relatively low cost. In addition, investing in multiple sets of psychophysiological equipment may facilitate rapid data collection. For example, an experienced researcher who is working remotely and coordinating with multiple research assistants who are delivering supplies could complete more virtual assessments in a day than would be feasible using traditional in-person methods in participants' homes or in the laboratories. Using this data collection strategy, a small number of specialized researchers could acquire psychophysiological data from many participants across a broad geographic region. In contrast, traditional home-based research methods require an on-site researcher with technical expertise, reducing the possible number of assessments that can be conducted per day. Further, the present study extends available methods for purely remote research or mailing supplies for saliva collection (e.g., Gunnar et al., 2021; Tran et al., 2017), which are not conducive to autonomic nervous system data collection. In addition, the virtual assessment model does not presume that families have their own video-capable devices or stable internet connection, which is a limitation of purely remote data collection.

The virtual assessment model can also be extended to additional tasks and measures. For example, in addition to collecting behavioral and physiological ANS data during the SFP, the researchers at both sites also collected behavioral and ANS data during a 10-min play interaction to assess parental sensitivity. Setup for this task was largely the same, except only one camera was used for video recording. The University of Delaware researchers also coached mothers through saliva collection for DNA methylation analyses for themselves and their infants. One of the benefits of the virtual assessment model is its flexibility to adapt to various dyadic and individual assessments.

4.2 | Challenges

Despite the many strengths of this virtual assessment model, several challenges remain that may serve as barriers to both researchers and participants. Because this model increases the burden on the research team in terms of travel time and effort with transporting equipment, it may not be feasible for groups without significant grant funding and staffing. These challenges may create a barrier for researchers in rural areas, although travel burden is similar in traditional home-based research and does not represent a specific challenge of virtual assessment. Although it is ideal for on-site research assistants to remain nearby, it may be uncomfortable for research assistants during certain times of the year or in geographic regions where temperatures may not permit researchers to safely remain in their vehicles for extended periods of time. In addition, this model relies heavily on technological savvy and effective communication skills of the researcher conducting the visit remotely, the on-site research assistant dropping off supplies, and the participant. This additional coordination on behalf of the research team may help to explain why data collected through virtual assessments were more susceptible to technological problems than data collected in the lab, resulting in a lower proportion of usable data (75% vs. 98%). It is important to note that the proportion of usable data from virtual assessments is likely to increase over time as teams become more experienced. However, certain problems may still occur at a higher rate for virtual assessments compared to laboratory assessments (e.g., poor internet connection in a participant's home, interference from siblings and pets). Although we supplied mobile hotspots to increase accessibility for families who do not have home Wi-Fi networks, there are also some rural locations in which mobile hotspots do not function well. As described above, we elected to not collect cardiac impedance data due to the difficulty of placing electrodes on the back without help and high susceptibility to motion artifacts, even in laboratory conditions, which is a limitation. Researchers could elect to collect electrodermal activity as a measure of sympathetic nervous system functioning, although electrodermal assessments are also vulnerable to motion artifacts and are quite sensitive to environmental differences across homes (e.g., humidity, temperature). Thus, collecting sympathetic nervous system measures remotely will likely require further troubleshooting.

In terms of the technology itself, a strong internet connection is necessary for video calling and for accessing the computer remotely for physiological data collection. Mobile hotspots were found to work well, although strength of signal varied; researchers should consider their geographic region and investigate the best provider for that region. Additionally, signal interference within the home may cause the internet connection or physiological recording device to intermittently disconnect, which is disruptive to data collection. If this disruption occurs, the researcher may need to call the participant on the phone to coach them through internet reconnection. Alternately, if the participant is comfortable with the on-site research assistant entering the home, the assistant may help to reconnect devices from a safe distance while wearing appropriate personal protective equipment. To best support the participant, the research team should investigate the easiest method for reconnecting the specific devices being used in a given virtual assessment. In addition, pilot work is essential to ascertain which tablets and mobile devices work best for the research team's purposes. Considerations may include cost, sturdiness, camera quality for video recordings, and speaker and microphone volume/quality for communicating with participants and administering tasks.

Collecting data in the home was challenging for researchers before the pandemic due to lack of environmental control (e.g., noise made by other home occupants and limited space). The virtual assessment model removes another aspect of researcher control in that the researcher cannot navigate the physical space or handle the equipment directly. As a result, remote data collection may require more time than in-person assessments for visit setup and administration of tasks, and adaptations to standard procedures may be necessary to ensure feasibility.

Conducting virtual assessments with participants who are high in impulsivity and emotion dysregulation presents further challenges. Equipment setup and task administration are particularly difficult with participants who may proceed without receiving complete instructions or become easily frustrated by technological challenges. However, we believe that virtual assessments may be reasonably attempted with any family who consents and has physical space available, and we encourage researchers to prepare to be flexible and adjust expectations in the moment. Researchers are advised to empathize with the participant's experience, using active listening and validation to support and empower participants as they navigate each step of the assessment. For example, researchers may draw attention to the mother's ability to effectively calm her infant while coaching the mother through infant electrode placement. The researcher can also comment on how tricky it can be to follow multiple instructions while also monitoring their infant. Paper-based instructions may be useful for more independent participants. Additionally, the research team should be thoughtful in organizing and clearly labeling supplies for participants (e.g., keeping supplies for the mother separate from supplies for the infant).

In general, it may be helpful to create a priority list of measures in case it feels necessary to simplify the data collection plan for a given participant.

5 | CONCLUSION

In summary, we made strategic adaptations to traditional data collection procedures to prioritize safety during the COVID-19 pandemic, while optimizing data quality and minimizing participant burden. Overall, we demonstrate that the virtual assessment model is feasible and has preliminary validity for physiological data collection with infants. Future research may further test the validity of these methods with larger samples. Most assessments produced usable physiological data from the SFP, and data reflected the expected patterns of physiological reactivity to the task. Although there are challenges to the virtual assessment model, there are significant strengths. We hope that researchers are empowered to continue psychophysiological data collection during COVID-19 and to consider virtual assessment beyond the pandemic to increase equity and representation in infant mental health research.

ACKNOWLEDGMENTS

This work was supported by funding from the National Institutes of Health (F31DA050426 to AT, R01HD098525 to MD, R01MH119070 to SC and EC), Purdue Pharma, L. P., the University of Utah Consortium for Families and Health Research, the Interdisciplinary Research Pilot Program, and an anonymous donor.

CONFLICT OF INTEREST

The authors have no conflicts of interest to disclose.

DATA AVAILABILITY STATEMENT

Data are available upon request.

ORCID

Mary Dozier https://orcid.org/0000-0002-3618-4361

REFERENCES

Bar-Haim, Y., Marshall, P. J., & Fox, N. A. (2000). Developmental changes in heart period and high-frequency heart period variability from 4 months to 4 years of age. *Developmental Psychobiology*, *37*(1), 44–56. https://doi.org/10.1002/1098-2302(200007)37:1<44::AID-DEV6>3.0.CO;2-7

Beauchaine, T. P. (2001). Vagal tone, development, and Gray's motivational theory: Toward an integrated model of autonomic nervous system functioning in psychopathology. *Development and Psychopathology*, *13*, 183–214. https://doi.org/10.1017/S0954579401002012

Beauchaine, T. P. (2015). Respiratory sinus arrhythmia: A transdiagnostic biomarker of emotion dysregulation and psychopathology. *Current Opinion in Psychology*, *3*, 43–47. https://doi.org/10.1016/j.copsyc.2015.01.017

Brown, S. M., Doom, J. R., Lechuga-Peña, S., Watamura, S. E., & Koppels, T. (2020). Stress and parenting during the global COVID-19 pandemic. *Child Abuse & Neglect*, *110*, 104699. https://doi.org/10.1016/J.CHIABU.2020.104699

Bush, N. R., Jones-Mason, K., Coccia, M., Caron, Z., Alkon, A., Thomas, M., Coleman-Phox, K., Wadhwa, P. D., Laraia, B. A., Adler, N. E., & Epel, E. S. (2017). Effects of pre- and postnatal maternal stress on infant temperament and autonomic nervous system reactivity and regulation in a diverse, low-income population. *Development and Psychopathology*, *29*(5), 1553–1571. https://doi.org/10.1017/S0954579417001237

Centers for Disease Control and Prevention (2021, March 8). How to Protect Yourself & Others. https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/prevention.html

Dozier, M., & Bernard, K. (2019). *Coaching parents of vulnerable infants: The attachment and biobehavioral catch-up approach*. The Guilford Press.

Goldman, E. (2020). Exaggerated risk of transmission of COVID-19 by fomites. *The Lancet Infectious Diseases*, *20*(8), 892–893. https://doi.org/10.1016/S1473-3099(20)30561-2

Gunnar, M. R., Reid, B. M., Donzella, B., Miller, Z. R., Gardow, S., Tsakonas, N. C., Thomas, K. M., DeJoseph, M., & Bendezú, J. J. (2021). Validation of an online version of the Trier Social Stress Test in a study of adolescents. *Psychoneuroendocrinology*, *125*, 105111. https://doi.org/10.1016/J.PSYNEUEN.2020.105111

Haley, D. W., Handmaker, N. S., & Lowe, J. (2006). Infant stress reactivity and prenatal alcohol exposure. *Alcoholism: Clinical and Experimental Research*, *30*(12), 2055–2064. https://doi.org/10.1111/j.1530-0277.2006.00251.x

Jones-Mason, K., Alkon, A., Coccia, M., & Bush, N. R. (2018). Autonomic nervous system functioning assessed during the Still-Face Paradigm: A meta-analysis and systematic review of methods, approach and findings. *Developmental Review*, *50*, 113–139. https://doi.org/10.1016/j.dr.2018.06.002

Lin, B., Kaliush, P. R., Conradt, E., Terrell, S., Neff, D., Allen, A. K., Smid, M. C., Monk, C., & Crowell, S. E. (2019). Intergenerational transmission of emotion dysregulation: Part I. Psychopathology, self-injury, and parasympathetic responsivity among pregnant women. *Development and Psychopathology*, *31*(3), 817–831. https://doi.org/10.1017/S0954579419000336

Mesman, J., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (2009). The many faces of the Still-Face Paradigm: A review and meta-analysis. *Developmental Review*, *29*(2), 120–162. https://doi.org/10.1016/j.dr.2009.02.001

Ostlund, B. D., Vlisides-Henry, R. D., Crowell, S. E., Raby, K. L., Terrell, S., Brown, M. A., Tinajero, R., Shakiba, N., Monk, C., Shakib, J. H., Buchi, K. F., & Conradt, E. (2019). Intergenerational transmission of emotion dysregulation: Part II. Developmental origins of newborn neurobehavior. *Development and Psychopathology*, *31*(3), 833–846. https://doi.org/10.1017/S0954579419000440

Sheskin, M., Scott, K., Mills, C. M., Bergelson, E., Bonawitz, E., Spelke, E. S., Fei-Fei, L., Keil, F. C., Gweon, H., Tenenbaum, J. B., Jara-Ettinger, J., Adolph, K. E., Rhodes, M., Frank, M. C., Mehr, S. A., & Schulz, L. (2020). Online developmental science to foster innovation, access, and impact. *Trends in Cognitive Sciences*, *24*(9), 675–678. https://doi.org/10.1016/J.TICS.2020.06.004

Tabachnick, A. R., Raby, K. L., Goldstein, A., Zajac, L., & Dozier, M. (2019). Effects of an attachment-based intervention in infancy on children's autonomic regulation during middle childhood. *Biological Psychology*, *143*. https://doi.org/10.1016/j.biopsycho.2019.01.006

Tran, M., Cabral, L., Patel, R., & Cusack, R. (2017). Online recruitment and testing of infants with Mechanical Turk. *Journal of Experimental Child Psychology*, *156*, 168–178. https://doi.org/10.1016/J.JECP.2016.12.003

Tronick, E., Als, H., Adamson, L., Wise, S., & Brazelton, T. B. (1978). The infant's response to entrapment between contradictory messages in face-to-face interaction. *Journal of the American Academy of Child Psychiatry*, *17*(1), 1–13. https://doi.org/10.1016/S0002-7138(09)62273-1

Zaadnoordijk, L., Buckler, H., Cusack, R., Tsuji, S., & Bergmann, C. (2021). A global perspective on testing infants online: Introducing ManyBabies-AtHome. PsyArXiv. https://doi.org/10.31234/OSF.IO/CNWH5