

Does Temperament Differentially Influence Study Completion Rates for 18 and 24-  
Month-Olds Participating in a Remote Study?

Caroline Hamilton Taylor

Dissertation submitted to the faculty of the Virginia Tech Polytechnic Institute and State  
University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy  
In  
Psychology

Committee  
Robin Panneton (Chair)  
Martha Ann Bell  
Koeun Choi  
Jungmeen Kim-Spoon

June 5<sup>th</sup>, 2023  
Blacksburg, VA

Keywords: temperament, remote testing, attrition, toddlerhood, engagement

# Does Temperament Differentially Influence Study Completion Rates for 18 and 24-Month-Olds Participating in a Remote Study?

Caroline Hamilton Taylor

## ABSTRACT

Since the COVID-19 pandemic, many researchers turned to remote testing to continue to collect data. This shift provided new insights and exciting opportunities for researchers, including the ability to gain access to larger demographic pools. However, remote work has come with unique challenges. One factor that has proved to be challenging is the level of control researchers have when conducting remote studies. For example, compared to in-lab room setup and design, participants' home environments have numerous distractions for young toddlers (e.g., family members, pets, tv, toys). Thus, the increased variability has led to important questions regarding methodology, implementation, and in the current study, variability in participation. That is, are there systematic biases in final samples due to differences in participant characteristics, such as temperament? Particularly for remote work, the increased variability has created challenges for researchers to collect data but also exciting opportunities to understand how temperament may interact with participation and study completion rates. The current study aimed to understand whether temperament plays a role in study completion rates in remote research for toddlers ages 18 and 24 months. It was expected that effortful control would significantly influence participation and study completion, and that both negative affectivity and surgency would negatively influence participation in the remote study. Both effortful control and negative affectivity were not significantly related to participation, while surgency positively related to participation.

# Does Temperament Differentially Influence Study Completion Rates for 18 and 24-Month-Olds Participating in a Remote Study?

Caroline Hamilton Taylor

## GENERAL AUDIENCE ABSTRACT

Since the pandemic, many developmental researchers shifted from in-person settings to online testing. Although remote testing is not new, this shift provided many research labs with new opportunities to reach more participants. However, with this shift to remote testing, researchers relinquished control that has been designed in the lab, in that each home environment is different. For example, noise, internet speed, and computer size will vary across participants, whereas in the laboratory, these factors remained constant. Moreover, toys, tvs, pets, and/or siblings may be distracting for toddlers participating in a remote study. Thus, these differences across home environments have illuminated differences in toddlers' successful participation. For example, are some toddlers more likely to maintain their attention to the current task? Temperament was investigated in the current study to determine if certain skills may influence study participation and completion rates for toddlers ages 18 to 24-months. It was expected that effortful control (e.g., inhibitory control, attention focusing) would significantly influence the likelihood of a toddler completing the remote task. On the other hand, it was expected that both surgency (e.g., impulsivity, activity level) and negative affectivity (e.g., fear, shyness) would negatively influence task participation. Only surgency positively influenced task participation, and will be discussed in the context of toddler motivation, and engagement in the task.

## Acknowledgements

I am eternally grateful for the love and support I've received throughout my journey. I was able to reach this point with the support and guidance from my family, friends, and mentors along the way.

Thank you to my mentor, Dr. Robin Panneton for taking me on as your student. I am thankful for your countless hours of time, feedback, and guidance on every project and milestone. Thank you for pushing me to become a better researcher. I am also grateful to my lab mates, students, and families who have dedicated their time towards this project to make it possible.

Thank you to my committee members, Dr. Martha Ann Bell, Dr. Koeun Choi, and Dr. Jungmeen Kim-Spoon. Your feedback and guidance helped enrich this project and my academic career. I also want to thank my mentors at Emory & Henry College, Dr. Kimberly Baranowsky, Dr. Celeste Gaia, and Dr. Christopher Qualls for inspiring me to pursue Psychology and for igniting my love of research.

Thank you to my friends and cohort for adding so much joy and happiness into this journey. You helped to make Blacksburg a very special place and I will always cherish the memories we created.

Thank you to my parents, who have been my unwavering role models in life and for supporting me throughout every step of this journey. Even before I can remember, you never missed a chance to cheer me on. Thank you for encouraging me to pursue this endeavor and for your endless love, guidance, and support throughout it all. Your encouragement and belief in me has been and will always be a source of strength and motivation.

I would also like to thank my twin sister (and most recently, Dr.!). Going through graduate school together has been an incredible blessing. Thank you for your friendship, laughs, facetimes, and endless moral support as we were able to share this experience together.

Thank you to Frank, for your endless love and joyful spirit. I'm forever grateful graduate school led me to you.

## Table of Contents

List of Figures .....	vi
List of Tables .....	vii
Introduction.....	1
Method .....	12
Results.....	17
Discussion .....	23
References .....	32
Supplemental Materials .....	42

## **List of Figures**

Figure 1: Multisensory Attention Assessment Protocol (MAAP).....	14
Figure 2: Intersensory Processing Efficiency Protocol (IPEP).....	14

## List of Tables

Table 1: Cross Sectional and Longitudinal Participant Demographics .....	12
Table 2: Correlations between ECBQ Temperament Higher-Order Factors for Cross-Sectional Sample (18 and 24-month-olds included) .....	17
Table 3: Logistic Regression Predicting Completion Rate .....	18
Table 4: Linear Regression Predicting Completion Rate .....	19
Table 5: Means & Standard Deviations of ECBQ Dimensions.....	20
Table 6: Correlations between ECBQ Temperament, MAAP and IPEP Total Duration of Participation, and MAAP and IPEP Total Looking Proportion for 18-month-olds and 24-month-olds .....	21
Table 7: Correlations between Lower-Order Dimensions of Surgency, MAAP and IPEP Total Duration of Participation, and MAAP and IPEP Total Looking Proportion for 18-month-olds .....	22
Table 8: Correlations between Lower-Order Dimensions of Surgency, MAAP and IPEP Total Duration of Participation, and MAAP and IPEP Total Looking Proportion for 24-month-olds .....	23

## Introduction

The COVID-19 pandemic served as a catalyst for many researchers to turn to creative methods to continue to collect data, whether cross-sectional or longitudinal in design. From this shift to remote data collection, many benefits and advantages emerged, such as providing increased flexibility for families to participate in research studies, and importantly, the potential to mitigate sampling bias by obtaining a larger demographic pool (Nielsen et al., 2017).

However, even with increased access to a larger demographic pool, extant work has shown that samples are skewed towards WEIRD (Western, educated, industrialized, rich, and democratic) participants. A recent investigation has shown that 96% of research is based on people from Western societies; this percentage only represents ~12% of the world's population (Henrich et al., 2010). These authors caution researchers against generalizing without recognizing the limitations associated with overgeneralizing from this select subset. An important extension of this work will be to understand if there are person-centered biases as well. More specifically, are there certain characteristics of participants, particularly when considering infant and toddler samples, that may increase the likelihood that they will be included in the final sample for data analysis?

Developmental researchers are exceedingly familiar with dropout rates across all types of methodologies (e.g., fNIRS, eye-tracking, looking time studies) (Laroy et al., 2011). During testing, many infants and toddlers will 'fuss out' of the study, sometimes due to crying, restlessness, or drowsiness. Some toddlers will need breaks during the session to increase the likelihood they will sit through the duration of the testing, and in some cases, the researchers will choose to reschedule the session to optimize their chances at obtaining data from the participant. Miller (2017) proposed that infant researchers face roughly ~50-60% greater dropout rates than



research with older participants. Dropout rate is defined as the proportion of collected data that was excluded from the final analysis. For the purpose of the current study, this metric does not include participants who choose to not partake in the study, but only those who begin the study and do not finish.

Though dropout rates are reported, and methods of managing missing data are suggested (Asendorpf et al., 2014, Nicholson et al., 2017) there is less attention to dropout rates at a person-centered level. Considering the ways in which infants and toddlers may fuss out of a study, it is important to note that these examples may explain state factors that are not entirely predictable. For example, if testing overlaps with the toddler's regular routine, (e.g., eating or napping) fussing out of the study would likely reflect a behavioral state change that disrupts performance. In this particular scenario, the toddler will experience a change in their routine, thus leading to distress. On the other hand, if the behaviors are well-known to the caretaker and are predictable (e.g., distress to strangers or if they are restless when seated for a period of time), then a toddler dropping out of specific trials or an entire procedure (and thus may not be included in the final sample) may be an example of exclusion on the basis of temperamental characteristics. In this case, the final sample would not be representative across various temperamental profiles. This bias could lead to greater difficulties in the replication of the study's findings. As noted by Yu and colleagues (2020, p. 13) "sampling and generalizability are the methodological bedrocks of behavioral science, and knowing whether the sample is representative of the population is critical to the validity and generalizability of research findings." Thus, the purpose of the current study was to investigate how various temperamental characteristics may influence study completion rates for toddlers participating in a remote study.

Importantly, there are numerous factors that influence whether or not there will be participant dropout during testing. These factors may include experimental error, parental interference, internet problems, and more. Moreover, these factors will interact with 1) age, and 2) method (Baek et al., 2021; Stets et al., 2012). For example, some procedures may elicit higher dropout rates than others based on the type of methodology used (e.g., fNIRS vs. looking time studies) (Slaughter & Suddendorf, 2007). According to Graham (2009) dropout rates are ubiquitous in longitudinal research. Coupled with age (particularly for infants and toddlers) and the need for multiple lab visits, dropout rates are even more prevalent (Miller, 2017). Missing data in developmental research is not a novel occurrence or experience, but identifying any potential systematic patterns in dropout rates (not at random) is an important question to address when evaluating the generalizability, external validity, and reproducibility of results.

Not only have the findings related to temperament and dropout rates been mixed, they have also been conducted with much younger infants (i.e., 12-months or younger) and for studies using different methodological approaches (e.g., fNIRS, electroencephalogram (EEG), etc.). Importantly, by 18-months, toddlers have gained increased independence, language skills, and mobility from the time they were 12-months. Thus, this age group is an important one to investigate potential systematic patterns in dropout due to potential temperamental differences. Moreover, these studies have investigated dropout for studies that have been conducted in a lab setting. In-person testing involves highly controlled environments specifically designed for the study, and involves strict protocol adherence and room configuration. While remote research also involves strict protocol adherence, remote studies have inherently less control than in-person studies. Thus, there is significantly more variability in the room configuration alone, and there are likely more opportunities for toddlers to explore the room, and move around as they please in

their home. Again, it is important to understand whether temperament plays a role in dropout, but possibly even more so for remote studies due to reduced control of the participant setup and room configuration.

### **What is temperament and how could it relate to dropout rates in developmental studies?**

According to Rothbart and Bates (2006) temperament is the “behavioral expression of underlying neurobiological substrates” (p.106). Temperament contributes to individual differences across emotions, activity levels, attention, regulation, and reactivity. Moreover, temperament may be influenced over time by genetic, epigenetic, and/or non-genetic experiences (Rothbart & Bates, 2006). Rothbart and colleagues (Rothbart et al., 2006; Rothbart & Bates, 2006) propose three dimensions of temperament: 1) Effortful Control, 2) Surgency/Extraversion, and 3) Negative Affectivity. Each dimension will be described in more detail below.

#### **Effortful control**

Effortful control is related to the regulatory component of temperament. Effortful control encompasses the child’s ability to regulate behaviors, demonstrate executive attention skills, and demonstrate the ability to inhibit a dominant response (Rothbart & Bates, 2006). Effortful control also involves the ability to voluntarily focus and shift attention, and/or regulate emotions (Caspi et al., 2005; Rothbart et al., 1998). This dimension involves the increasing regulatory ability to control motor movements and behaviors, in part due to maturation of attention systems (Posner & Rothbart, 1992). For example, high effortful control involves the ability to suppress inappropriate responses, demonstrate better self-regulation, and maintain focus on task-related activities. Effortful control also involves attentional focusing, inhibitory control, low intensity pleasure, and perceptual sensitivity. Putnam and Stifter (2002) suggest that inhibitory control improvements are seen between 6 and 12 months, become more refined by 18 months, and

continue to advance between 24 and 36-months (Mezzacappa, 2004; Rothbart & Bates, 2006). Although the ability to inhibit behaviors will improve significantly in the third and fourth year, these skills will continue to improve throughout childhood (Rothbart et al., 2006). Research shows higher levels of effortful control is linked to higher social competence, conscience, and adjustment for young children (Eisenberg et al., 2001; Eisenberg et al., 2005). Higher levels of effortful control skills promote positive outcomes later in life, such as lower negative emotion, increased compliance, and better adjustment (Eisenburg et al., 2001; Kochanska & Knaack, 2003). For example, Eisenberg and colleagues (2007) suggest that children with better behavioral control are more likely to engage in socially appropriate behaviors with their peers.

Effortful control has also been tied to early attention skills. For example, Kochanska and colleagues (2000) found that focused attention at 9 months was predictive of effortful control at 22 months. As previously mentioned, refining executive attention (i.e., voluntary deployment of attention) is a key component of effortful control. Nakagawa and Sukigara (2013) investigated temperament and attention longitudinally in toddlers from 12 to 36-months. In their study, they found a positive relationship between effortful control and sustained attention. This finding supports Ruff and Rothbart (1996) in that effortful control is related to a child's refining ability to exhibit sustained attention.

Effortful control encompasses subscales including inhibitory control, which involves the "ability to suppress inappropriate approach responses under instructions or in novel or uncertain situations" and attentional focusing "maintaining attentional focus upon task-related channels" both of which are important when completing a remote study (Wolfe & Bell, 2007, 433). For example, when asked to sit still and watch a video, (especially in the context of their home environment where they are surrounded by siblings, toys, pets, or other distractors) toddlers must

suppress these dominant responses in order to follow instructions, and sit through the duration of the video. Thus, effortful control is an important factor that may contribute to a toddler's ability to complete the task and thus be included in the final sample.

### **Surgency**

Surgency encompasses an individual's disposition related to positive emotions, reward seeking, impulsivity, reactions to stimulus intensity, sociability, and level of comfort in social situations (Putnam et al., 2006; Rothbart et al., 2011). This dimension is known as the "approach system" and encompasses activity level, sociability, and positive affect (e.g., smiling, laughing). Research suggests that aspects of surgency emerge around 2-3 months and the expression of positive emotions, for example, increase during the first year (Rothbart et al., 2000). During this time, activity level, smiling, laughing, vocal reactivity, and perceptual sensitivity can be used to measure surgency. Some aspects of surgency have been linked to positive outcomes later in childhood, such as sociability or social competence. For example, some children will be more sociable or impulsive than their peers. Children who are lower in surgency will often display more shyness (Rubin et al., 1999) and lower levels of social competence (Fox & Calkins, 1993). Research has shown that surgency is also negatively related to attention and effortful control (Rothbart & Rueda, 2005).

Surgency also plays an important role in mitigating (or augmenting) internalizing and externalizing behaviors. In a sample of toddlers from low-income families with young mothers and at risk for contextual risk, surgency and internalizing problems were investigated (Eisenberg et al., 2009). In their sample, higher levels of surgency predicted fewer internalizing problems (e.g., unhappy, sad, depressed, sulks a lot, etc.). Research also indicates that children lower in surgency are at greater risk of experiencing internalizing behaviors (Nilzon & Palmerus, 1998).

On the other hand, some surgent behaviors may not be as beneficial across all circumstances, as some highly surgent children may struggle to control their impulsivity (Rothbart et al., 2000). Rigato and colleagues (2022) recently found that infant surgency in the first year was later associated with hyperactivity-inattention at 36 months. In older age groups, high levels of surgency have been shown to elicit more negative perceptions from same-aged peers at age four (if the children are lacking in self-regulatory skills) (Dollar & Stifter, 2012). Highly surgent children were also rated as more aggressive by their mothers. The highly surgent children who used more self-soothing techniques (e.g., thumb sucking, rocking, etc.) were rated lower in social competence by their mothers. These findings highlight the nuances in surgency across age and context. That is, in terms of social competencies and internalizing and externalizing behaviors, surgency may be beneficial for children. On the other hand, too much surgency without self-regulatory skills may lead to increased impulsivity and decreased social competence. In the case of the current study, highly surgent children without self-regulatory skills may have greater difficulties following instructions or sitting in front of a screen for an extended period of time.

### **Negative affect**

Negative affect encompasses mood, anger, fear, and soothability. Also known as the “avoidance system” this dimension also includes frustration, discomfort, sadness, and falling reactivity. Negative affectivity can be measured early in life around three months of age (Gartstein and Rothbart, 2003). Negative affectivity has been linked to many outcomes later in life such as self-regulation and internalizing and externalizing symptoms (Gartstein et al., 2012; Putnam et al., 2008).

Negative affectivity is linked to self-regulation in toddlerhood in that increased negative affect is related to decreased self-regulation (Putnam et al., 2008; Raikes et al., 2007). Higher levels of negative affectivity at 10 and 24-months is linked to lower levels of effortful control (but only for toddlers with right frontal EEG asymmetry) (Smith et al., 2016).

Northerner and colleagues (2016) found that higher levels of negative affectivity predicted externalizing problem behaviors for 24-month-olds such as “hitting others” or “demonstrating disobedience”. These researchers also found that low negative affectivity acted as a protective factor against cumulative risk factors (e.g., neighborhood dangerousness, low social support) at 24-months. Here, the authors suggest negative affectivity plays a significant role in early child behavior problems. Rigato and colleagues (2022) also found that infant negative affectivity (measured in the first year) predicts child conduct problems at 36-months. Thus, dropout rates may be more prevalent for children who experience higher levels of negative affectivity than those with lower levels of negative affectivity, especially considering high negative affectivity is negatively associated with self-regulation skills (Bridgett et al., 2013; Putnam et al., 2008).

### **How could facets of temperament relate to completion rates?**

Exploring these three facets of temperament and their relationship to success or failure in research protocols should be done individually and in interaction with each other. Few prior studies have been conducted to investigate whether temperament plays a role in completion rates in infant and toddler studies. From the few that have been conducted, they have provided mixed results. Segal and colleagues (2021) investigated whether temperament predicted completion rates for 3.5-month-olds, and then at 7 months across four tasks (free-play, still-face, eye tracking, ERP). The authors did not find that temperament predicted completion rates at either

age, and suggested that differences in temperament are not systematically related to study completion rates. In contrast, Mink and others (2013) found that for 6 and 12-month-olds, temperament may influence dropout for tasks presented at the end of a session. For example, they found that if an ERP was conducted at the end of a session, the infants were more likely to drop out than if it was first. They also found that Duration of Orienting (attention to and/or interaction with a single object for extended periods of time) had the most impact on completion rates. That is, lower levels of duration of orienting corresponded to decreased completion rates. Klein and Radukic (2015) evaluated temperament and dropout rates at 6 and 9 months. They found that specific components of temperament were related to dropout rates at each time point. They found that compared to “completers” of the study, the “non-completers” were more likely to fuss and cry during the test session, and were slower to recover from distress or excitement. Unlike the findings by Mink and colleagues, (2013) they did not find a relationship between Duration of Orienting and completion rates. Furthermore, Rothbart and colleagues (2003) noted in their study with 2 to 3-year-olds that toddlers with lower effortful control and higher negative affect scores were not included in the final analysis due to differences in completion rates (i.e., these toddlers did not complete at least 50% of the trials across each task).

Investigating behavioral patterns is essential when considering the likelihood of a toddler not completing a remote study due to fussing out, physically moving away from the computer, or playing with their toys (to name a few examples). As outlined above, the three facets of temperament that are commonly measured (Effortful Control, Surgency/Extraversion, and Negative Affectivity) all show important correlations to behavior, reactivity, and regulatory skills in toddlerhood. Although some toddlers will fuss out due to random factors, some dropout may be more predictable. For example, in Klein and Radukic’s (2015) findings, they show that



infants that dropped out were more likely to fuss and cry, or were slower to recover from distress or excitement. Thinking about the three dimensions of temperament, could these two examples mentioned above map onto the negative affectivity or surgency dimensions? Understanding the nuances in performance and likelihood for toddlers to complete a remote session is important to understand if there are systematic biases in the reported final samples of infants and toddlers, and for future work to better understand how to support children with various temperamental characteristics to mitigate biased or unreliable estimates.

Unique to the current study, completion rates will be investigated in older toddlers (e.g., 18 and 24-months). Up to this point, there have been no studies that have investigated temperament (or specific dimensions of temperament) and dropout rates at 18-24-months of age. This study will investigate the relationship between completion rates and toddler temperament in an effort to more thoroughly understand 1) if samples are representative and generalizable on a person-centered basis, and 2) to understand if there are some toddlers who may be more likely to complete a study. With refining attention skills, refining language skills, increased mobility, and many more factors that have shifted since early in the first year, it is important to consider and contextualize completion rates in this age group. Additionally, due to the nature of remote studies, there are numerous distractors present for the toddler at any given moment (e.g., siblings, toys, pets) that make remote work more challenging than in-person studies, where there is less uncertainty in the participant's immediate environment. Nonetheless, remote research has largely become more prominent since the COVID-19 pandemic and has opened new doors for data collection that is not possible in a lab setting. Remote work has widened the demographic pool for researchers (Nielsen et al., 2017), as well as created improved accessibility and flexibility for participants. Given these benefits and increased implementation, it will be

important to understand if some toddlers may be more likely to be excluded on the basis of their likelihood of completing the task for remote studies.

### **Aims of the Current Study**

The current study aimed to address potential systematic biases in study completion rates in a remote study of toddlers ages 18 and 24 months. This study investigated potential differences between participants who complete a remote study vs those who do not (i.e., participants who do not complete a full test session). Completion was investigated in regards to the three dimensions of temperament: Effortful Control, Surgency/Extraversion, and Negative Affectivity (Putnam, Gartstein, & Rothbart, 2006). There have been no studies that have investigated systematic biases of final sample characteristics at a person-centered level for this age group. Although there have been a small number studies conducted to investigate temperament and completion rates, these studies have been conducted with infants 12 months and younger, and for different methodological approaches (e.g., fNIRs, EEG; Baek et al., 2021; Stets et al., 2012). The current study aimed to fill this gap to understand whether facets of temperament may influence how a toddler participates in a remote test session and whether they complete the test session. In the current study, it was hypothesized that toddlers' 1) effortful control would be positively correlated with their task completion, whereas surgency and negative affectivity would be negatively correlated with their task completion; and 2) effortful control would moderate the relationship between both surgency, negative affectivity, and completion. That is, a toddler with high levels of surgency (e.g., impulsivity, activity level) may complete the session if they also demonstrate higher levels of effortful control. This same concept applies to negative affectivity, in that children with high levels of negative affectivity (e.g., shyness, discomfort) may not drop out if they also demonstrate higher levels of effortful control.

## Method

### Participants

There were two separate samples of toddlers involved in this study. In the cross-sectional sample, a total of 17 toddlers participated: 6 toddlers (3 = females) were tested at 18-months and 11 toddlers (6 = females) were tested at 24-months. In the longitudinal sample, there were a total of 11 participants (9 = females) that were first tested at 18-months and then again at 24-months. Thus, analyses were conducted for the cross-sectional sample of participants, and when applicable, within age (i.e., 18, 24-months) analyses were conducted that included the longitudinal sample (see Table 1; total sample size was  $n=28$ ). A power analysis informed by prior effect sizes from two studies on attrition and temperament in toddlers (Klein-Raducik & Zmyj, 2015; Mink et al., 2013) determined that a minimum of 36 participants would be required for the current analyses. In total, 32 additional participants completed the initial screener form but did not participate in the current study, so we were not able to reach our target sample size.

Participants were recruited remotely and were able to participate from anywhere in the United States. Recruitment took place via email listservs, social media, (e.g., Facebook, Twitter, Instagram) and snowball sampling. Participants were monolingual English speakers with no hearing or vision impairments (by parental report).

Table 1. *Cross Sectional and Longitudinal Participant Demographics*

	Cross-Sectional Participants		Longitudinal Participants	
	<i>n</i>	%	<i>n</i>	%
Age				
18-month-olds				
Female	3	50.0	9	81.2
		50.0	2	18.2

---

Male	3			
24-month-olds		54.5	9	81.2
Female	6	45.4	2	18.2
Male	5			

### Protocol and Measures

**Video Events.** Each participant watched two short movies on their parent’s computer: The Multisensory Attention Assessment Protocol (MAAP) and the Intersensory Processing Efficiency Protocol (IPEP). These movies were shown via Gorilla Experiment Builder. The mother received a direct link in the Zoom meeting and was instructed to share her screen, click the link, and maximize her screen. These movies are dynamic (i.e., they contain moving visual images and sound-tracks), each comprising a series of short videos interspersed with “attention getter” clips. Both the MAAP and IPEP have two condition types: 1) social and 2) non-social. For the MAAP, there were two trial types: 1) low competition (no distractor) and 2) high competition (with distractor) (see Figure 1 for an example of each MAAP condition). On the MAAP, each trial features two lateral events in which only one event is synchronous to the soundtrack. On the IPEP, each trial features six events in which only one event matches the soundtrack (see Figure 2 for an example of each IPEP condition). For both the MAAP and IPEP, the social condition involves women telling stories in child-directed speech. The non-social condition involves small objects (e.g., blocks) striking a surface and their sounds. The MAAP consists of 32 trials, with 16 blocks of social (i.e., 8 with no distractor; 8 with a distractor) and 16 blocks of non-social trials. The IPEP consists of 48 trials with 12 blocks of social and 12 blocks of non-social trials. For the IPEP, once all 24 trials are presented for both social and non-social, the blocks repeat. Prior to each trial, a dynamic, silent attention getter appears in the center of the screen for 3 seconds. For half of the trials on the MAAP (high competition) the attention getter remains on the screen in between each lateral event. For the other half of the trials on the MAAP

(low competition) the attention getter does not remain on the screen. The MAAP and IPEP each have four orders and are counterbalanced.



*Figure 1.* Multisensory Attention Assessment Protocol (MAAP): Social trials (left) and nonsocial trials (right); low competition (top row) and high competition (bottom row). The soundtrack matched only one of the two lateral events within each trial. The match was counterbalanced on the left and right across trials.



*Figure 2.* Intersensory Processing Efficiency Protocol (IPEP): Social trials (left) and nonsocial trials (right). The soundtrack matched only one of the six events within each trial. The location of the match was counterbalanced across trials.

**Procedure.** Prior to the appointment, the parent completed the consent form via REDCap. Once completed, the remaining documents were delivered electronically via REDCap for the parent to complete prior to the moderated remote session. During the remote appointment, the parent first discussed the procedure with the experimenters to calculate internet speed, measure their screen size, and review seating positions for the toddler. After the researchers ensured the parent had no further questions and had reviewed the protocol, the parent was asked if their child would be more comfortable sitting on their own or in their lap. Then, the child was positioned either on their lap ( $n=14$  at 18-months and  $n=17$  at 24-months) or by themselves facing the computer screen ( $n=3$  at 18-months and  $n=5$  at 24-months). Each child watched the MAAP and IPEP movies on their parents computer screen. Both the MAAP and IPEP were counterbalanced and each had 4 orders. The participants then completed a 10-minute free play in between the two protocols. After finishing the first protocol, the parents were then instructed to “play with your child as you normally would” while being positioned in front of the computer screen. The parents provided three soft, silent toys for this play episode. The entire session was recorded on Zoom.

**Temperament.** To investigate temperament, parents completed the Early Childhood Behavior Questionnaire - Short Form (ECBQ; Putnam, Gartstein, & Rothbart, 2006). This scale is most appropriate for this age group (i.e., 18-24 months) and was created for children ages 1.5 to 3 years. Parents completed the form electronically via REDCap. This form has 107 items and parents responded on a 7-point likert scale ranging from 1= never to 7 = always. From this scale, 18 dimensions and three factors were calculated. The 18 dimensions include: Discomfort, Fear, Motor Activation, Sadness, Perceptual Sensitivity, Shyness, Soothability, Frustration, Impulsivity, Activity Level, High-Intensity Pleasure, Sociability, Positive Anticipation,

Inhibitory Control, Attention Shifting, Low-Intensity Pleasure, Cuddliness, and Attention Focusing. These 18 dimensions load onto 3 higher order factors: Effortful Control, Surgency/Extraversion, and Negative Affectivity. Effortful Control includes the sum of Inhibitory Control, Attention Shifting, Low-Intensity Pleasure, Cuddliness, and Attention Focusing. For Surgency/Extraversion, scores are averaged for Positive Anticipation, Sociability, High-Intensity Pleasure, Activity Level, and Impulsivity. Negative Affectivity is calculated by averaging the scores from Discomfort, Fear, Sadness, Frustration, Soothability (reversed) Motor Activation, Perceptual Sensitivity, and Shyness. Overall, this scale demonstrates good internal consistency. Moreover, each dimension (of the 18) demonstrates good internal consistency at both 18 and 24-months (with the exception of impulsivity at 18 months) and show stability (via caregiver ratings) between 18 and 36-months (with the exception of positive anticipation).

**Quantifying Protocol Performance.** First, two different classifications for completion profiles were created. In one, toddlers were placed in one of two groups depending on whether they sat through both of the protocols or not. In the second, toddlers were placed in one of three groups depending on whether they successfully sat through zero, one, or both protocols (i.e., the entire MAAP sequence and/or the entire IPEP sequence). If the participant sat through the full protocol, the toddler was placed into the “completer” group. If they sat through one of the two protocols, they were placed in the “partial completer” group. If they sat through neither of the two protocols, they were placed in the “non completer group.” These two variables are referred to as completion rates below.

Second, total duration of participation (total number of trials completed) and proportion of attention (divided by total trials they completed) were calculated. These two values were calculated because many studies use the total number of trials completed as a guideline to

determine which toddlers to keep in the final analyses and could help target completion rates and exclusion on the basis of trials completed. The proportion of attention value was created to create a more nuanced understanding of participation. This value would help differentiate children who “completed” 32/32 trials, but only looked during 12 of those trials. This value was essential because many toddlers would glance away from the screen throughout their test session. So, while these participants were not looking at the screen for the full duration of the test session, they would often return their attention back to the screen. The test session would continue until the toddler became fussy or physically moved away from the screen. The experimenter would then unmute herself and ask the mother to exit the protocol.

### Results

Preliminary analyses were conducted to examine if there were any sex or order effects in the cross-sectional sample. An Independent t-test was conducted to investigate if there were any sex-related differences on completion rate, total duration of participation (for both MAAP and IPEP), and proportion of attention (for both MAAP and IPEP). There were no significant sex effects. A Multivariate General Linear Model was then conducted to investigate if there were any effects of order. There were no significant differences, thus remaining analyses were collapsed across order and sex. Correlations were conducted between each higher-order factor to investigate if the three factors were significantly related (see Table 2). There were no significant relationships between the higher-order factors.

Table 2. Correlations between ECBQ Temperament Higher-Order Factors for Cross-Sectional Sample (18 and 24-month-olds included) (sample size is within the parentheses)

<i>ECBQ Dimension</i>	1	2	3
Effortful Control	-		



Surgency	.18	-	
	(n=17)		
Negative Affectivity	-.10	-.01	-
	(n=17)	(n=17)	

---

All  $ps > .05$

### **Effortful Control, Surgency, & Negative Affectivity and Completion Rates**

First, to determine whether facets of temperament influenced completion rates, Effortful Control, Surgency, and Negative Affectivity were regressed onto completion in two separate ways. For the logistic regression, temperament was regressed onto the binary classification for completing both protocols (MAAP, IPEP) or not. Here, temperament did not significantly influence completion rate ( $R^2 = .05, p = .83$ ) (see Table 3).

Table 3. Logistic Regression Predicting Completion Rate

Predictor	<i>B</i>	SE	$\beta$	<i>p</i>
Intercept	-.24	7.27	.08	.73
Effortful Control	-.39	1.07	.67	.71
Surgency	.82	1.00	2.27	.41
Negative Affectivity	-.25	.65	.77	.70

---

$n = 17, (R^2 = .05, p = .83)$

For the linear regression, Effortful Control, Surgency, and Negative Affectivity were regressed onto the completion rate that indicated whether the toddler completed both protocols, one of the two protocols, or zero protocols. This regression was not statistically significant for the cross-sectional group of toddlers ( $R^2 = .08, F(3, 13) = .42, p = .75$ ).

Table 4. Linear Regression Predicting Completion Rate

Predictor	<i>B</i>	SE	$\beta$	<i>p</i>
Intercept	-.85	2.76	--	.76
Effortful Control	-.12	.41	-.08	.76
Surgency	.41	.37	.29	.28
Negative Affectivity	.01	.26	.01	.94

$n = 17, R^2 = .08, F(3, 13) = .42, p = .75.$

Correlations between each of the three facets of temperament were calculated with MAAP total duration of participation and IPEP total duration of participation and MAAP/IPEP proportion of attention. For the cross-sectional participants only, there were no significant correlations between any temperamental factors and total duration of participation for MAAP or IPEP ( $n = 17$ ).

On average, toddlers completed 21.70 out of 32 trials for MAAP, and 38 out of 48 trials for IPEP. For the cross-sectional sample (excluding longitudinal participants) 18-month-olds ( $n = 6$ ) watched an average of 13.66 trials for MAAP and 33.66 trials for IPEP whereas 24-month-olds ( $n = 11$ ) watched an average of 26.09 trials for MAAP and 40.60 trials for IPEP. After conducting an Independent Samples T-Test, there was a significant difference between 18-month-olds and 24-month-olds total completion  $t(15) = -1.16, p = .004$ . Thus, analyses were conducted within age for both 18-month-olds and 24-month-olds (i.e., the cross-sectional and longitudinal samples were combined at each age). Below, I present analyses on the relationship between temperament and task performance within each age group for each of the two protocols

(see Table 5 for average temperament scores on the main three higher-order dimensions along with their lower-order dimensions).

**Table 5. Means & Standard Deviations of ECBQ Dimensions**

<i>ECBQ Dimension</i>	<i>18 Months (n=17)</i>		<i>24 Months (n=22)</i>	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Effortful Control	4.93	0.34	5.00	0.50
Inhibitory Control	4.10	0.64	4.45	0.70
Attention Shifting	4.87	0.51	5.05	0.63
Low-Intensity Pleasure	5.40	0.90	5.18	0.86
Cuddliness	5.35	0.96	5.48	0.64
Attention Focusing	4.88	0.85	4.86	0.80
Surgency	5.03	0.60	4.96	0.55
Positive Anticipation	5.04	1.07	5.69	0.71
Sociability	5.87	0.95	5.88	1.14
High-Intensity Pleasure	4.91	1.22	4.80	0.98
Activity Level	4.65	0.78	4.35	0.74
Impulsivity	4.46	0.88	4.03	1.03
Negative Affectivity	3.58	0.85	3.29	0.57
Discomfort	2.09	0.92	2.27	0.76
Fear	2.50	0.66	2.21	0.58
Sadness	3.25	0.70	2.71	0.77
Frustration	3.79	0.75	3.43	0.83
Soothability	5.41	0.79	5.34	0.65
Motor Activation	2.35	0.69	2.22	0.82
Perceptual Sensitivity	4.28	0.98	3.98	0.94
Shyness	3.88	1.20	4.25	1.35

**MAAP x 18-month-olds.** Correlations were conducted to determine if any of the three temperament facets were significantly related to MAAP total duration of participation or looking proportion ( $n = 17$ ). There were no significant correlations with either of these performance measures (see Table 6).

**MAAP x 24-month-olds.** The same analyses were conducted for 24-month-olds ( $n = 22$ ).

Surgency was significantly positively correlated with MAAP total duration of participation ( $r =$

+ .45,  $p = .04$ ). However, there were no other significant correlations for Effortful Control or Negative Affectivity and total duration of participation.

**IPEP x 18-month-olds.** Correlations between effortful control, surgency, negative affectivity, and IPEP total duration of participation and proportion of attention were calculated ( $n = 17$ ). For 18-month-olds, Surgency and IPEP total duration of participation were significantly, positively correlated ( $r = +.53, p = .04$ ). There were no other significant relationships.

**IPEP x 24-month-olds.** The same correlations were conducted for both IPEP total duration of participation and IPEP proportion of attention ( $n = 22$ ). There were no significant correlations.

Table 6. Correlations between ECBQ Temperament, MAAP and IPEP Total Duration of Participation, and MAAP and IPEP Total Looking Proportion for 18-month-olds and 24-month-olds (sample size is within the parentheses)

<i>ECBQ Dimension</i>	MAAP Total Duration of Participation	MAAP Total Looking Proportion	IPEP Total Duration of Participation	IPEP Total Looking Proportion
Effortful Control				
18-month-olds	-.28 (17)	-.27 (17)	.11 (17)	.09 (17)
24-month-olds	.03 (22)	.29 (22)	-.01 (22)	.01 (22)
Surgency				
18-month-olds	.22 (17)	.20 (17)	.47* (17)	.46 (17)
24-month-olds	.44* (22)	.28 (22)	.30 (22)	.30 (22)
Negative Affectivity				
18-month-olds	-.13 (17)	-.16 (17)	-.01 (17)	-.02 (17)
24-month-olds	.05 (22)	.12 (22)	.25 (22)	.25 (22)

$p < .05 = *$

### **Lower-Order Dimensions of Surgency Related to Total Duration of Participation and Proportion of Attention**

As surgency was positively correlated to total duration of participation for MAAP and IPEP, exploratory analyses were conducted to determine which components of surgency were correlated to task performance. Correlations were run between each of the five lower-order

dimensions of surgency with total duration of participation and looking proportion for each protocol separately (see Tables 7 and 8). Given that these were exploratory correlations, Holms-Bonferroni correction was applied to significance levels (Abdi, 2010).

**MAAP x 18-month-olds.** There were no significant relationships between lower-order dimensions and total duration of participation or MAAP looking proportion ( $n = 17$ ).

**MAAP x 24-month-olds.** At 24-months, MAAP total duration of participation was not significantly correlated with any lower-order dimensions ( $n = 22$ ). There were no other significant correlations.

**IPEP x 18-month-olds.** For the 18-month-olds, IPEP total duration of participation was significantly, positively correlated with high-intensity pleasure ( $r = +.67, p = .01$ ) ( $n = 17$ ).

Additionally, IPEP proportion of attention was positively correlated with high-intensity pleasure ( $r = +.61, p = .01$ ). There were no other significant correlations.

**IPEP x 24-month-olds.** For the 24-month-olds, IPEP looking proportion was also positively correlated with activity level ( $r = +.58, p = .02$ ) ( $n = 21$ ). There were no other significant correlations.

Table 7. Correlations between Lower-Order Dimensions of Surgency, MAAP and IPEP Total Duration of Participation, and MAAP and IPEP Total Looking Proportion for 18-month-olds

<i>Surgency Lower-Order Dimensions</i>	MAAP Total Duration of Participation	MAAP Total Looking Proportion	IPEP Total Duration of Participation	IPEP Total Looking Proportion
Positive Anticipation	.01	-.08	-.29	-.39
Sociability	-.01	.02	.21	.10
High-Intensity Pleasure	.14	.13	.67**	.61**
Activity Level	.15	.04	.33	.26
Impulsivity	.33	.35	.32	.27

$p < .05 = *$ ,  $p < .01 = **$

Table 8. Correlations between Lower-Order Dimensions of Surgency, MAAP and IPEP Total Duration of Participation, and MAAP and IPEP Total Looking Proportion for 24-month-olds

<i>ECBQ Lower-Order Dimensions</i>	MAAP Total Duration of Participation	MAAP Total Looking Proportion	IPEP Total Duration of Participation	IPEP Total Looking Proportion
Positive Anticipation	-.17	-.04	-.12	-.22
Sociability	.34	.33	.10	.03
High-Intensity Pleasure	.09	.01	.13	.16
Activity Level	.25	.03	.59	.58*
Impulsivity	.35	.10	.26	.28

$p < .05 = *$

### Discussion

Considering experimental attrition that early development researchers face, it is important to understand whether study completion rates may be a product of infants' and toddlers' temperamental factors. That is, are some toddlers fussing out of the study due to factors that are predictable of the toddler (e.g., distress to strangers)? A recent meta-analysis found that across 272 experiments with infants and toddlers from 0-24 months, average attrition rate was 34.23% (Baek et al., 2023). Out of the publications that reported specific reasons for why participants were excluded, 21.50% were subject-driven. Although these findings were in the context of functional near-infrared spectroscopy (fNIRS) experiments, the analysis also found that block/trial design and stimulus type significantly predicted subject-driven attrition rates, and that when comparing subject-driven attrition rates across fNIRS, EEG, and infant looking time studies, there were no significant differences across methodological approaches. Thus, they suggest that subject-driven attrition (of which is reported; 83% of studies in this meta-analysis reported information attrition) “might be similar across methods and literature”. Given the attrition rates (Miller, 2017) and sampling strategies used to overcome dropout in infant and toddler research (Nicholson et al., 2017), this question is important to understand whether final

samples are representative across toddlers and temperamental profiles. Understanding potential biases at a person-centered level will help promote replicability and generalizability (Yu et al., 2020). The purpose of the current study was to understand how temperament may be related to completion for toddlers participating in a remote study.

The three temperamental factors measured in the current study were 1) Effortful Control, 2) Surgency/Extraversion, and 3) Negative Affectivity. Research shows that toddlers with high levels of effortful control are better equipped to voluntarily focus and shift attention, and/or regulate emotions (Caspi et al., 2005; Rothbart et al., 1998). Moreover, research indicates that effortful control is positively correlated with sustained attention (Nakagawa & Sukigara, 2013; Ruff & Rothbart, 1996) and compliance during a structured task (Kuo & Braungart-Reiker, 2022). Studies have also shown that shorter attention spans and the likelihood of the toddler to become jittery during a task may also increase attrition (Marshall et al., 2009).

In the current study, it was hypothesized that toddlers' 1) effortful control would be positively correlated with their task completion, whereas surgency and negative affectivity would be negatively correlated with their task completion; and 2) effortful control would moderate the relationship between both surgency, negative affectivity, and completion. These hypotheses were not supported in that no significant correlations were found between effortful control and any completion metric. This was also true for negative affectivity. However, results indicated significant, positive correlations between surgency and toddlers' protocol performance. Surgency was not correlated with completion or participation metrics for the cross-sectional sample. When these analyses were conducted within groups for 18 and 24-month-olds there were significant, positive (rather than the predicted negative) correlations between surgency and completion metrics. These results will be discussed in more detail below.

## **Effortful Control, Negative Affectivity, and Task Completion**

It was predicted that toddlers' effortful control would positively correlate with completion, and toddlers' negative affectivity would negatively correlate with task completion. It was also expected that effortful control would moderate the relationship between negative affectivity and completion. However, these hypotheses were not supported; there were no significant correlations between effortful control, negative affectivity, and protocol performance.

One potential explanation for why there was not a significant relationship between negative affectivity and protocol performance may be that in the current sample, variability for negative affectivity (and lower-order dimensions of negative affectivity) was low. Studies show that differences in parenting styles are related to negative affectivity (Al Van Den Akker et al., 2010), so it is possible that parenting styles were too homogenous in the current sample. Though parenting styles were not measured here, future studies should seek to understand more about these relationships and should obtain a larger, and more diverse sample to increase variability (Song et al., 2018; Spinrad et al., 2012). Negative affectivity (such as shyness or fear) has been shown to negatively correlate with attention and positively with fear, frustration, or withdrawal (Nakagawa & Sukigara, 2013; Posner & Rothbart, 2007). Research shows that negative affectivity is also inversely associated with self-regulation (Bridget et al., 2013; Raikes et al., 2007) which was expected to be an important component of test completion in the current task. This finding will be discussed below in the context of task engagement.

By 18-months, effortful control and the more fine-grained factors of effortful control can be measured (Putnam et al., 2006) and numerous studies have shown that effortful control plays an important role in attention and self-regulation during toddlerhood (Kochanska et al., 2000;



Nakagawa & Sukigara, 2013). Effortful control includes inhibitory control, attention shifting, low-intensity pleasure, cuddliness, and attention focusing. Effortful control is also known as the ability to inhibit a dominant response to perform a subdominant response (Diamond, 1990). Additionally, as discussed by Rothbart and colleagues (2005) it also creates opportunities to react in ways beyond “affectively driven behavior” and is linked to planning and inhibiting a response. From this perspective, why was effortful control not significantly correlated with any performance measure? One possible explanation is that the variability in effortful control across this sample of participants was not high enough to reveal correlation, and with a larger, more variable sample, this correlation would be significant. Another possible explanation extends from the engaging nature of the tasks themselves, which could minimize the need for inhibitory control. In the current task, many toddlers were highly engaged with the protocols employed (e.g., women telling stories, blocks falling). In fact, utilizing multimodal stimuli (which was implemented in the current study) has been shown to elicit greater attention (as opposed to audio-only or visual-only) and has been employed as one way to potentially reduce attrition in infant and toddler studies (Reynolds & Guy, 2012; Richards, 2003). Often while the movies played, the toddlers could be seen smiling and laughing at the screen during both MAAP and IPEP. Analyses indicated that on average, toddlers were looking at the screen more than 50% of the time for both MAAP (56%) and IPEP (66%). Completion and participation will now be discussed with respect to toddlers’ surgency scores as this dimension was significantly related to task completion for toddlers.

### **Surgency and Task Completion**

According to the ECBQ, surgency is a composite of multiple factors. It is a composite of sociability, impulsivity, high-intensity pleasure, positive anticipation, and activity level/energy.

Surgency is also known as the tendency to actively engage with one's environment (Oddi et al., 2013). In the present study, it was expected that distractors present in the home could make participation more difficult for toddlers higher in surgency (e.g., impulsivity, activity level) and lower in effortful control (e.g., attentional focus, inhibitory control) and that sitting for a duration of time in front of a screen would be more challenging. This hypothesis was not supported in the present study; the relationship between surgency and protocol participation was positive. These findings are similar to the recent findings of Moyano and colleagues (2022). In this study, it was predicted that effortful control would be positively related to attentional processes (i.e., endogenous orienting, context monitoring) whereas surgency and negative affectivity would negatively relate to these processes. Similar to the current study, these authors found no significant relationship between effortful control and attentional processes. However, they did find a significant, positive correlation between *surgency* and sustained attention across the task.

At this age, could the behavioral demand of sitting in front of a screen be less important to task completion than the active engagement that may be elicited from the task? That is, certain elements of surgency (such as positive anticipation or high-intensity pleasure) may translate to greater attention via more positive engagement. This participation is supported by the relationships between the lower-order dimensions of surgency and overall protocol performance. Significant correlations were found at 18-months between high-intensity pleasure and IPEP proportion of attention, and then at 24-months, activity level, positive anticipation, and IPEP proportion of attention. Considering that high-intensity pleasure involves enjoyment related to situations with high stimulus intensity, rate, complexity, and novelty, this may explain why significant relationships were found for IPEP looking patterns at both ages. Positive anticipation involves excitement about pleasurable activities. Both positive anticipation and high-intensity

pleasure involve components that relate to enjoyment and excitement about situations and activities. These measures may help to further explain differences in completion or participation based on differences in task engagement.

For example, one recent study investigated potential associations between maternal education, home literacy environment, and overall interest and engagement in literacy activities for toddlers ages 20 to 36 months to determine how these variables influenced vocabulary skills (Dicataldo & Roch, 2022). These authors found that the child's *interest and engagement* in the literacy activities significantly influenced vocabulary. These authors conclude that leveraging a toddler's interests can help promote task engagement and subsequent language skills at this age. Another study found that when conducting a sorting task with two-year-olds, the social interaction that was initiated by the toddlers was positively associated with performance (Antrilli & Wang, 2018). Underlying the findings from both of these studies is that the toddlers' initiation was positively associated with outcomes both for vocabulary skills and the social interactions.

Similar findings can also be seen when considering media usage for toddlers. For example, studies show that interactive media use can improve toddlers' experience and engagement when using screens compared to passive screen exposure (Kikorian et al., 2020). There have also been studies to show that joint media engagement or "screen scaffolding" can bolster screen time and related outcomes for toddlers (Arundell et al., 2020; Wood et al., 2016). Screen scaffolding involves the parent taking a more active role and participating while the toddler plays on a screen. This interaction is now a shared interaction as opposed to the child passively watching a video or playing a game independently. Moreover, Choi and Kikorian (2016) found that when toddlers are using screens, contingency can strengthen mental representations via increased engagement and arousal, but only for younger toddlers. These

authors suggested that for younger toddlers, guiding selective attention helped promote a greater outcome to successfully complete the task as opposed to increasing overall attention.

Studies show that toddlers learn more when they experience a live, reciprocal interaction (via video chat or in-person) as opposed to a non-reciprocal interaction (Golinkoff, 2014). An extension of this study determined that while one benefit of contingency is that “it increases engagement and enables self-pacing”, some types of contingency may be less effective for older toddlers (Kikorian et al., 2014). For younger toddlers, directing their attention can be particularly helpful to support less refined attention skills. In another study, Bruce and colleagues (2022) demonstrate enhanced attention facilitation (i.e., appropriateness of attention directing) from mothers during a 10-minute free play was predictive of vocabulary for 24-month-olds. Paired with these findings, could engagement act as a segue into toddlers attention and task performance?

Finally, guided by prior studies investigating completers vs. non-completers (Fagen et al., 1987; Mink et al., 2013; Wachs & Smitherman, 1985), the current study created additional variables to capture nuances in “completion”. As previously discussed, researchers have to define cutoffs when deciding which participants should be included in the final sample. Exclusion may involve number of trials completed (Klein-Radukic & Zmyj, 2012) total fixation time (Rothbart & Bates, 2003) and more. Thus, it is important to know whether certain criteria may be more likely to exclude participants due to person-centered factors. Here, it was determined that surgency was positively associated with total duration of participation for IPEP and looking proportions. Neither effortful control nor negative affectivity played a significant role in completion for MAAP or IPEP. The findings from Baek and colleagues (2022) suggest that age is an important factor when considering how to prevent attrition in infant and toddler

research. That is, researchers should consider how to balance engagement and compliance by “not making the task too long or boring for them to complete” (Baek et al., 2022). Future research should investigate what specific components of tasks such as these may help to promote engagement and compliance.

### **Limitations**

A few limitations should be noted in regards to the current study. First, research has shown that while parents’ scoring of their toddler’s temperament show high stability across age (Carranza Carnicero, et al., 2000) completion rate was evaluated during a one hour test session. This is important to note because although the experimenters asked the parent to schedule the testing session around eating or nap times, there was no guarantee the toddler would be prepared for the test session. For example, there were a few occasions during the test session that if a toddler became fussy, the parent would apologize and express that the toddler did not get a good night sleep, the child was sick, etc. These sessions could have been rescheduled to ensure the best and most accurate data could be obtained. However, studies have shown significant and positive relationships between parent-report of child temperament and tasks conducted in the lab for inhibitory control (Carlson & Moses, 2001; Morasch & Bell, 2010). Nonetheless, it is important to note that temperament and performance during the task may have been influenced by additional factors that could not have been controlled (e.g., if the toddler was not feeling well, if she was not well-rested). Studies show that parent reports of their child’s temperament may be influenced by additional factors such as maternal extraversion, marital stress, parental stress, or depression (Bayly & Gartsten, 2013). These variables were not measured in the current study, so these potential confounding factors could not be controlled in the current analyses. As mentioned in the Methods section above, the sample size of this study is underpowered for the kinds of

analyses that were planned. Future work will need to seek higher numbers of toddlers at each age to better assess relationships between performance and temperament.

High rates of attrition in developmental work may threaten the generalizability or replicability of these studies. Here, surgency was a significant factor when considering task completion and participation. Thus, this finding extends prior work by highlighting a potential avenue to prevent attrition to better understanding person-centered task completion and participation during a task. As researchers continue to test remotely and return to testing in-person, it will be important to determine if these findings replicate to in-person studies as well, or if these factors shift as a function of context across studies.

## References

- Abdi, H. (2010). Holm's sequential Bonferroni procedure. *Encyclopedia of Research Design*, 1(8), 1-8. <https://doi.org/10.4135/9781412961288>
- Antrilli, N. K., & Wang, S. H. (2018). Toddlers on touchscreens: immediate effects of gaming and physical activity on cognitive flexibility of 2.5-year-olds in the US. *Journal of Children and Media*, 12(4), 496-513. <https://doi.org/10.1080/17482798.2018.1486332>
- Arundell, L., Parker, K., Timperio, A., Salmon, J., & Veitch, J. (2020). Home-based screen time behaviors amongst youth and their parents: familial typologies and their modifiable correlates. *BMC Public Health*, 20, 1-11. <https://doi.org/10.1186/s12889-020-09581-w>
- Asendorpf, J. B., Van De Schoot, R., Denissen, J. J., & Hutteman, R. (2014). Reducing bias due to systematic attrition in longitudinal studies: The benefits of multiple imputation. *International Journal of Behavioral Development*, 38(5), 453-460. <https://doi.org/10.1177/0165025414542713>
- Baek, S., Marques, S., Casey, K., Testerman, M., McGill, F., & Emberson, L. (2021). Attrition rate in infant fNIRS research: A meta-analysis. *BioRxiv*. <https://doi.org/10.1371/journal.pone.0272438>
- Bayly, B., & Gartstein, M. (2013). Mother's and father's reports on their child's temperament: Does gender matter?. *Infant Behavior and Development*, 36(1), 171-175. <https://doi.org/10.1016/j.infbeh.2012.10.008>
- Bridgett, D. J., Oddi, K. B., Laake, L. M., Murdock, K. W., & Bachmann, M. N. (2013). Integrating and differentiating aspects of self-regulation: Effortful control, executive functioning, and links to negative affectivity. *Emotion*, 13(1), 47. <https://doi.org/10.1037/a0029536>

- Carlson, S. M., & Moses, L. J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development*, 72(4), 1032-1053. <https://doi.org/10.1111/1467-8624.00333>
- Caspi, A., Roberts, B. W., & Shiner, R. L. (2005). Personality development: Stability and change. *Annu. Rev. Psychol.*, 56, 453-484. <https://doi.org/10.1146/annurev.psych.55.090902.141913>
- Choi, K., & Kirkorian, H. L. (2016). Touch or watch to learn? Toddlers' object retrieval using contingent and noncontingent video. *Psychological Science*, 27(5), 726-736. <https://doi.org/10.1177/0956797616636110>
- Constantino, J. N., & Gruber, Charles, P. (2005). *The social responsiveness scale*. Los Angeles, CA: Western Psychological Services.
- Dicataldo, R., & Roch, M. (2022). How does toddlers' engagement in literacy activities influence their language abilities?. *International Journal of Environmental Research and Public Health*, 19(1), 526. <https://doi.org/10.3390/ijerph19010526>
- Dixon Jr, W. E., Lawman, H. G., Johnson, E. B., May, S., Patton, L. A., Lowe, A. K., & Snyder, C. M. (2012). Effects of exogenous and endogenous distracters on immediate and long-term recall in toddlers. *Infancy*, 17(5), 525-557. <https://doi.org/10.1111/j.1532-7078.2011.00090.x>
- Dollar, J. M., & Stifter, C. A. (2012). Temperamental surgency and emotion regulation as predictors of childhood social competence. *Journal of Experimental Child Psychology*, 112(2), 178-194. <https://doi.org/10.1016/j.jecp.2012.02.004>
- Eisenberg, N., Cumberland, A., Spinrad, T. L., Fabes, R. A., Shepard, S. A., Reiser, M., Murphy, B., Lasoya, S. H., & Guthrie, I. K. (2001). The relations of regulation and emotionality to



- children's externalizing and internalizing problem behavior. *Child Development*, 72(4), 1112-1134. <https://doi.org/10.1111/1467-8624.00337>
- Eisenberg, N., Hofer, C., & Vaughan, J. (2007). Effortful control and its socioemotional consequences. *Handbook of Emotion Regulation*, 2, 287-288.
- Eisenberg, N., Valiente, C., Spinrad, T. L., Cumberland, A., Liew, J., Reiser, M., Zhou, Q., & Losoya, S. H. (2009). Longitudinal relations of children's effortful control, impulsivity, and negative emotionality to their externalizing, internalizing, and co-occurring behavior problems. *Developmental Psychology*, 45(4), 988. <https://doi.org/10.1037/a0016213>
- Eisenberg, N., Zhou, Q., Spinrad, T. L., Valiente, C., Fabes, R. A., & Liew, J. (2005). Relations among positive parenting, children's effortful control, and externalizing problems: A three-wave longitudinal study. *Child Development*, 76(5), 1055-1071. <https://doi.org/10.1111/j.1467-8624.2005.00897.x>
- Fagen, J. W., Ohr, P. S., Singer, J. M., & Fleckenstein, L. K. (1987). Infant temperament and subject loss due to crying during operant conditioning. *Child Development*, 497-504.
- Fox, N. A., & Calkins, S. D. (1993). Pathways to aggression and social withdrawal: Interactions among temperament, attachment, and regulation: In KH Rubin & J. Asendorpf (Eds.), *Social Withdrawal, Inhibition, and Shyness in Childhood* (pp. 81–100).
- Gartstein, M. A., & Rothbart, M. K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior and Development*, 26(1), 64-86. [https://doi.org/10.1016/S0163-6383\(02\)00169-8](https://doi.org/10.1016/S0163-6383(02)00169-8)
- Gartstein, M. A., Putnam, S. P., & Rothbart, M. K. (2012). Etiology of preschool behavior problems: Contributions of temperament attributes in early childhood. *Infant Mental Health Journal*, 33(2), 197-211. [https://doi.org/10.1016/S0163-6383\(02\)00169-8](https://doi.org/10.1016/S0163-6383(02)00169-8)

- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60(1), 549-576. <https://doi.org/10.1146/annurev.psych.58.110405.085530>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, 466(7302), 29-29. <https://doi.org/10.1038/466029a>
- Igartua, J. J., & Hayes, A. F. (2021). Mediation, moderation, and conditional process analysis: Concepts, computations, and some common confusions. *The Spanish Journal of Psychology*, 24, e49. <https://doi.org/10.1017/SJP.2021.46>
- Johansson, M., Marciszko, C., Gredebäck, G., Nyström, P., & Bohlin, G. (2015). Sustained attention in infancy as a longitudinal predictor of self-regulatory functions. *Infant Behavior and Development*, 41, 1-11. <https://doi.org/10.1016/j.infbeh.2015.07.001>
- Klein-Radukic, S., & Zmyj, N. (2015). Dropout in looking time studies: The role of infants' temperament and cognitive developmental status. *Infant Behavior and Development*, 41, 142-153. <https://doi.org/10.1016/j.infbeh.2015.10.001>
- Kochanska, G., & Knaack, A. (2003). Effortful control as a personality characteristic of young children: Antecedents, correlates, and consequences. *Journal of Personality*, 71(6), 1087-1112. <https://doi.org/10.1111/1467-6494.7106008>
- Kochanska, G., Murray, K. T., & Harlan, E. T. (2000). Effortful control in early childhood: Continuity and change, antecedents, and implications for social development. *Developmental Psychology*, 36(2), 220. <https://doi.org/KM037//0O12-1649.36.2.220>
- Kuo, P. X., & Braungart-Rieker, J. M. (2022). Attachment configurations to mothers and fathers during infancy predict compliance, defiance, and effortful control in toddlerhood. *Early Childhood Research Quarterly*, 58, 188-197. <https://doi.org/10.1016/j.ecresq.2021.09.006>

- Madigan, S., Browne, D., Racine, N., Mori, C., & Tough, S. (2019). Association between screen time and children's performance on a developmental screening test. *JAMA Pediatrics*, *173*(3), 244-250. <https://doi.org/10.1001/jamapediatrics.2018.5056>
- Mezzacappa, E. (2004). Alerting, orienting, and executive attention: Developmental properties and sociodemographic correlates in an epidemiological sample of young, urban children. *Child Development*, *75*(5), 1373-1386. <https://doi.org/10.1111/j.1467-8624.2004.00746.x>
- Miller, S. (2017). Infancy. *Developmental Research Methods* (pp. 195–232). SAGE Publications, Inc. <https://doi.org/10.1177/10775595221098029>
- Mink, D., Henning, A., & Aschersleben, G. (2013). Visual habituation tasks: The role of infant temperament. *Infant Behavior and Development*, *36*(3), 377-390. <https://doi.org/10.1016/j.infbeh.2013.03.005>
- Morasch, K. C., & Bell, M. A. (2011). The role of inhibitory control in behavioral and physiological expressions of toddler executive function. *Journal of Experimental Child Psychology*, *108*(3), 593-606. <https://doi.org/10.1016/j.jecp.2010.07.003>
- Moyano Flores, P. S., Conejero Barbero, Á., Fernández Cabezas, M., Serrano Chica, F., & Rueda Cuerva, M. D. R. (2022). Development of visual attention control in early childhood: Associations with temperament and home environment. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2022.1069478>
- Nakagawa, A., & Sukigara, M. (2013). Individual differences in disengagement of fixation and temperament: Longitudinal research on toddlers. *Infant Behavior and Development*, *36*(4), 728-735. <https://doi.org/10.1016/j.infbeh.2013.08.001>

- Nicholson, J. S., Deboeck, P. R., & Howard, W. (2017). Attrition in developmental psychology: A review of modern missing data reporting and practices. *International Journal of Behavioral Development, 41*(1), 143-153. <https://doi.org/10.1177/0165025415618275>
- Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology, 162*, 31-38. <https://doi.org/10.1016/j.jecp.2017.04.017>
- Nilzon, K. R., & Palmerus, K. (1998). Anxiety and withdrawal of depressed 9-11 year olds three years later: A longitudinal study. *School Psychology International, 19*(4), 341-349. <https://doi.org/10.1177/0143034398194005>
- Northerner, L. M., Trentacosta, C. J., & McLear, C. M. (2016). Negative affectivity moderates associations between cumulative risk and at-risk toddlers' behavior problems. *Journal of Child and Family Studies, 25*(2), 691-699. <https://doi.org/10.1007/s10826-015-0248-x>
- Oddi, K. B., Murdock, K. W., Vadnais, S., Bridgett, D. J., & Gartstein, M. A. (2013). Maternal and infant temperament characteristics as contributors to parenting stress in the first year postpartum. *Infant and Child Development, 22*(6), 553-579. <https://doi.org/10.1002/icd.1813>
- Posner, M. I., & Rothbart, M. K. (1992). Attentional mechanisms and conscious experience. In A. D. Milner & M. D. Rugg (Eds.), *The neuropsychology of consciousness* (pp. 91-111). Academic Press. <https://doi.org/10.1016/B978-0-12-498045-7.50010-4>
- Putnam, S. P., & Stifter, C. A. (2002). Development of approach and inhibition in the first year: Parallel findings from motor behavior, temperament ratings and directional cardiac response. *Developmental Science, 5*(4), 441-451. <https://doi.org/10.1111/1467-7687.00239>

- Putnam, S. P., Rothbart, M. K., & Gartstein, M. A. (2008). Homotypic and heterotypic continuity of fine-grained temperament during infancy, toddlerhood, and early childhood. *Infant and Child Development*, 17(4), 387–405. <https://doi.org/10.1002/ICD.582>
- Raikes, H. A., Robinson, J. L., Bradley, R. H., Raikes, H. H., & Ayoub, C. C. (2007). Developmental trends in self-regulation among low-income toddlers. *Social Development*, 16(1), 128–149. <https://doi.org/10.1111/j.1467-9507.2007.00375.x>
- Reynolds, G. D., & Guy, M. W. (2012). Brain–behavior relations in infancy: Integrative approaches to examining infant looking behavior and event-related potentials. *Developmental Neuropsychology*, 37(3), 210–225. <https://doi.org/10.1080/87565641.2011.629703>
- Rigato, S., Charalambous, S., Stets, M., & Holmboe, K. (2022). Maternal depressive symptoms and infant temperament in the first year of life predict child behavior at 36 months of age. *Infant Behavior and Development*, 67, 101717. <https://doi.org/10.1016/j.infbeh.2022.101717>
- Rothbart, M. K., & Bates, J. E. (2006). Temperament. In N. Eisenberg, W. Damon, & R. M. Lerner (Eds.), *Handbook of child psychology: Social, emotional, and personality development* (pp. 99–166). John Wiley & Sons, Inc. [https://doi.org/10.1007/978-1-4899-0643-4\\_5](https://doi.org/10.1007/978-1-4899-0643-4_5)
- Rothbart, M. K., Derryberry, D., & Hershey, K. (2000). Stability of temperament in childhood: Laboratory infant assessment to parent report at seven years. In V. J. Molfese & D. L. Molfese (Eds.), *Temperament and Personality Development Across the Lifespan* (pp. 85–119). Lawrence Erlbaum Associates Publishers.

- Rothbart, M. K., Ellis, L. K., Rosario Rueda, M., & Posner, M. I. (2003). Developing mechanisms of temperamental effortful control. *Journal of Personality*, 71(6), 1113-1144. <https://doi.org/10.1111/1467-6494.7106009>
- Rothbart, M. K., & Jones, L. B. (1998). Temperament, self-regulation, and education. *School Psychology Review*, 27(4), 479-491.
- Rothbart, M. K., Posner, M. I., & Kieras, J. (2006). Temperament, attention, and the development of self-regulation. In K. McCartney & D. Phillips (Eds.), *Blackwell handbook of early childhood development* (pp. 338–357). Blackwell Publishing. <https://doi.org/10.1002/9780470757703.ch17>
- Rothbart, M. K., & Posner, M. I. (2005). Genes and experience in the development of executive attention and effortful control. *New Directions for Child and adolescent development*, 2005(109), 101-108. <https://doi.org/10.1002/cd.142>
- Rothbart, M. K., Sheese, B. E., Rueda, M. R., & Posner, M. I. (2011). Developing mechanisms of self-regulation in early life. *Emotion Review*, 3(2), 207-213. <https://doi.org/10.1177/1754073910387943>
- Rubin, K. H., Nelson, L. J., Hastings, P., & Asendorpf, J. (1999). The transaction between parents' perceptions of their children's shyness and their parenting styles. *International Journal of Behavioral Development*, 23(4), 937-957. <https://doi.org/10.1080/016502599383612>
- Ruff, H. A., & Rothbart, M. K. (2001). *Attention in early development: Themes and variations*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195136326.001.0001>
- Rutter, M., Bailey, A., & Lord, C. (2003). *The social communication questionnaire*. Los Angeles, CA: Western Psychological Services.

[https://doi.org/10.1007/978-1-4419-1698-3\\_1651](https://doi.org/10.1007/978-1-4419-1698-3_1651)

Segal, S. C., Marquis, A. R., & Moulson, M. C. (2021). Are our samples representative?

Understanding whether temperament influences infant dropout rates at 3 and 7 months.

*Infant Behavior and Development*, 65, 101630.

<https://doi.org/10.1016/j.infbeh.2021.101630>

Slaughter, V. & Suddendorf, T. (2007). Participant loss due to “fussiness” in infant visual paradigms: A review of the last 20 years. *Infant Behavior and Development*, 30(3), 505-514. <https://doi.org/10.1016/j.infbeh.2006.12.006>

Smith, C. L., Diaz, A., Day, K. L., & Bell, M. A. (2016). Infant frontal electroencephalogram asymmetry and negative emotional reactivity as predictors of toddlerhood effortful control. *Journal of Experimental Child Psychology*, 142, 262-273.

<https://doi.org/10.1016/j.jecp.2015.09.031>

Song, J. H., Miller, A. L., Leung, C. Y., Lumeng, J. C., & Rosenblum, K. L. (2018). Positive parenting moderates the association between temperament and self-regulation in low-income toddlers. *Journal of Child and Family Studies*, 27, 2354-2364.

<https://doi.org/10.1007/s10826-018-1066-8>

Stets, M., Stahl, D., & Reid, V.M., (2012) A meta-analysis investigating factors underlying attrition rates in infant ERP studies. *Developmental Neuropsychology*, 37(3), 226-252.

<https://doi.org/10.1080/87565641.2012.654867>

Wachs, T. D., & Smitherman, C. H. (1985). Infant temperament and subject loss in a habituation procedure. *Child Development*, 861-867.

- Wolfe, C. D., & Bell, M. A. (2007). Sources of variability in working memory in early childhood: A consideration of age, temperament, language, and brain electrical activity. *Cognitive Development, 22*(4), 431-455. <https://doi.org/10.1016/j.cogdev.2007.08.007>
- Wood, E., Petkovski, M., De Pasquale, D., Gottardo, A., Evans, M. A., & Savage, R. S. (2016). Parent scaffolding of young children when engaged with mobile technology. *Frontiers in Psychology, 7*, 690. <https://doi.org/10.3389/fpsyg.2016.00690>
- Van Den Akker, A. L., Deković, M., Prinzie, P., & Asscher, J. J. (2010). Toddlers' temperament profiles: Stability and relations to negative and positive parenting. *Journal of Abnormal Child Psychology, 38*, 485-495. <https://doi.org/10.1037/fam0000574>
- Yu, Y., Shafto, P., & Bonawitz, E. (2020). Inconvenient samples: Modeling biases related to parental consent by coupling observational and experimental results. *Open Mind, 4*, 13-24. [https://doi.org/10.1162/opmi\\_a\\_00031](https://doi.org/10.1162/opmi_a_00031)



## Supplemental Materials

### Dimensions of Temperament defined

- **Activity Level/Energy:** Level (rate and intensity) of gross motor activity, including rate and extent of locomotion.
- **Attentional Focusing:** Sustained duration of orienting on an object of attention; resisting distraction.
- **Attentional Shifting:** The ability to transfer attentional focus from one activity/task to another.
- **Cuddliness:** Child's expression of enjoyment in and molding of the body to being held by a caregiver.
- **Discomfort:** Amount of negative affect related to sensory qualities of stimulation, including intensity, rate or complexity of light, sound, texture.
- **Fear:** Negative affect, including unease, worry, or nervousness related to anticipated pain or distress and/or potentially threatening situations; startle to sudden events.
- **Frustration:** Negative affect related to interruption of ongoing tasks or goal blocking.
- **High-intensity Pleasure:** Pleasure or enjoyment related to situations involving high stimulus intensity, rate, complexity, novelty and incongruity.
- **Impulsivity:** Speed of response initiation.
- **Inhibitory Control:** The capacity to stop, moderate, or refrain from a behavior under instruction.
- **Low-intensity Pleasure:** Pleasure or enjoyment related to situations involving low stimulus intensity, rate, complexity, novelty and incongruity.
- **Motor Activation:** Repetitive small-motor movements; fidgeting.

- **Perceptual Sensitivity:** Detection of slight, low intensity stimuli from the external environment.
- **Positive Anticipation:** Excitement about expected pleasurable activities.
- **Sadness:** Tearfulness or lowered mood related to exposure to personal suffering, disappointment, object loss, loss of approval, or response to other's suffering.
- **Shyness:** Slow or inhibited approach and/or discomfort in social situations involving novelty or uncertainty.
- **Sociability:** Seeking and taking pleasure in interactions with others.
- **Soothability:** Rate of recovery from peak distress, excitement, or general arousal.

