What is Risky Sexual Behavior? It Depends on Who You Ask…

Serena Agterberg

Department of Psychology, University of Oregon

Honors Thesis

Advised by: Dr. Sanjay Srivastava & Bradley T. Hughes

Abstract

Many college students engage in sexual behavior with consequences, known as risky sexual behavior. The goals of this study are to identify themes in self-described risky sexual behavior in college students and to compare these themes to previous measures. To do this, a sample of 151 student participants wrote descriptive narratives of stories regarding risky sexual behavior. An inductive thematic analysis was then conducted and seven themes were identified: Unprotected Sex, Sex in Public, Technology, Emotionally Risky Behavior, Stranger Danger, Under the Influence, and Non-Consensual Behaviors. These themes varied somewhat from previous measures of risky sexual behavior. The results could assist in creating an accurate new measure of risky sexual behavior in college students and to inform risk education and reduction strategies.

What is Risky Sexual Behavior? It Depends on Who You Ask…

Many young adults begin exploring their identity as sexual beings during college. While this is a healthy and important part of growing up, it also means that some students engage in sexual behavior with consequences. This behavior can be referred to as *risky sexual behavior* (RSB). Consequences of risky sexual behavior may include contraction of sexually transmitted infections (STI), unintended pregnancy, and emotional or psychological distress. Understanding what behaviors fall under the category of "risky" is the first step to reducing the consequences of risk. In order to identify these behaviors, and subsequently develop effective risk reduction strategies, measurement techniques must be accurate and targeted to the population of interest. This can be accomplished by asking and listening to the needs and concerns of said population. While there are sexual behaviors that are objectively risky and lead to specific consequences, such as unprotected sex leading to possible STI transmission or unintended pregnancy, other risks are more subjective. Taking a participant-centered approach to RSB allows researchers to understand intentions and specific behaviors that they may not have considered previously. Additionally, a participant-centered approach better reflects the values of the population rather than those of the researchers which results in a better picture of what is actually occurring within the population.

What is Risky Sexual Behavior?

If a group of individuals were asked to list the top five sexual behaviors that are risky, it is likely that there would be some crossover, but also significant variation. One of the greatest challenges of implementing risk reduction strategies and identifying which sexual behaviors are of concern to students is defining what "risky sexual behavior" actually entails. Depending on the source, risky sexual behavior can include a variety of actions and inactions. Perhaps the most

common behavior that comes to mind when talking about risky sex is engaging in unprotected sexual intercourse. However, even "unprotected sex" is defined uniquely from study to study and person to person. Many identify unprotected sex as penetrative, vaginal sex between a cisgender man and woman with the "risky" aspect mostly referring to the risk of unwanted pregnancy (Jaspal, 2019), reflecting a heteronormative worldview. Others extend this definition to include vaginal sex without condoms and/or birth control, as well as anal and oral sex without protection from sexually transmitted diseases (Turchik & Garske, 2009). It should be noted that early literature on the subject focused almost solely on the contraction of HIV/AIDS and the sexual behaviors of homosexual men.

Beyond unprotected sex, previous literature does not agree on what else is considered "risky sexual behavior". Another relatively common theme is sexual activity while under the influence of drugs and/or alcohol (Turchik & Garske, 2009; Fino et al., 2021). However, it is up for debate whether consent is possible when someone is intoxicated. Planned Parenthood's website states that consent should be "freely given", made without the pressure or while under the influence of drugs or alcohol (Planned Parenthood). If this is the case, should sex while intoxicated fall into the category of "risky sexual behavior", "non-consensual sexual contact", or even "sexual assault"? Along with this, should non-consensual behavior be included in measures of risky sexual behavior? Some researchers seem to equate sexual health risking behaviors with sexual victimization, especially in the context of alcohol use before sex. However, the idea of "sexual risk taking" implies autonomous decision making so there may be an argument for omitting any behaviors that are not consensual, including sex under the influence of substances.

Previous research on RSB has focused nearly exclusively on health risking outcomes while ignoring other possible types of risk, such as emotional harm or damage to ones'

reputation. However, risky sexual behavior encompasses a wider range of behaviors dependent on personal experience and beliefs, making it hard to define. Additionally, as times change, new opportunities to engage in risky behaviors emerge. For example, "sexting", hook-up apps like Tinder, and "camming" or engaging in online sexual behaviors for money have all become relevant in the past 10 years. Previous measures of risky sexual behavior rarely take online behaviors into account. Social norms also change drastically and quickly. Therefore, sexual behaviors that were considered deviant and risky a decade ago may not be widely considered as risky in the current social world.

Measuring Risky Sexual Behavior

A variety of measures have been created to study participation in RSB, though the most frequently used measure specific to college students is the Sexual Risk Survey (SRS; Turchik & Garske, 2009). The SRS was created to fill a gap in sexual risk measures which previously focused on certain populations, such as homosexual men, or specific behaviors, such as sex while under the influence of alcohol. Students ($n= 72$) had limited input into which behaviors were included in the final measure and the items were planned before students were surveyed. Ultimately, the researchers decided which items to include based on a literature review, adaption of items from other measures of sexual risk, the student survey, as well as some behaviors the researchers decided to include based on their own beliefs.

Researchers then distributed a 37-item version to students ($n= 613$) to assess reliability and validity. It should also be noted that nearly all the students that were surveyed identified as White, heterosexual, and Christian. Some items were excluded based on low responses, though researchers chose to include some items that had few responses because they were deemed "useful". The final SRS included 23 items that fit into five factors: Sexual Risk Taking with

Uncommitted Partners, Risky Sex Acts, Impulsive Sexual Behavior, Intent to Engage in Risky Sexual Behaviors, and Risky Anal Sex Acts. Intent to Engage in RSB was a factor that researchers included to capture students who had not yet engaged in sexual behavior. Risky Anal Sex Acts was psychometrically weak, though it was retained for use in "more diverse populations" (SRS; Turchik & Garske, 2009). One reason this may be included as a separate factor from general Risky Sex Acts is due to latent heteronormativity and subconscious homophobia which is a remanent from older RSB studies that treated any homosexual sex act as risky.

The SRS was developed between 2006 and 2009 and likely does not capture the attitudes and beliefs of today's college students. The measure has not been updated to reflect changing times, though it still functions as the most popular measure of risky sexual behavior specific to college students and is widely used by researchers in various fields. It is apparent that the morals and beliefs of the researchers biased the choices made when deciding which behaviors to include in the final measure. Other measures of RSB are similarly outdated and include theorizing about risk rooted in moralization. These biases disallow researchers to collect accurate data about sexual risk-taking behaviors. Developing a measure of RSB specific to college students that centers the voices of this population without moral constraints from the researcher is imperative to understanding and reducing the consequences of risk.

Present Research

Previous theories and measures of RSB are outdated, rooted in heteronormative and/or homophobic beliefs, and based on the morals and values of researchers. So, what is the alternative? A quantitative study in which students rate how risky they consider various sexual behaviors on a scale could be useful. However, this would still not allow the students' voices to

be heard nor would it center their perspectives. Instead, this study uses a qualitative approach to allow participants to define risk in their own language with the researcher imposing fewer assumptions and constraints than a structured questionnaire. The research question that guides this study is: What sexual behaviors do college students consider to be risky? While this study takes an exploratory approach, there are two goals that we focused on:

G1: Identify themes in self-described sexually risky behavior in college students.

G2: Compare the identified themes to the five factors of the Sexual Risk Survey (Turchik & Garske, 2009).

Methods

Participants

Participants (n= 151) for this study were recruited from introductory psychology and linguistics courses at the University of Oregon. They were college-aged (m= 20, sd= 4), mostly female (62% female, 33% male, 3% non-binary, 2% other), White (69% White, 13% Latino/a, 11% Asian, 3% Black, 4% other), and heterosexual (71% heterosexual, 17% bisexual, 5% gay/lesbian, 7% other). Participants received partial course credit for taking part in this study.

Materials

Participants were asked to write two narratives using the following prompts:

Describe a time when you or someone you're close to engaged in sexual behavior that you consider risky.

Describe a time someone you know (but aren't necessarily close to) engaged in sexual behavior that you consider risky.

They were told to write at least one paragraph and given additional prompts such as: what happened? Who was involved? How did you feel when taking part in or hearing about this

situation? If they told a story about themselves in the first prompt, they were asked two additional questions: *Did you consider the behavior you engaged in risky before engaging in it? And, How much did you consider risk when engaging in the behavior?*

Participants filled out additional questionnaires including the Big Five Inventory-2 (BFI-2; Soto & John, 2017) to assess personality traits. They were also asked to indicate the extent to which they agreed that items from the SRS were risky on a Likert-style scale with 1= disagree strongly and 5= agree strongly that the behavior is risky. These items were reframed as statements, for example, "How many times have you had a sexual encounter you engaged in willingly but later regretted? (Turchik & Garske, 2009)" was restated as "Having a sexual encounter you engaged in willingly but later regretted". Participants also answered sexual history questions including what type of sex education they received (abstinence-only/abstinence focused at school and/or at home, comprehensive at school and/or at home, none, or other) and demographic information. These data will not be used for the current study but will be analyzed in future research.

Procedure

The data for this study was collected using online questionnaires that could be taken at participant's leisure. Informed consent was obtained before participants began answering questions.

Analytic Strategy

Themes of RSB in the narratives were identified using an inductive thematic analysis. The first step in this process was to read through all of the narratives and take notes on details that seemed important. Using these notes, I created an initial coding scheme which asked: who, what, where, why, and which emotions were present in the narrative. *Who* included any of the

people involved, *what* was the behavior or behaviors mentioned, *where* the behaviors took place as well as the context of the narrative, *why* did the storyteller consider the behaviors risky, and which emotions did the storyteller and/or subject express during or after the events of the narrative? I then trained two other researchers on the coding scheme and how to read through the narratives. Separately, we read through the narratives again and used the codes to find the themes we found to be most important to the participants. We then came together to share our lists and further discuss which themes were to be included as the final results of this study.

We looked for themes at both the semantic and latent level. At the semantic level, themes were identified based on explicitly stated language that did not need further interpretation. Identifying latent level themes involved deeper interpretation of the narratives to find underlying ideas and concepts not explicitly stated by the storyteller (Braun & Clarke, 2008). This process is more complex than finding semantic themes because it requires theorizing about the meanings of the narratives and understanding how the values of the storyteller influenced their process of choosing a narrative as well as what information they decided to include. Latent themes are subjective and harder to identify. However, when I and the other two researchers discussed our findings, we found that we identified many overlapping latent themes. Semantic level themes generally focused on the *who, what*, and *where* of the narratives, while latent theme analysis helped us find the *why* and *emotions* featured in the narratives, though there was intersection of themes.

Results

Using an inductive thematic analysis, we found seven themes of risky sexual behavior that we identified as important to the participants. Many of the narratives featured more than one theme and most of the themes encompass a variety of behaviors. The final themes of self-

described risky sexual behavior in college students are: Unprotected Sex, Sex in Public, Technology, Emotionally Risky Behavior, Stranger Danger, Under the Influence, and Non-Consensual Behaviors. Some behaviors that were mentioned but that were not included in the themes due to them being mentioned infrequently include: engaging in sexual behavior for money/other products, sex before marriage, and sexual behavior while in the same house as family.

Unprotected Sex

The theme with highest salience within the narratives was Unprotected Sex. Generally, this was referred to in the context of heterosexual, vaginal intercourse. However, there were other unprotected behaviors mentioned as well, such as this participant who mentioned having "unprotected anal sex with someone I didn't know too well". Non-heterosexual unprotected sex was discussed, including a story about how perceived norms around unprotected homosexual sex resulted in potential STI exposure:

> Upon getting out of a long term relationship, I went to Tinder for hookups (I am a lesbian woman). I was talking to a few girls, and I decided to meet up with one on a Friday night for the first time… A few weeks later, I got a text from [her] in which she told me that she found out she had chlamydia when we met. I was super upset with her because she admitted to having not even used a condom with the man she slept with before me. Lesbians know that dental dams are not a thing people actually use, so I recall being very disappointed in how big of a risk she took by not simply using a condom with the random man that she met before me.

Other unprotected sexual behavior was not specified. For example, one participant told a story about a friend who "hooked up with a random girl and did not use protection". "Hooked-up" likely implies sexual intercourse, but it may refer to other behavior as well.

Participants also had different ideas about what constituted "unprotected" sex. Many narratives focused on condom usage, some describing a single incidence of sex without a condom while others discuss ongoing failure to use a condom. As one participant says, "My friend never wears a condom." This was a fairly common occurrence, especially in stories about others, though many expressed their concern about their friends' and peers' choices. The participant above continued: "I'm worried about him but he's grown and I can't control him." Self-stories about condom usage tended to focus on a specific occurrence. These narratives often involved a long-term, trusted partner and participants expressed enjoying the experience, even if they perceived it as risky.

Along with condom usage, others consider vaginal sex without some form of birth control as risky behavior. One participant discussed her former roommate who "would meet up with guys from tinder that she had never met and have sex while not being on birth control leading to multiple pregnancy scares." For some, the risk of unintended pregnancy seemed more serious than the risk of contracting an STI. Others considered sex without condoms and birth control to be riskiest, such as the story told by a participant:

> When one of my best friends was dating her past boyfriend they often had sex without condoms. To me, that is considered risky because she was not on any form of birth control, and he did not get tested for anything before they started having sex. When she told me this had been happening I was shocked that she was not concerned for herself, not only for the possibility of STDs, but for possibility of pregnancy.

As this story illustrates, some were shocked at their friends' lack of concern regarding the consequences of unprotected sex, though a wide range of emotions accompanied the narratives featuring this theme.

Sex in Public

Participants told stories about themselves and others engaging in sexual behaviors in a variety of public places. Many of these narratives took place when the storyteller or the subject of the story were in high school. Cars were frequently mentioned as a risky location to take part in sexual behavior. One participant knew about two people from high school who "decided to engage in a sexual act in the backseat of a moving car". This and other narratives described instances where people were not only engaging in sexual behavior in public places, but also knowingly around others. The subjects of these stories did not seem worried about being seen or judged by those around them. Many narratives about public sexual behavior were told in a lighthearted way, despite the possibility of serious consequences. Take this story, for example:

> A time one of my friends engaged in sexual behavior that I considered risky was when they had sexual intercourse in a public pool. It was two people who did not know each other very well. This situation was funny when talked about but also could have led to worse decisions and consequences.

Others were more concerned about themselves or their friends/peers engaging in public sexual behavior and expressed feelings of anxiety and disapproval.

Risky sexual behavior is often thought of as health risking behavior, but the primary consequence of sex in public is getting caught. This involves two different types of risk: social/reputational and legal risk. One participant described a situation where they had sex with a partner on a public beach and stated, "doing this in a public place was not only illegal, but

stressful as anyone could have caught us." Not only is getting caught potentially embarrassing, but it could also result in a fine or even jail time. While those who engaged in public sex often expressed enjoying their experience in the moment, some felt more negatively about the situation as time passed. From the same participant as above: "During the experience I enjoyed myself, however as time passed I became less and less satisfied with my decision." A significant portion of the public sex narratives took place in high school or during freshman year of college when the subject of the story did not have access to a private location. This indicates that the choice to participate in this risky behavior was often out of necessity and not a behavior that the person would engage in if they had another option.

Technology

Technology is a theme that encompasses a wide range of behaviors. One of these behaviors is using dating or hook-up phone applications or websites, such as Tinder, Bumble, and YikYak, to meet people with the purpose of engaging in sexual behavior. Notably, participants only considered this behavior to be risky if they or the subject of the story engaged in sexual behavior with someone the first time they met them in person. One participant described a scenario about a friend of a friend:

> [She] met a guy on tinder and they barely talked before meeting up. We told her it didn't seem smart or safe to meet a complete stranger off the internet that she never knew prior but she didn't listen to us.

As with the above story, participants find meeting up with someone with the sole purpose of engaging in sexual behavior to be dangerous, especially if they know little about the other individual. However, there were no narratives that mentioned meeting someone online, going on dates, and then engaging in sexual behavior. Additionally, no one discussed meeting someone

online that they have mutual friends with as being risky. This may suggest that college students are mistrustful of meeting people online because they cannot be sure that the person is accurately representing themselves or their intentions. It seems that participants felt that this behavior is no longer risky if they meet the individual at least once beforehand or if they have mutual friends.

Other technology related risky behaviors focused on individual behavior such as sending nude pictures, or participating in "camming" in which people livestream or video chat with others for money, usually in a sexual context. Participants discussed the dangers associated with this behavior, and how easy it is for pictures and videos to be "leaked" or spread around online. One participant described a consequence associated with this risk:

My friend used to exchange nudes quite frequently with people on the internet without

knowing much about them first, and quite a few times they have ended up

communicating with dangerous people and had their nudes leaked or was threatened.

Similar stories described the subject getting blackmailed to send money or the other person would share their explicit pictures/videos with others online. This could have life changing consequences as it can be very difficult to get content taken off the internet. Another consequence of concern to participants was parents and family seeing explicit content. As one participant stated: "Someone I knew was camming online. [This is] risky if their parents found out." Others mentioned taking videos of sexual acts as risky, similarly due to the risk of it being seen by unintended parties, such as this participant: "Someone that I know recorded a sex tape and this seems risky to me because there is always a chance that those things could get leaked." Narratives featuring technology in this way emphasize how easily it is for sexual content to get into the wrong hands. The loss of control and autonomy associated with digital sexual behaviors

is of great concern to many college students. This is another example of a type of social/reputational risk.

Emotionally Risky Behaviors

Some participants described situations where the main consequence of the exhibited behaviors was emotional distress. These narratives often involved someone experiencing heartbreak. Take this story for example:

My friend started having sexual relations with this one girl that we all knew. He was very naïve about relationships. We told him that he should not pursue a relationship with this specific girl because we thought that we was going to get his heart broken. He went against our warnings and decided to pursue the relationship anyway. He then got very hurt and changed a bit after everything was said and done.

Though the subject of the story did not engage in typical risky sexual behaviors, the participant still considered their friend's actions as risky because he suffered from negative consequences directly relating to the behavior. Another participant described a time when a friend wanted to be in a relationship with someone who did not reciprocate these feelings, but who did want a sexual relationship. The friend ended up beginning a purely sexual relationship with this person, despite wanting a serious relationship. The participant stated that the situation was "extremely emotionally risky from the beginning and did not turn out well." Again, no physical risk was discussed, but the emotional damage experienced was just as harmful.

The other type of emotional risk that participants expressed concern about were situations regarding infidelity or perceived betrayal. One participant described their own actions: "I almost cheated on my partner when I was under the influence, which I considered extremely emotionally risky." Not only could this hurt their partner, but the guilt associated with this

behavior evidently had a lasting effect on the participant. Some participants noted the health risks associated with infidelity, but it seemed that the emotional distress was more damaging in the long run. In one narrative, a participant described how a friend had unprotected sex with multiple individuals while also having unprotected sex with their partner. While this situation involved physical risk, the participant said that, "Everything turned out okay physically but the situation was met with some emotional damage as the relationship ended because of this."

Other situations involved perceived betrayal. For instance, one narrative discussed how a friend began having a sexual relationship with someone new after a breakup:

> My close friend was engaged in risky sexual behavior, although the risk had more to do with emotions than what actually went on. My friend and her boyfriend had broken up, and [she had sex with] another guy she had been talking to…Though they were never dating, her ex-boyfriend got upset with her. They eventually got back together, but it took a long time for him to not see her actions as cheating. I felt bad for my friend hearing about this, because she didn't deserve to be seen as in the wrong for what she did.

The friend's ex-boyfriend felt hurt, but his reaction resulted in emotional distress for the participant's friend as well. This and similar narratives illustrate how, for many, emotions are just as pertinent as the physical aspects of sexual behavior. Even in narratives where other risks were featured more prominently, emotions often played a significant role. Emotional risky behavior can result in consequences that are often longer-lasting and more harmful than consequences of other types of RSB.

Stranger Danger

There were a number of contexts and environments that the narratives took place in, but regardless of the situation, many stories featured sexual encounters with strangers. As mentioned

previously, participants expressed concern regarding meeting strangers online or on dating apps. Another common location for meeting and engaging in sexual behaviors with strangers was at parties, bars, or nightclubs. One participant describes a story about this as well as their opinion about the topic:

> It was just a normal weekend and me and my girl friends went out like normal and were drinking and my roommate met a guy at the party and it was the first time they had ever interacted. They hung out at the party and my roommate ended up going home with the guy that night and spending the night with him at his apartment. This is something that I don't believe that I would ever do because I don't know how safe I would feel going home with a complete stranger and it leading to any type of sexual encounter.

While this narrative does not go in depth about the sexual encounter, it is evident that the participant was less concerned about the details of the encounter and considered this to be risky simply because it took place with a stranger. Others discussed why they considered sexual behavior with strangers to be risky, often mentioning that it is harder to know their sexual history. Some told stories about themselves or others contracting STIs after sexual encounters with someone they did not know. One participant described a time they had sex with "someone random": "This girl really wanted to have sex without a condom. I probably should have said no but I didn't. I actually ended up getting Chlamydia". Though it may be just as likely to contract an STI from someone known, participants seem to perceive unprotected sex with strangers as more dangerous in this regard.

Another reason sex with strangers was considered risky was due to the age of the stranger. These stories featured people who were of consenting age, but who were engaging in sexual behaviors with strangers much older than them. This was considered risky by participants

who questioned the intentions of the older person. In one narrative, the participant described how her friend who had recently turned 18 began talking to a much older man. The participant stated that this made them feel "extremely uneasy" and that they believe that "men who are [older] that seduce young women are dangerous to me". Not only did participants question strangers' sexual histories and safety concerns, they did not trust the intentions of strangers, especially if they were not a part of their greater social circle.

Under the Influence

Previous research has focused heavily on sex while under the influence of drugs and/or alcohol and this was also a prevalent theme for our participants. Even if drugs or alcohol was not the primary focus of a narrative, it often played a role. Most narratives relating to this theme centered around alcohol use, though some also mentioned drug usage or a combination of drugs and alcohol. While substance use was only briefly mentioned by some, others focused on this as the main risky behavior in their narratives. A variety of behaviors were mentioned as being risky to participate in while intoxicated, some that are rarely considered to be risky when those involved are sober. For instance: "My friends often times will get drunk and make out with other drunk frat men at parties. I consider this risky because both of [the] parties are not clear headed in being able to make decisions." In this story and others, the exhibited behaviors are considered risky simply because alcohol was involved. Other stories discussed people engaging in behavior they would not normally engage in due to substance use. In one narrative, a participant discussed how someone they knew had sex in a public restroom with their partner and someone walked in on them. The participant stated that, "they were drunk and acting wild when they maybe wouldn't have had they been sober." Evidently, students believe that alcohol plays a significant

role in risky sexual behavior, either by fueling it or by turning otherwise non-risky situations into risky ones.

Another issue of engaging in sexual behavior while under the influence of drugs and/or alcohol is that one's ability to consent is complicated or eliminated. This is true even for those with committed partners. Take this participant's story for example:

> Over a year ago, a close friend of mine became drunk at a summer party. After he went to bed, his girlfriend at the time joined him. Although he offered verbal consent to sex, he did not remember, and likely was not capable of offering legal consent as he was heavily intoxicated. The next morning, when he woke up, his girlfriend began to discuss the previous night's sex to him before realizing that he had no recollection of the night before. I felt bad for my friend and became concerned about his mental health after the incident…The girlfriend felt horrible and apologized to my friend, and he seemed to brush it off, explaining that although drunk, he offered consent and that if he were sober, he likely would have consented. After this event, from my understanding, they have been explicitly careful in asking for consent.

While the subject of this story likely would have consented had he been sober, alcohol took away this ability and resulted in the possibility of negative emotional or psychological consequences. In this narrative, the subject was with someone they trusted, though other stories involved people that were not quite as trustworthy. In these narratives, the subject of the story was unable to consent and would likely not have consented had they been sober. These stories may be considered "risky behavior" due to alcohol use, but not "risky sexual behavior" because the person was unable to consent to the sexual acts. In this study, these situations were placed under the theme of Non-Consensual Behaviors, which will be discussed in more detail. But as alcohol

usage was featured heavily, these narratives fit into both themes. However, this does not mean that the subjects chose to engage in the sexual behavior and many of the narratives were stories of sexual assault.

In some narratives, the subject of the story was not the victim, but the perpetrator, of sexual misconduct while they were under the influence of substances. One participant tells a story about someone they know:

> [He] was drunk and didn't get the social ques of the person that he went to party with. It resulted in her asking him to walk her home, but since he was so drunk, he walked her to his house instead and proceeded to have sex with her while she was blacked out.

In this and other stories where the subject of the story was the perpetrator, the participant did not seem to consider the behavior as sexual assault and they generally discussed it in terms of reputational/social, not legal risk. It is evident that many students know someone who has been sexually assaulted or who has sexually assaulted someone else while under the influence. Because drug and alcohol use is a part of college life for many, this type of behavior is perhaps the hardest to target from a risk reduction standpoint. Students are taught about consent and how to make safe choices when using substances, though when they and others are under the influence, any prior education becomes irrelevant.

Non-Consensual Behaviors

In some narratives, the main risk involved was the intentions and actions of other people. These stories focused on situations where the subject of the story was unable to consent due to their levels of intoxication (as mentioned above) or their age. In these narratives, the participant considered the behavior to be risky because others put the subject of the story at risk. Often, they blamed themselves or their friends/peers even though they were the victim of sexual assault or

statutory rape. Some of these stories included the subject getting drugged at a party or bar. One participant described how they were drugged and nearly sexually assaulted before they were rescued by a friend. Despite this situation being out of the participant's control, they stated, "When I found out what happened the next day, I felt ashamed of myself." This sentiment was echoed by others who had similar experiences, often describing feelings of guilt and regret.

Other narratives regarding the intentions and actions of others focused on relationships or sexual encounters where the subject was underaged and the other person was not. These stories were also not consensual due to the age of one of the people involved. Consider the example from this participant:

> When I was in middle school, I remember a classmate telling me about her boyfriend. He was 19 and she was 14 at the time. She told me they were having sex and only using a condom. I was young enough to not know about statutory rape or the laws that determine legal age gaps and consent, but I still had a clear gut feeling that she was being taken advantage of and it was an unsafe relationship and sexual relationship. Later that year she got pregnant and then stopped coming to school.

While this study is focusing on college students, it is important to understand how they perceive risk generally, and stories such as this are likely to affect the victim throughout their lives. It is often the case in situations of statutory rape that the minor involved does not consider the behavior of the other person to be dangerous, but the consequences of their actions may result in psychological, emotional, and even physical harm.

Including Non-Consensual Behaviors as a theme was not an easy decision. Research on risky sexual behavior generally focuses on changeable behavior and places blame on the individual engaging in risk. However, this theme was of great concern to the participants and

these narratives brought to light a different type of risk that is not often discussed in relation to risky sexual behavior: contextual risk. This is risk due to factors that are out of control of the person involved based on outside factors. It is not likely that this theme would be included in a measure of RSB because measures tend to focus on individual behaviors and not the actions of others. Because this study is centered on the beliefs of the participants, we chose to include Non-Consensual Behaviors because the students frequently mentioned this as a risky behavior. The frequency with which this behavior was mentioned as a risk taken by the subject of the story even though they could not consent indicates that college students need more education on what is and is not consensual sex.

Discussion

G1: Identify themes in self-described sexually risky behavior in college students.

The themes we identified as self-described risky sexual behavior in college students were Unprotected Sex, Sex in Public, Technology, Emotionally Risky Behavior, Stranger Danger, Under the Influence, and Non-Consensual Behaviors. Some of these themes were expected, including Unprotected Sex, Stranger Danger, and Under the Influence, though other themes were surprising and have interesting implications. As with previous research examining RSB, we initially looked at this behavior through a health risking lens only. However, the participants were concerned about a wide range of risk. Social/reputational, emotional, legal, and contextual risk were just as salient to the participants as health risks. Some narratives focused on one type of risk or one behavior while others expressed concern about various types of risk and behaviors.

In addition to various types of risk, the way that participants perceive risk exists on a spectrum. For some, risk did not stop them from enjoying a sexual experience. If their narrative featured a friend or peer, they often described the risk involved as "not a big deal" or that the

subject had a good time. Some even describing certain situations as "funny", often in the context of narratives regarding Sex in Public. Even if the risky behavior may result in health risks, such as STI contraction or unplanned pregnancy, some were not very concerned about these consequences. At the opposite end of the spectrum, some participants viewed risk as extremely harmful. This was evident in the behaviors they identified as risky including sexual assault and statutory rape. These participants expressed feelings of anxiety and distress whether they were discussing themselves or others. Some participants that wrote about others conveyed feelings of judgment and disapproval, even in situations where the subject of the story was a victim of sexual assault. The majority of the participants fell somewhere in the middle of the spectrum. They took a balanced approached to risk, avoiding it if possible but not condemning themselves or others who engaged in occasional RSB. This balanced approach may be used as a model for researchers studying risky sexual behavior.

G2: Compare the identified themes to the five factors of the Sexual Risk Survey

The second goal of this study was to compare the five factors of the Sexual Risk Survey to the themes that were identified in the narratives. We expected some crossover and also some divergence. After identifying the themes, we found only minimal overlap between the SRS factors and our results (see Figure 1). The SRS factor Risky Sex Acts includes items relating to vaginal and oral sex without protection as well as sex while under the influence of substances. This was similar to our themes of Unprotected Sex and Under the Influence. However, the SRS includes a separate category for Risky Anal Sex Acts while we included this within the theme of Unprotected Sex. The SRS factor of Sexual Risk Taking with Uncommitted Partners was similar to our theme of Stanger Danger, but there was also some discrepancy. Similar to our findings, this factor encompasses sex with strangers, sex with untested partners, and sex before discussing

risk factors. However, Uncommitted Partners in the SRS included items which asked about sex with any uncommitted partner, number of sexual partners, and sex with partners who have other partners. Our participants did not seem to consider sex with uncommitted partners (such as friends or friends with benefits) to be risky. They also did not mentioned number of partners or sex with partners who have other partners as risky behavior.

The other factors of the SRS that were not discussed by the participants in this study were Impulsive Sexual Behavior, Intent to Engage in Risky Sexual Behaviors, and as mentioned, Risky Anal Sex Acts. Impulsive Sexual Behavior involved questions about number of sexual behavior partners (including kissing and fondling), sexual behavior with acquaintances, and regretted or unexpected sexual encounters. While our participants did mention unexpected encounters, this behavior was always accompanied by other risks which the participants seemed more concerned about. The participants may have felt regret associated with RSB, but their experiences were much more complex. They discussed deeper emotions such as shame, loss, and grief. Intent to Engage in Risky Sexual Behaviors was a factor included in the SRS because researchers wanted to include students who had not yet engaged in sexual behavior (Turchik & Garske, 2009). These items involved situations where someone went to a party or other social event with the intent to engage in sexual behavior but who did not engage is any behavior. Intent without engaging in sexual behavior is unlikely to result in negative consequences and this factor was not reflected in our results. Finally, Risky Anal Sex Acts does not need to be a category separate from Unprotected Sex. It seems that this was included as a factor due to antiquated beliefs that associate specific consequences (HIV/AIDS) solely with specific behaviors (anal sex between homosexual men). Our results indicate that students with a variety of identities engage in anal sex. Additionally, HIV/AIDS and STIs contracted during vaginal, oral, or anal sex affect

all types of people. Because of this, we feel that unprotected vaginal, oral, and anal sex should be included under one theme.

The current study found themes that were not represented in the SRS. The themes Sex in Public, Technology, Emotionally Risky Behavior, and Non-Consensual Behaviors did not relate to any items in the measure. It is likely that researchers did not ask about technology related sexual risk taking in the SRS because the measure was creating between 2006 and 2009, a time when smartphones were much less common and Tinder was yet to exist. Additionally, the SRS focused mostly on health risking behaviors while our themes of Sex in Public, Emotionally Risky Behavior, and Technology involved other types of consequences. The SRS involved individual behaviors and did not examine how others' actions can result in consequences. Non-Consensual Behaviors is an important theme that we found within the narratives and this theme encompassed stories where someone else perpetrated the risky behavior. We feel that this theme should be included though it does not focus on the subject's own behavior.
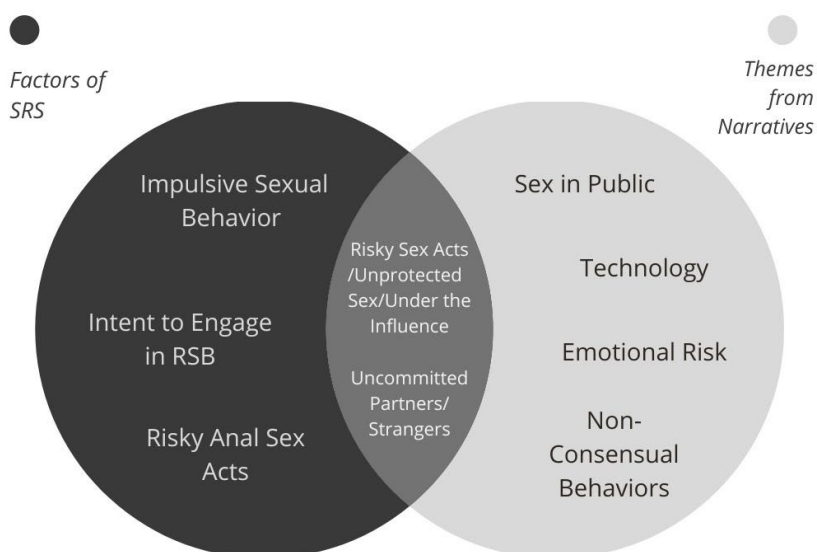

Figure 1.

Implications

The results of this study indicate that college students have clear ideas about what they consider to be risky sexual behavior. The study design allowed them to express their thoughts and feelings in their own words. They shared their thoughtful, complex, and at times, heartbreaking experiences with RSB. Sometimes, they participated in the behavior themselves while others witnessed or heard about the experiences of others. Regardless, it is evident that college students are more than capable of contributing their knowledge and understanding of RSB to the existing research on the topic. It is also important to acknowledge that students conceptualize and perceive risk on a spectrum, and that there is not one "correct" answer to the question, "What is risky sexual behavior?".

Due to the variability in college students' experiences and beliefs regarding RSB, it is essential that researchers look at this topic through a wide lens. This includes limiting their own biases and moralization of risk. Previous research on RSB in all populations takes an approach that focuses heavily on negative outcomes and included behaviors are decided by researchers with little to no input from the population of interest. These studies rarely examine consequences other than health risk outcomes. Researchers do not seem to consider other types of outcomes, such as emotional or reputational consequences, as relevant to their research. Additionally, there is a belief that risk should be approached in a "one-size-fits-all" way. That is, that there is a "correct" way to conceptualize risk and that everyone will have the same negative experience if they choose to engage in risky behavior. In reality, different types of risk effect some people more than others and in unique ways. Overall, the ways in which researchers theorize about risk should place the values of the population, rather than their own values, at the forefront. A good

beginning point for this is allowing the population to speak for themselves and to communicate their feelings in their own words.

Limitations

Similar to the limited population that the SRS sampled to develop their measure, our sample also lacked diversity. The majority of our participants were White (69%), female (62%), and heterosexual (71%). Additionally, most of our participants were middle class and likely come from families with liberal or moderate political beliefs. A significant portion of University of Oregon students come from Oregon, California, and Washington (80% of the student population are from these three states alone [University of Oregon]), indicating that our sample may reflect the views of West Coast college students, but the findings may not be generalizable to college students in the greater United States or other countries. Future studies should focus on a wider population of students from diverse backgrounds.

Future Directions

The themes identified in this study did not align with the behaviors included in the Sexual Risk Survey, which is currently the most widely used measure of RSB specific to college students. This suggests that a new measure is needed to accurate assess how current students are thinking about and engaging in RSB. Due to the rapid changes in social norms, especially on college campuses, a new measure should be flexible and updated regularly. The development of this measure should heavily take into account the opinions of students, though it is also important to engage the opinions of experts, such as health professionals and sex educators. The goal for this measure should not be to judge or shame students, but to accurately understand which risky sexual behaviors they are engaging in to better employ methods of risk reduction or otherwise gain valuable knowledge to be used for other purposes.

In terms of future applications for this research, risk education and reduction strategies could center around the identified themes. For example, a sex education course or workshop for college students in which our themes were used as a basis for the curriculum would likely be more impactful than a program based solely on issues that educators or researchers deem as important. The more that information is relevant to the students, the more they will engage with the content. This would hopefully lead to a decrease in consequences of risky behavior by giving students knowledge that could help them make safer choices, even if some of these choices involved risk. For example, talking openly with students about testing and treatment of STIs without stigmatizing them could encourage regular screenings which would likely decrease STI transmission, even in those who choose to have sex without a condom. Another possibility for risk reduction would be to encourage compassion and respect for others. Many of the narratives featured situations where the person or people involved did not seem to care about how their actions may affect others. Stimulating feelings of empathy and consideration for others by building community and fostering honest conversations may help to decrease certain risky behaviors.

Conclusions

Risky sexual behavior (RSB) is a complex topic that is challenging to define yet critical to understand due to the wide range of negative consequences associated with it. This study used a thematic analysis to identify how students define RSB by asking them to explain it in their own words. We found specific patterns in descriptive narratives told by the students and our analysis resulted in the formation of seven themes: Unprotected Sex, Sex in Public, Technology, Emotionally Risky Behavior, Stranger Danger, Under the Influence, and Non-Consensual Behaviors. These themes partially aligned with previous research on RSB in college populations.

However, our results also differed significantly and we identified behaviors that may be uniquely relevant to modern college students. Future studies should aim to collect similar qualitative responses on other college campus to get a better understanding of how risk is perceived in diverse populations. The next step is to develop a new measure that reflects the values of current students as a way to accurately assess their participation in RSB. This information can be used to develop risk education and reduction strategies, with a long term goal of decreasing the consequences of risky sexual behavior nationally, or perhaps globally.

References

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Facts at a Glance Fall Term 2021, Fourth Week- University of Oregon. (n.d.). [Statistics & Historical Data]. University of Oregon. https://registrar.uoregon.edu/sites/registrar1.uoregon.edu/files/pdf/uo-facts-at-a-glance-202101-fall-2021.pdf

Fino, E., Jaspal, R., Lopes, B., Wignall, L., & Bloxsom, C. (2021). The Sexual Risk Behaviors Scale (SRBS): Development & validation in a university student sample in the UK. *Evaluation & the Health Professions*, *44*(2), 152–160. https://doi.org/10.1177/01632787211003950

Jaspal, R. (2019). Why a Social Psychology of Gay Men? In R. Jaspal, *The Social Psychology of Gay Men* (pp. 3–15). Springer International Publishing. https://doi.org/10.1007/978-3-030-27057-5_1

Sexual Consent. (n.d.). *Planned Parenthood*. https://www.plannedparenthood.org/learn/relationships/sexual-consent#:~:text=%7C%20Planned%20Parenthood%20Video,is%20rape%20or%20sexual%20assault.

Soto, C. J., & John, O. P. (2017). The next Big Five Inventory (BFI-2): Developing and assessing a hierarchical model with 15 facets to enhance bandwidth, fidelity, and predictive power. *Journal of Personality and Social Psychology*, *113*(1), 117–143. https://doi.org/10.1037/pspp0000096

Turchik, J. A., & Garske, J. P. (2009). Measurement of Sexual Risk Taking Among College Students. *Archives of Sexual Behavior*, *38*(6), 936–948. https://doi.org/10.1007/s10508-008-9388-z

Who are These Virtue Signalers Anyway?:

Answers from a Sustainable Transportation Survey

Bridgette Bammann

University of Oregon

Abstract

Virtue signaling – advertising one's morally respectable actions – is a popular cultural concept, but there has been little research exploring who in the general population participates in virtue signaling. The present study aims to describe virtue signaling as observed in a naturalistic setting. Data were obtained from the open-ended responses of a larger survey about Eugene/Springfield Oregon residents' transportation habits before and during the COVID-19 lockdown of spring 2020 (n = 673). In coding survey responses, we found unprompted comments highlighting their sustainable transportation habits and attitudes – i.e., virtue signaling. We expected younger participants and participants with higher incomes to be the more likely virtue signalers but did not find support for this hypothesis. We expected those who scored higher on a measure of beliefs and attitudes regarding environmental policy to be more likely to virtue signal, and this hypothesis was supported. Labelling something as "virtue signaling" is often uncomplimentary, but we explore whether virtue signalers may have other qualities that make them effective targets for future environmental protection interventions.

Who are these Virtue Signalers Anyway?: Answers from a Sustainable Transportation Survey

We may very well be headed toward a total climate change catastrophe. While most accountability can be placed on commercial industry and agriculture for their record-setting contribution of greenhouse gases to the atmosphere, in the United States our individual transportation accounts for a total of 37% of carbon emissions (Energy Information Administration, 2019). In Oregon, rates of single-driver car use continue to rise and account for 40% of the greenhouse gases emitted in the state (Oregon Global Warming Commission, 2018). And in Eugene, roughly 65% of the population drives alone to work (U.S. Census Bureau, 2018). Concern about the carbon-contributions of individual drivers led the city of Eugene to partner with researchers at the University of Oregon to learn ways to reduce these numbers. Together, they developed a survey to learn more about transportation habits while also searching for insight into the attitudes and beliefs that deter people from using alternative methods of transportation such as walking, biking, or riding the bus.

This transportation research group was initially interested in conducting a survey about sustainable transportation in Eugene but when the COVID-19 pandemic hit, the survey transitioned to questions about people's transportation habits before and during the COVID-19 "lockdown" of Spring 2020. Throughout the survey, there were four separate instances when people were prompted to answer questions like, "Is there anything else you'd like to add or clarify about your answers to the questions about your habits during COVID-19 shelter-in-place?" Upon reviewing these open-ended responses, we were struck by how frequently people volunteered information pertaining to how they were "good environmentalists." This style of communication could be viewed as 'virtue signaling.' While the origin of this phrase is disputed, James Bartholomew who claims to have coined it says virtue signaling means "indicating that

you are kind, decent, and virtuous (Bartholomew, 2015, para. 2)." When applied with motive, the phrase 'virtue signaling' describes an empty message in which a person attempts to display their socially pleasing, "politically-correct," or environmentally friendly attitude in order to boost their status (Griskevicius, Tybur, & Van den Bergh, 2010) or social desirability, including romantic value (Palomo-Vélez, Tybur, & van Vugt, 2021). Many use 'virtue signaling' as a pejorative term, but for the purpose of this research I will be using virtue signaling as a neutral label for when someone voluntarily expresses information conveying their values.

In the present study, we are interested in who in the general population takes part in virtue signaling, and if virtue signaling is actually reflective of the beliefs and habits that signalers engage in. Previous research has attempted to discover virtue signaling primarily in a lab setting (e.g., Griskevicius, et al., 2010; Wallace & De Chernatony, 2020). The goal of this investigation is to describe virtue signaling as it occurred free from experimenter demand effects. Many responses received from the survey made mention of various virtuous behaviors, including messages about wearing a mask to prevent COVID-19 or how Black lives matter but for the purpose of this study we are specifically interested in virtue signaling related to sustainable transportation use. Because the survey was about transportation habits, and participants had already answered multiple questions providing this information, we felt most confident identifying additional responses about sustainable transportation practices as virtue signaling.

Once we finished reading and coding all of the responses, we examined how various demographics collected along with the survey were related to virtue signaling. Who did or did not respond to the open-ended questions? Who did or did not choose to explain more about their transportation habits or provide some explanation for why others chose their transportation

habits? Who *are* the virtue signalers anyway? Particular areas of interest included age, gender, and income.

We predict that younger participants as well as higher earners will be more likely to respond to the open-ended questions and will be more likely to virtue signal. Previous research supporting our age hypothesis suggests young people are more likely to be concerned about the destruction of the environment than older people (Liere & Dunlap, 1980). One possible explanation for this age difference is that younger people are less integrated in social order and because environmental issues are a threat to social order, young people are more likely to support actions against environmental destruction (Liere & Dunlap, 1980). Relevant to our income hypothesis is the social-class hypothesis that states environmental concern is positively associated with education and income (Liere & Dunlap, 1980). This hypothesis is based on Maslow's theory of hierarchy of needs, suggesting that because middle and upper classes have satisfied their basic material needs, they are able to focus on 'higher' needs, like environmental welfare. Individuals of higher income may be more likely to afford the choice of primarily using alternative methods of transportation.

We are also interested in seeing how the people who virtue signal responded to other areas of the survey that asked questions about attitudes and behaviors. Do those who virtue signal also report using more sustainable transportation modes such as walking, biking, or riding the bus as their primary transportation method? Are those who virtue signal voting in the local elections? Essentially, are the signalers backing up their words with corresponding action?

By taking a closer look at the habits that signalers engage in, we might be able to better understand the role of virtue signaling in a social context. Studies examining people's spending habits show how people are more likely to buy an "environmentally friendly" product if they are

shopping in a public store versus online (Griskevicius, et al., 2010). These results suggest the importance of appearing prosocial when it comes to how we choose to spend our money. By signaling a consideration for the environment, and thus society, the "green" shoppers may be more likely to reap rewards of status gains.

Moral grandstanding or using moral rhetoric to promote one's own interests (Tosi & Warmke, 2016) is also often ridiculed across multiple social spheres. Humble bragging or masking a statement in the form of a complaint in order to appear humble (Sezer, Gino, & Norton, 2018) has been cited as conveying a similar sentiment and thus serving a similar purpose; however, the phrases moral grandstanding and humble bragging imply that the accused is acting out of vanity. Virtue signaling differs in that it is more ambiguous. To be called virtuous is considered a compliment on one's high moral standards. And signaling is simply transmitting information. It is *how* someone uses the phrase that ultimately dictates its meaning. Furthermore, a key part of my investigation will be highlighting various ways in which people attempt to display these values.

It's easy to spot how businesses and companies signal their moral values. They consistently change marketing strategies with shifts in social hot topics. All of a sudden, following national attention on a topic, they're best friends with the polar bears, or active agents in battling racial inequality. Their messages show up plastered on every window of every store, or in the unavoidable advertisements we scroll past on a regular basis. The main message from these companies: We care. Studies have shown how corporations have created an entirely new sector of the industry called 'business sustainability' motivated solely on exhibiting their consideration for social and environmental well-being (Gray, Sütterlin, Siegrist, & Árvai, 2020). In the context of the business world, companies are willing to sacrifice a marginal level of profit

in order to appease consumer expectations, suggesting the powerful effect that virtue signaling has on a broad base of people.

Virtue signaling reflects social norms and values (Zaki & Cikara, 2020). When information was released about the harmful effects that straws were having on the sea turtle population in 2019, people around the world started asking for their drinks to be served without them. Quickly to follow, policy changes were implemented. Cities such as Eugene, Oregon banned the use of all single-use plastic utensils unless requested by the customer (City of Eugene, 2019).

This summarizes why signaling is important. While signalers may have no intention of dedicating their time to support their cause, the accumulation of many voices toward one objective can create social change (Zaki & Cikara, 2020). It is possible that virtue signalers may be more susceptible to future interventions of societal do-gooding. One study explored cognitive dissonance among a group of academic professionals and their relationship with traveling by plane (Schrems & Upham, 2020). The dissonance in this case was related to inconsistency between pro-environmental attitudes and flying for academic purposes. This study found that in order to resolve feelings of dissonance, the participants reported changes in behavior such as reducing that number of flights taken and engaging with a planned compensation program to offset their carbon impact (Schrems & Upham, 2020). This same style of intervention may be translated to the virtue signalers. Say a community decides to hold a bicycle-to-work day in order to lower the use of single-driver cars: virtue signalers may be good targets as their values are already known and they may be more likely to get involved.

Previous studies have been designed so that participants have the opportunity to choose between two or more limited behaviors in a lab setting (Jung, Nelson, Gneezy, & Gneezy, 2017;

Zentz, 2021), with one being designated as virtue signaling. In our study we will attempt to see if we can reliably identify people's behavior (in the form of people's comments) that could be considered virtue signaling. In the form of "direct" virtue signaling, people declare their allegiance toward sustainable transportation habits. There are however other methods of virtue signaling that occur through explanations for why either the self or other are unable to practice sustainable transportation habits. These explanations allow the signaler to share why they cannot be more virtuous but still display their value of pro-environmental behavior. Through describing the ways others are not able to be more virtuous, the signaler is able to share their values, and might even be able to make themselves appear more virtuous.

This definition of virtue signaling may be a bit broader than is used colloquially, but we believe through the extra effort that people are making to share with us what they find important, they are virtue signaling. After answering several questions about their transportation habits, these people provide even more information in an attempt to share what they care about. If virtues signaling is defined by the observer, and depends heavily on how the signaling is delivered, then this study represents a similar construct.

Method

Participants

Six hundred and seventy-three participants completed the transportation survey online from June 3, 2020 through July 31, 2020. Participants were recruited in the Eugene/Springfield region of Oregon via word-of-mouth, e-mail lists, Facebook ads, and reddit pages. Attempts were made to recruit a sample that was representative of the Eugene/Springfield area in terms of major demographic variables (age, income, and ethnicity). An opportunity to win a $50 gift card to a regional or national business was offered to those who took part in the survey.

Participants ranged from ages 18 to 82 years old ($M = 47.2$, $SD = 15.9$). The sample was predominately White with 78.7% identifying as White ($n = 406$). The next largest group was Hispanic/Latino at 5.0% ($n = 26$). All other racial and ethnic groups constituted 11.43% of the sample ($n = 59$) and 4.8% chose not to say ($n = 25$). In the sample, 18.0% identified as male ($n = 121$), 53.2% identified as female ($n = 358$), 4.9% identified as gender non-conforming ($n = 33$), and 23.9% chose not to say ($n = 161$ – it's not entirely clear why so many chose not to report their gender; most of these participants did complete the rest of the survey). Participants were given categories of ranges to describe their household income; the most commonly chosen category was $75k or higher annually. The survey was also made available in Spanish.

Survey

The data used in this project were drawn from a larger survey (see Appendix 2 for full set of survey questions). The survey was administered online via Qualtrics (an online survey software tool) and was divided into four sections. The first section contained 32 questions about participants' habits prior to the Covid-19 lockdown of Spring 2020. The next section, consisting of 25 questions asked participants about their habits during the "shelter-in-place" order issued by the state of Oregon in Spring 2020. Also included were seven Likert-style questions asking participants about their attitudes and beliefs towards transportation and environmental policies. The measurement of these scales' ranges from 1= strongly disagree to 5= strongly agree. Example questions from the survey include, "I believe government should adopt policies to reduce driving," and "I believe that people driving less is good for the environment." Responses to these attitude and belief questions were highly intercorrelated (Cronbach's $\alpha = .78$) so responses on the seven items were aggregated into one mean score. Lastly, the survey asked 20 categorical questions in order to collect demographic information. While participants were given

the choice among nine categories to describe their annual household income, these categories were collapsed to four levels: less than $25,000, $25,000 to $50,000, $50,000 to $75,000, and $75,000 or higher. Similarly, five categories were initially presented for educational level, but these were ultimately collapsed into four: no college, some college, bachelor's degree, and graduate degree or higher. The question about gender originally had four options that were collapsed to three: male, female, and other, and the eight options for race and ethnicity were collapsed into three: White, Latino/Hispanic and Non-White/Non-Hispanic. Each demographic question excluding education, also included the option of "prefer not to say."

Throughout the survey, there were four separate instances when participants were given the opportunity to provide open-ended responses. The first of these open-ended questions stated, "Is there anything else you'd like to add or clarify about your answers to the questions in Part 1 about your habits pre-COVID-19?" This question occurred after the portion of the survey asking about pre-lockdown travel behavior. Next there was a brief question prompting participants on whether they wanted to be entered into the gift card drawing. Immediately following was the next open-ended question which stated, "Are there any other details, comments, or clarifications that you'd like to add here that are related to this survey, the questions, and/or your answers?." The next open-ended question, "Is there anything else you'd like to add or clarify about your answers to the questions in Part 2 about your habits during COVID-19 shelter-in-place?" followed the portion of the survey asking participants about their transportation during the COVID-19 "shelter-in-place" of Spring 2020. The last open-ended question, "Is there anything else you'd like to add or clarify about your answers on beliefs and attitudes in Part 3?" immediately followed the portion of the survey that asked participants about their attitudes and

beliefs towards transportation and environmental policies. It was in these responses that we looked for virtue signaling.

Virtue Signaling Coding

We designed a coding scheme for participants' open-ended responses. The guidelines established can be found in the attached codebook (Appendix 1). First, we coded whether a participant provided any response to the open-ended questions. (Responses such as "no" or "nope" in response to whether participants had anything to add were considered as not providing a response.) While the majority of participants did not respond, for those who did, we then divided responses into those that were relevant to transportation and those that were not. The former category included responses such as "I used to bike more but my bike needs repair." The latter category included responses that had to do with clarifying previous answers or responses related to COVID-19 that had nothing to do with transportation such as, "There's an error in the previous question comparing frequency of walking/biking now vs. last year." All quoted examples are from actual respondents.

Many of the open-ended responses were related to transportation but did not include a form of virtue signaling. Some referenced not being able to afford a car, others talked about transportation habits changing due to precautions taken to protect against COVID-19. Some expressed opinions that were not in favor of sustainable transportation habits such as, "you can take my gas guzzling, 16 mpg on high octane only, sports car from my cold dead hands…" Designating responses about transportation was an important first step in developing our coding scheme.

For the responses that were about transportation, we then coded whether the responses contained three separate categories of virtue signaling: "direct" virtue signaling; personal

explanation/excuse; and public/policy explanation as virtue signaling. All three kinds of

statements were considered virtue signaling and are described in detail below.

   Direct virtue signaling. "Direct" virtue signaling included participants' statements of

describing their own sustainable transportation "virtuous behavior". The responders that directly

virtue signaled gave a clear indication that they value environmental sustainability by saying

things like, "We have a solar charging station for our electric car." Other examples in this

category were responses discussing sustainable transportation as exercise, an increase in

sustainable transportation use due to weather, and many responses about continued use of

sustainable transportation despite not travelling to work anymore (presumably due to the

COVID-19 lockdown).

   Personal explanation or excuse. Some responders gave a personal explanation or excuse

for why they did or did not use alternative transportation methods. For example, one response

received stated, "Because of where I live, I have no other choice than to use my car for

transportation for work." Also considered relevant to transportation were responses that included

information about their amount of travel and related travel energy increasing or decreasing (e.g.

"I am using way less gas"), discussion of proximity to stores or errands such as "8 miles to the

nearest grocery store," and any commentary on a trip to or from the home. Not included were

comments about the location of participant's homes such as, "I live in Creswell."

   Public/policy explanation. Lastly, for our public/policy explanation category of virtue

signaling, we coded for responses that included reasons why participants believed other people

("the community") may not use sustainable transportation more, such as "more people would

commute by bike in my hilly neighborhood if it were easier to get up the hill." Included in this

category of coding were responses that cited public policy challenges such as "Safe places to

bike still remain a challenge." This category also included any explanation stemming from an issue with city infrastructure or lack of access.

In order to ensure that the coding scheme was reliable, two researchers coded three subsets of the responses individually, and then discussed any discrepancies and refined the coding scheme. A fourth subset with data from 71 participants was coded individually by the two coders without discussion and responses were then compared. The two coders agreed 89.6% of the time on their codes. The remainder of the data were coded solely by the primary researcher.

A third party was contacted to translate open-ended responses in Spanish. The translator provided English translations for all but one response that the translator was unsure of. After further discussion, the translator noted that the response did not have anything to do with transportation.

Results

Responses

Of the 2,692 opportunities to respond to the open-ended questions from the survey, 496 responses were received. Two-hundred and ninety-one of these responses were relevant to transportation; 205 comments were not. Of the 291 comments related to transportation, 89 did not include any type of virtue signaling. Within the remaining 201 comments we found 116 instances of "direct" virtue signaling, 73 instances of providing an explanation for one's own transportation choices, and 25 instances of either explanations for others' transportation habits or some type of public policy constraint. (Note that some responses included more than one type of virtue signaling per response.) A graphic of the breakdown of responses can be found in *Figure 1*.

Describing the Signalers

Of the 673 participants, 144 of them, or a little over 20% of the sample engaged in some form of virtue signaling. There were no significant gender effects when we compared virtue signalers to non-signalers among participants who provided gender info, $\chi^2 (2) = 2.38$, $p = .30$. Among the virtue signalers, 67% were female ($n = 97$), 19% were male ($n = 28$), and 8% identified as gender non-conforming ($n = 12$; see Table 1). In comparison, for the rest of the survey sample (those who did not leave any response or left a response that did not include any type of virtue signaling), 70% were female ($n = 261$), 25% were male ($n = 93$), and 5% were non-conforming ($n = 21$). Age tended to be slightly higher for signalers but not significantly so ($t(243) = -1.65$, $p = .10$) and ranged from 18 to 81 years old ($M = 51.3$, $SD = 15.3$; mean age for non-virtue signalers was 45.7, $SD = 15.9$).

Mirroring the whole sample, we found the modal household income category selected by virtue signalers was the "$75,000 or more annually" category. This was also the modal household income category for those who did not respond or made a comment that did not include any type of virtue signaling. Modal education level with 34.8% of virtue signalers was college graduate. Out of the 136 virtue signalers who were registered to vote, 127 reported having voted in the 2018 midterm general election, 9 did not, and 8 chose not to say.

Using the mean response to the seven items assessing attitudes and beliefs about environmental policy, on average, participants in the overall sample scored 3.81 out of a 5-point scale. A one-way ANOVA found that the combined group of virtue signalers (direct, personal excuse, and public/policy excuse) scored significantly higher on the beliefs and attitude measure ($M = 3.96$, $SD = .71$, $n = 142$) compared to those who didn't respond in the open-ended section ($M = 3.74$, $SD = .75$, $n = 275$), and those who responded but did not virtue signal ($M = 3.78$, SD

= .72, n = 117), $F(2, 671)$ = 4.60, p = .011, η^2 = .11. This finding indicates that virtue signalers

report being more likely to support initiatives to increase sustainable transportation use. Those

who scored higher on this measure were more likely to comment in the open-ended response

section and when they did comment, they were more likely to engage in virtue signaling.

Virtue signalers were more likely to report that biking or walking was their primary form

of transportation and were less likely to report driving as their primary form of transportation, as

compared to non-signalers, χ^2 (3) = 13.91, p = .003. Thirty-nine percent of virtue signalers

reported that driving was their primary method of transportation (n = 35), 34% reported walking

or biking as their primary method (n = 31), 19% primarily used public transit (n = 17), and 8%

engage in some other method of transportation (n = 7). In contrast, the majority (60%) of

participants who did not virtue signal reported that driving was their primary method of

transportation (n = 212), and only 26% walk or bike (n = 93), 11% use public transit (n = 40),

and 3% rely on some other method (n = 12; see Figure 2).

As noted above, survey respondents were given four open-ended questions, which gave

them four opportunities to virtue signal. When we tallied instances of virtue signaling across the

four questions, the average virtue signaler provided a "virtuous" response twice. When

analyzing type of virtue signaling and gender together, some forms of virtue signaling were more

frequent than others F (2, 509) = 12.28, p < .001. The most common method of virtue signaling

was direct virtue signaling, used an average of .81 times per signaler. Personal explanations or

excuses were used on average .51 times per signaler while explanations for others or public

policy constraints were used only .17 times per signaler. Of the three types of virtue signaling,

those who gave a personal explanation or excuse were significantly more likely to be female

(post hoc t-test, $t(123) = -2.21$, $p = .03$). However, there were no gender effects for the other two types of signaling (see Table 2).

Whole sample

Looking at the environmental attitudes and beliefs section for the whole sample, we initially thought young people would score higher (more support for sustainable actions, more in favor of protecting the environment) than older people. However, we found no correlation between age ($M = 47.2$, $SD = 15.9$) and environmental attitudes difference $r(493) = -0.16$, $p = .72$.

Discussion

With limited previous research, this exploratory study expanded the literature on virtue signaling by capturing and cataloging instances of this phenomenon. We have also expanded on who in the general population participates in virtue signaling. By comparing the demographic information from the survey responses from those who were coded as engaging in virtue signaling and those who were not, we did not uncover any significant demographic effects. We did however learn that in the context of this survey on transportation habits, the virtue signalers were more likely to hold pro-environmental attitudes and beliefs that supported their signals. They were also more likely to engage in actions that match their signaled virtues. They are more likely to walk, bike, and use public transit, and less likely to drive. The virtue signalers are matching their actions to their words.

In our quest to understand more about who among the general population is virtue signaling, we also aimed to further explore the various forms it may take. The open-ended responses from participants allowed us to delineate the responses into three categories of virtue signaling: "direct" virtue signaling, excuses for self, and explanations for others/public policy

constraints. Our categories of virtue signaling included responses that provided some type of excuse for not engaging in alternative methods of transportation. Excuse making or shifting causal attributions for negative personal outcomes from sources that are relatively more central to the person's sense of self to sources that are relatively less central, results in perceived benefits to the person's image and sense of control (Snyder & Higgins, 1988). Thus, this form of virtue signaling is consistent with an attempt to create an image that a person thinks is either more in line with their own values or would be perceived as more pro-social.

Similarly, explanations for others' behavior sit at the core of social functioning. When participants explained why they thought others did or did not engage in sustainable transportation habits, they may be trying to understand their social world as well as regulate their own behavior (Böhm & Pfister, 2015). Explaining either self-behavior or behavior by others can act as an indirect message to display what one cares about. Someone might not wish to communicate directly their moral beliefs because this could convey arrogance and ultimately have a back-firing effect that would damage their self-presentation (Kray, 2016), but through virtue signaling one is able to more subtly provide evidence of their character through expressing their endorsements.

In the context of our study, we narrowed the frame of coding for virtue signaling to the realm of sustainable transportation habits in order to help us understand someone's displayed virtue. An example response we received, "I would ride more frequently but the bike theft in this community is outrageous," highlights the responders wish to tell the researcher that they want to ride their bike more however they are unable to do so based on some external factor.

While the participants in the survey overall reported strong support for pro-environmental attitudes and beliefs, we were surprised to learn there were no effects of age. This may reflect the

established values and norms for Eugene, Oregon, in that we found a uniform high endorsement of pro-environmental attitudes across all ages in our sample. Eugene has historically been home to numerous climate justice advocacy groups including Breach Collective, Thriving Earth Exchange, and Friends of the Trees. Current climate recovery goals planned by the city include phasing out fossil fuel usage to 50% of what it was in 2010 as well as converting all City of Eugene owned facilities to carbon neutral (City of Eugene, n.d.). Eugene's established values might contribute to the percentage of our responders who signaled their environmentally friendly virtues. While these are established values among the city of Eugene, it is possible that our results may not be generalizable to other regions of the country that place less value on the importance of climate change.

Virtue signaling may also be a way of communicating and reinforcing social norms. Social norms specify to a group what is or is not acceptable (Bicchieri & Muldoon, 2011). For the individual, social norms may influence how they think and act, likewise how they think others should think and act (Tankard & Paluck 2016). When the individual recommunicates their approved social rules, their audience may also be motivated to conform to these norms. The same is true for virtue signaling. Virtue signaling communicates one's social norms to one's audience. The accumulation of many people coming together to reiterate social norms becomes a powerful force in social change (Tankard & Paluck 2016; Zaki & Cikara, 2020).

Audience, may, however matter. People may be more or less likely to virtue signal depending on the platform in which they are being asked to share. Griskevicius, et al. (2010) found consumers were more likely to purchase a "green" product compared to its generic counterpart when they were in a public store versus online shopping. This makes sense, as there is more opportunity to reap social rewards in a public setting. However, when given the

opportunity to commit to certain charities on personal social media sites, users were more likely to display their social goodness in order to impress others (Wallace & De Chernatony, 2020). Given that the platform for this survey was online and the results were likely to be viewed by researchers who were interested in sustainable transportation habits, people may have been more likely to display their prosocial behavior due to the likely chance of gaining audience approval.

We are not aware of other studies that examine virtue signaling captured in an incidental context. Previous studies have taken place in a controlled setting where participants are offered constrained choices as way to measure their virtue signaling, such as how they answer on a 7-point Likert scale asking them how likely they are to support a company that gave direct versus indirect support of climate sustainability (Gray, et al., 2020). In contrast, our study contributes to this area of research by further exploring what spontaneous information people will add when given the opportunity. Other studies were designed specifically as an attempt to elicit virtue signaling (Jung, et al., 2017; Zentz, 2021). Under the circumstances of our study, we are able to reduce experimenter effects by investigating virtue signaling after the survey had been published and completed.

One drawback from the incidental nature of this study is the inability to explore the motivations behind people's virtue signaling. *Why* people feel they need to declare their moral commitments is a question to be left to further research. We speculate a few motivating factors play a role in this phenomenon. First, is being able to identify one's in-group. Social identity theory states that a set of individuals who hold a common social identity may categorize themselves as a social group (Stets & Burke, 2000). And through the process of social comparison, a social group who shares a common identity are labeled an in-group (Stets & Burke, 2000). In-group identification supports a variety of benefits to the self, including

verifying and validating the self's own identity, as well as enhancing one's self-esteem (Stets & Burke, 2000). Virtue signaling provides these same benefits by allowing the signalers to identify themselves and attract their like-minded in-group.

Next, it is possible that virtue signaling stems from a gap in attitude and behavior. While our results found a significant number of virtue signalers participated in pro-environmental transportation habits that would match their pro-environmental signaling, some still may have felt the need to affirm their support for pro-environmental behavior after self-reporting a different account in the many other questions on the survey, that may have highlighted ways that participants' transportation behavior was not particularly pro-environment (e.g., rarely or never using public transportation; driving for errands that could be easily done on foot). Drawing on the theory of cognitive dissonance, inconsistence between thoughts of behaviors and thoughts of attitudes can create tension or uncomfortable feelings in people (Stone & Cooper, 2001). This tension may lead some people to either change their behaviors, change their attitude or engage in further strategies in order to resolve the uncomfortable feelings (Schrems & Upham, 2020). Given that people couldn't change their *past* behavior while taking our survey, they may have resorted to virtual signaling as a way to bolster their pro-sustainability credentials or signal their future intentions.

It is important to acknowledge that people may just want to be viewed positively (Baumeister, 1982) or think of themselves positively. While this is often where the pejorative part of the virtual signaling label comes from, there is plenty of previous literature that supports the idea that people's desire to be viewed positively is accompanied by social benefits (Griskevicius, et al., 2010; Nowak & Sigmund, 1998). Virtue signaling taken in a neutral sense may be a potent avenue for exploring further the power that appearing pro-social has on the

social sphere. If virtue signaling is reflecting social norms, it is possible that it may play a role in producing positive social change (Zaki & Cikara, 2020).

Regardless of reasoning, people are virtue signaling. In our sample, about 20% of people spent time and effort to offer additional information about their virtuous behaviors. While some may find virtue signaling annoying and none of us like being accused of it, virtue signaling may play a role in establishing social norms and values. Through this study we have learned more about who these virtue signalers are. They may be doing more than just signaling; they are reportedly matching their actions to their words. Our study found they are more likely to hold pro-environmental beliefs as well are more likely to use alternative methods of transportation. Looking ahead we are interested in exploring if virtue signaling may result in changes for the better by reflecting the changing norms. In a time when climate change is mounting as a real threat, virtue signaling may be a powerful tool in assisting us on our path to change.

References

Bartholomew, J. Easy Virtue. *The Spectator*, 18 Apr. 2015, www.spectator.co.uk/article/easy-

virtue.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological

Bulletin, 91*(1), 3-26.

Bicchieri, C., Muldoon, R., & Sontuoso, A., (2011). Social norms. Edward N. Zalta (Ed.), *The

Stanford Encyclopedia of Philosophy* (Winter 2018 Edition), Retrieved from

https://plato.stanford.edu/archives/win2018/entries/social-norms/.

Böhm, G., & Pfister, H. R. (2015). How people explain their own and others' behavior: A theory

of lay causal explanations. *Frontiers in Psychology*, 6.

https://doi.org/10.3389/fpsyg.2015.00139

City of Eugene, (2019), *Single-Use Ordinances*. https://www.eugene-or.gov/4235/Single-Use-

Ordinances

City of Eugene. (n.d.). *Sustainability in Eugene.* Eugene, OR Website. www.eugene-

or.gov/509/Sustainability-in-Eugene.

Energy Information Administration. (2019). *State-Level Energy-Related Carbon Dioxide

Emissions, 2005-2016*. Retrieved from

https://www.eia.gov/environment/emissions/state/analysis/

Gray, S. G., Sütterlin, B., Siegrist, M., & Árvai, J. (2020). The benefit of virtue signaling:

Corporate sleight-of-hand positively influences consumers' judgments about "social

license to operate". *Journal of Environmental Management*, *260*, 110047.

https://doi.org/10.1016/j.jenvman.2019.110047

Griskevicius, V., Tybur, J. M., & Van den Bergh, B. (2010). Going green to be seen: Status,

    reputation, and conspicuous conservation. *Journal of Personality and Social Psychology,*

    98(3), 392–404. https://doi.org/10.1037/a0017346

Jung, M. H., Nelson, L. D., Gneezy, U., & Gneezy, A. (2017). Signaling virtue: Charitable

    behavior under consumer elective pricing. *Marketing Science*, *36*(2), 187-194.

Kray, T. R. (2016). On name-dropping: The mechanisms behind a notorious practice in Social

    Science and the Humanities. *Argumentation*, *30*(4), 423-441.

Liere, K. D. V., & Dunlap, R. E. (1980). The social bases of environmental concern: A review of

    hypotheses, explanations and empirical evidence. *Public Opinion Quarterly*, *44*(2), 181-

    197.

Mkono, M. (2020). Eco-hypocrisy and inauthenticity: Criticisms and confessions of the eco-

    conscious tourist/traveller. *Annals of Tourism Research*, *84*, 102967.

    https://doi.org/10.1016/j.annals.2020.102967.

Nowak, M. A., & Sigmund, K. (1998). Evolution of indirect reciprocity by image

    scoring. *Nature*, *393*(6685), 573-577.

Oregon Global Warming Commission. (2018). *2018 Biennial Report to the Legislature*.

    Retrieved from

    https://static1.squarespace.com/static/59c554e0f09ca40655ea6eb0/t/5c2e415d0ebbe8aa6

    284fdef/154 6535266189/2018-OGWC-Biennial-Report.pdf

Palomo-Vélez, G., Tybur, J. M., & van Vugt, M. (2021). Is green the new sexy? Romantic of

    conspicuous conservation. *Journal of Environmental Psychology*, *73*, 101530.

    https://doi.org/10.1016/j.jenvp.2020.101530.

Schrems, I., & Upham, P. (2020). Cognitive dissonance in sustainability scientists regarding air travel for academic purposes: A qualitative study. *Sustainability*, *12*(5), 2-10.

Semmann, D., Krambeck, H. J., & Milinski, M. (2005). Reputation is valuable within and outside one's own social group. *Behavioral Ecology and Sociobiology*, *57*(6), 611-616.

Sezer, O., Gino, F., & Norton, M. I. (2018). Humblebragging: A distinct—and ineffective—self-presentation strategy. *Journal of Personality and Social Psychology*, *114*(1), 52.

Snyder, C. R., & Higgins, R. L. (1988). Excuses: Their effective role in the negotiation of reality. *Psychological Bulletin*, *104*(1), 23-31.

Stets, J. E., & Burke, P. J. (2000). Identity theory and social identity theory. *Social Psychology Quarterly*, 224-237.

Stone, J., & Cooper, J. A. (2001) Self-standards model of cognitive dissonance. *Journal of Experimental Social Psychology, 37,* 228–243.

Tankard, M. E., & Paluck, E. L. (2016). Norm perception as a vehicle for social change. *Social Issues and Policy Review*, *10*(1), 181-211.

Tosi, J., & Warmke, B. (2016). Moral grandstanding. *Philosophy & Public Affairs*, *44*(3), 197-217.

U.S. Census Bureau. (2018). *2013-2017 American Community Survey Five-year Estimates*. Retrieved from https://www.census.gov/newsroom/press-kits/2018/acs-5year.html

Wallace, E., Buil, I., & De Chernatony, L. (2020). 'Consuming good' on social media: What can conspicuous virtue signalling on Facebook tell us about prosocial and unethical intentions?. *Journal of Business Ethics*, *162*(3), 577-592.

Westra, E. (2021). Virtue signaling and moral progress. *Philosophy & Public Affairs*, *49*(2), 156-178.

Zaki, J., & Cikara, M. (2020). "Don't Be Afraid to Virtue Signal." *Time*, Time, 25 June 2020, time.com/5859459/in-defense-of-virtue-signaling-2/.

Zentz, L. (2021). "I AM HERE AND I MATTER" Virtue signaling and moral-political stance in progressive activists' Facebook posts. *Narrative Inquiry, 30.* https://doi.org/10.1075/ni.20117.zen.

Table 1
Prevalence of response types by gender

Gender	Did not respond	Responded; did not virtue signal	Virtue signaled in any way	Total
Female	184	77	97	358
Male	63	30	28	121
Gender non-conforming	13	8	12	33
Total	260	115	137	512

Note: 161 respondents did not provide any answer to the gender question and are not included above.

Table 2

Prevalence of different types of virtue signaling by gender

		Direct Signaling	Personal Excuse Explanation	Public/Policy
Female Mean		0.80	0.62	0.13
	SD	0.69	0.67	0.37
Male	Mean	0.82	0.32	0.11
	SD	0.48	0.48	0.32
Gender non-conforming Mean		0.92	0.25	0.59
	SD	0.90	0.45	0.99

*Note: Means are derived from the 137 participants who virtue signaled in any way.
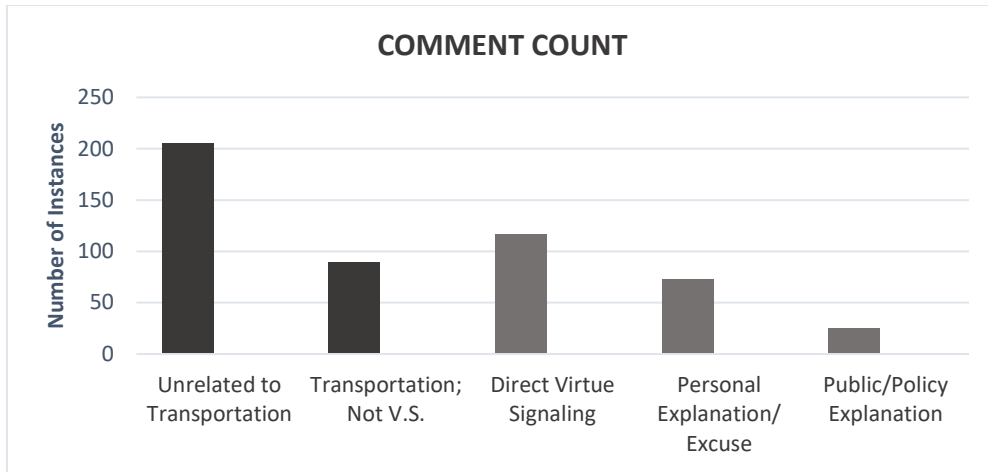
COMMENT COUNT

Figure 1.

Instances of each category were coded and counted from the survey responses received

SIGNALERS' PRIMARY METHOD OF TRANSPORTATION

Walk/Bike 34%

Drive 39%

Other 8%

Public Transit 19%

NON-SIGNALERS PRIMARY METHOD OF TRANSPORTATION
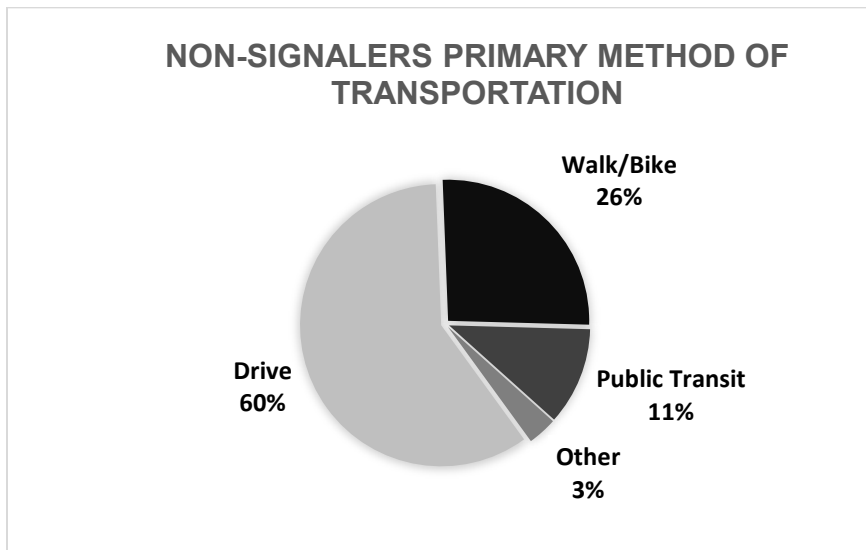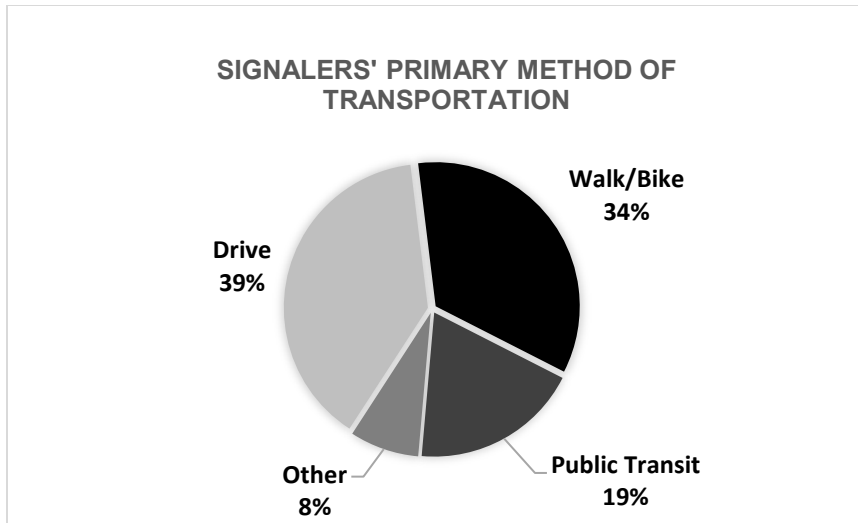
Walk/Bike 26%

Drive 60%

Public Transit 11%

Other 3%

Figure 2.

Virtue signalers (people who used any form of virtue signaling in their comments) were more likely to bike or walk and were less likely to use driving as their primary form of transportation than non-signalers

Appendix 1 - **Codebook**

Is text relevant to transportation? – yes/no
- Can be any form of transportation
- Response should be explicitly about transportation in some way
- Can talk about the amount of travel being increased/reduced
  (e.g. I am using way less gas…)
- Relevant- any trip to or from the home
- Can be a discussion of proximity to stores/errands but must also pertain to their thoughts of traveling there…
  - Yes- I am 10 miles to the nearest store
  - No- I live in Roseburg
- If response has to do with a previous survey item AND* provides more information on transportation habits, then it is still relevant

Personal Explanation-yes/no, if not relevant to transpo: leave blank
- Explaining why* they do/don't use other methods of transportation
- This can include for recreational purposes
- Yes, if they limit their sustainable transportation choices due to weather
- Yes, if Covid-19 is the excuse but they must explicitly say so

Public/Policy Explanation- yes/no, if not relevant to transpo: leave blank
- Explaining why they think <u>others</u> do/don't use alternative methods of transportation
- Explanation stems from an issue with city infrastructure/policy/lack of access

Virtue Signaling- yes/no, if not relevant to transpo: leave blank
- Specifically, about their care/consideration for the environment*
  - Does have to be directly related to sustainable transportation methods
  - Not any other moral cause such as Co-vid response, BLM, disabled rights, etc.
- Yes, if discussing their use of sustainable transportation as exercise
- Yes, if they increase their sustainable transportation due to weather
- Yes, if discussing preference/use of sustainable transportation
- Yes, if they mention limitation of *not going to work anymore* but they are still using sustainable transportation

Other notes:
- Explanation can be/usually is different from virtue signaling
- If response is a simple "lack of bike" or "I do not walk" then neither excuse or virtue signaling
- "Too poor to drive" is not excuse or virtue signaling

COVID-19 Impact on Transportation in Eugene/ Springfield Region

---

Start of Block: Part 1

Q1 **COVID-19 Impact on Transportation**
We are studying the impact of COVID-19 on everyday travel behavior (travel for work, school, errands, and recreation) -- before and during the pandemic. We are interested in how you traveled before the pandemic and how your use of public space for functional and recreational transportation (streets, paths and sidewalks) changed during shelter-in-place phase of the COVID-19 pandemic.   We are interested in hearing from residents of the Eugene/Springfield region who <u>lived in the Eugene/Springfield region from January 2020 to the present.</u>
 This survey should take between 20 and 30 minutes to complete. At the end of the survey, you can choose to enter a raffle to win one of ten $50 gift cards to a regional or national business. If you proceed but decide not to take the survey, you will still have an option to enter the raffle.
 This survey was developed by a team of researchers at the University of Oregon. Your answers are and will be completely confidential. Any personally identifying information will not be tied to any product this research produces. We will not share or sell your personally identifying information. By completing and submitting this survey you provide consent in allowing the University of Oregon research team to use these findings for research. You may choose not to participate in this survey without penalty.
 If you have any questions regarding the survey, please email Rebecca Lewis at the University of Oregon School of Planning, Public Policy and Management: rlewis9@uoregon.edu.
 Do you agree to participate in this survey? By clicking "Yes", you are consenting to participate in this survey and you are also informing us that you are age 18 or older. If you do not consent, or are not age 18 or older, please click "No" to navigate away from the survey.

○ Yes, I consent to participate in this survey (3)

○ No, I do not consent to participate in this survey (4)

Skip To: End of Block If Q1 = No, I do not consent to participate in this survey

Q76 **Key Definitions**

 For the purposes of this survey, we define types of travel in the following way: **Functional**

trips include traveling to work, school, religious events, social events, errands, sports, and organized activities.

 Recreational trips include walking, hiking, running or biking for the purpose of exercise, physical activity, walking a pet, or other recreational activities like bird-watching or flower-admiring.

 For the purposes of this survey, we are interested in your behavior during these periods: **Part 1:** Pre-COVID-19: January and February 2020 (unless otherwise specified) **Part 2:** During COVID-19:   **- Average Week** refers to a typical week during the weeks of shelter-in-place recommendations (March 16-May 14) **- Last 7 days** refers to the week before you took the survey.

Q2 Part 1 - Tell us about how you traveled pre-COVID-19 (specifically, in January and February 2020)

Q19 Describe your work or school situation in January/February 2020, before COVID-19 (select all that apply)

☐   I worked at a location that was different than my place of residence (1)

☐   I worked at my residence (2)

☐   My work location was split between my residence and a location different than my place of residence (3)

☐   I am a student (4)

☐   I am retired or do not work (5)

Skip To: Q8 If Q19 = I am retired or do not work

Display This Question:

    If Q19 = I am a student

Q20 **Did you live on campus?**

○ Yes (1)

○ No (2)

Q3 **Prior to shelter-in-place, how did you usually travel to <u>work and/or school</u>?** (Even if you use multiple methods, select what you consider as your **primary** method)

○ Motor vehicle (drive alone) (1)

○ Carpool (drive or ride with others) (2)

○ Ride-hail (taxi, Uber, Lyft) (3)

○ Transit (e.g. LTD bus, EmX, Ridesource or paratransit) (4)

○ Personally owned bicycle (5)

○ Bikeshare (PeaceHealth rides) (6)

○ Walk (7)

○ Work from home (8)

○ Other (Please specify) (9) _____

Q6 Prior to shelter-in-place, h**ow long did it take you to travel to <u>work and/or school</u> <u>one way</u> using your primary mode of travel?**

○ Less than 5 minutes (1)

○ 5-9 minutes (2)

○ 10-14 minutes (3)

○ 15-19 minutes (4)

○ 20-29 minutes (5)

○ 30-44 minutes (7)

○ 45-60 minutes (8)

○ More than 60 minutes (9)

Q7 Prior to shelter-in-place, w**hat was your target time to arrive at school or work?**

○ Before 7 am (1)

○ 7-8 am (2)

○ 8-9 am (3)

○ After 9 am (4)

Q4 Prior to shelter-in-place, a**part from your primary mode of travel, did you regularly use another mode of travel to get to and from <u>work and/or school</u> in January or February 2020 (i.e., a secondary travel mode)?**

○ Yes (1)

○ No (2)

Q5 Prior to shelter-in-place, w**hat was your <u>secondary</u> mode of travel to and from <u>work and/or school</u>? If you have more than one secondary mode, please choose only the secondary mode used most frequently.**

○ Motor vehicle (drive alone) (1)

○ Carpool (drive or ride with others) (2)

○ Ride-hail (taxi, Uber, Lyft) (3)

○ Transit (e.g. LTD bus, EmX, Ridesource or paratransit) (4)

○ Personally owned bicycle (5)

○ Bikeshare (PeaceHealth rides) (6)

○ Walk (7)

○ Other (Please specify) (8) _____

Q8 Prior to shelter-in-place, h**ow did you usually get <u>groceries?</u>**

○ Motor vehicle (drive alone) (1)

○ Carpool (drive or ride with others) (2)

○ Ride-hail (taxi, Uber, Lyft) (3)

○ Transit (e.g. LTD bus, EmX, Ridesource or paratransit) (4)

○ Personally owned bicycle (5)

○ Bikeshare (PeaceHealth rides) (6)

○ Walk (7)

○ Online Delivery (Please specify the platform or store.) (9)
_____

○ Other (Please specify) (8) _____

Skip To: Q12 If Q8 = Online Delivery (Please specify the platform or store.)

Q11 Prior to shelter-in-place, how **long did it take you to travel to where you buy <u>groceries</u>?**

     ◯ Less than 5 minutes (1)

     ◯ 5-9 minutes (2)

     ◯ 10-14 minutes (3)

     ◯ 15-19 minutes (4)

     ◯ 20-29 minutes (5)

     ◯ 30-44 minutes (7)

     ◯ 45-60 minutes (8)

     ◯ More than 60 minutes (9)

Q9 Apart from your usual mode of travel for <u>groceries,</u> did you use another mode of travel to get groceries (i.e., secondary travel mode)?

     ◯ Yes (1)

     ◯ No (2)

Display This Question:

     If Q9 = Yes

Q10 Prior to shelter-in-place, what was your secondary mode of travel to get groceries? If you have more than one secondary mode, please choose only the secondary mode used most frequently.

○ Motor vehicle (drive alone) (1)

○ Carpool (drive or ride with others) (2)

○ Ride-hail (taxi, Uber, Lyft) (3)

○ Transit (e.g. LTD bus, EmX, Ridesource or paratransit) (4)

○ Personally owned bicycle (5)

○ Bikeshare (PeaceHealth rides) (6)

○ Walk (7)

○ Online Delivery (Please specify the platform or store.) (9)
_____

○ Other (Please specify) (8) _____

Q12
Think about another pre-COVID-19 <u>routine trip</u> you took at least once a week, like going

to the gym, the coffee shop, kids activity, nail salon, or music lessons. What trip are you thinking of?

○ Exercise (gym, group exercise, etc.) (1)

○ Eating/Drinking establishment (2)

○ Clubs or activities (3)

○ Religious activity (church, synagogue, mosque, etc.) (4)

○ Personal care (hair salon, nail salon, massage) (5)

○ Kids activity (6)

○ Other (Please specify) (7) _____

Q13 Prior to shelter-in-place, w**hat form of transport did you usually take for this** <u>**routine trip**</u>**?**

○ Motor vehicle (drive alone) (1)

○ Carpool (drive or ride with others) (2)

○ Ride-hail (taxi, Uber, Lyft) (3)

○ Transit (e.g. LTD bus, EmX, Ridesource or paratransit) (4)

○ Personally owned bicycle (5)

○ Bikeshare (PeaceHealth rides) (6)

○ Walk (7)

○ Other (Please specify) (8) _____

Q14 Prior to shelter-in-place instructions during COVID-19, how frequently did you walk, run, hike, or jog in <u>your neighborhood</u> for recreation/physical activity? *(e.g., a walk, run, hike, or jog which began and ended at your house)*

○ I did not walk in my neighborhood (1)

○ Once a month (2)

○ Once a week (3)

○ 2-3 times a week (4)

○ Between 4-6 times a week (5)

○ Every day (6)

○ More than once a day (7)

Q15 Prior to shelter-in-place instructions during COVID-19, how frequently did you take a bicycle trip from <u>your home</u> for recreation/physical activity?

○ I did not bike (1)

○ Once a month (2)

○ Once a week (3)

○ 2-3 times a week (4)

○ Between 4-6 times a week (5)

○ Every day (6)

○ More than once a day (7)

Q18 Prior to shelter-in-place instructions, on average, how frequently did you use each of the following modes of transportation for <u>functional trips like work, errands and other trips</u> (not including walks or bike rides for recreation.)?

	Never (1)	A few times a year (2)	Once a month (3)	Once a week (4)	2-3 times a week (5)	Between 4-6 times a week (6)	Every day (7)	Multiple times per day (8)
Motor Vehicle (drive alone) (1)	○	○	○	○	○	○	○	○
Carpool (ride/drive with others) (2)	○	○	○	○	○	○	○	○
Ride-hail (Taxi, Uber, Lyft) (3)	○	○	○	○	○	○	○	○
Transit (LTD, EmX) (4)	○	○	○	○	○	○	○	○
Personally Owned Bicycle (5)	○	○	○	○	○	○	○	○
Peace Health Bikeshare (6)	○	○	○	○	○	○	○	○
Walk (7)	○	○	○	○	○	○	○	○

Q21 Is there anything else you'd like to add or clarify about your answers to the questions in Part 1 about your habits pre-COVID-19?

_____

Q68 Do you want to be entered in a drawing to win a $50 gift card? If yes, enter your email address.

_____

Q69 Are there any other details, comments, or clarifications that you'd like to add here that are related to this survey, the questions, and/or your answers?

_____

Q22 Part 2 – Tell us about how you have been traveling DURING "stay at home" shelter-in-place instructions during COVID-19.
For this section:- Average Week refers to a typical week during the weeks of shelter-in-place recommendations (March 16-May 14)- Last 7 days refers to the week before you took the survey.

Q23 Describe your work or school situation during COVID-19 shelter-in-place instructions (select all that apply).

☐   I am working my same job I had before in the same location. (1)

☐   I am working my same job I had before but now from my place of residence. (2)

☐   I am working my same job I had before but now from new location that is not my place of residence. (3)

☐   I am no longer working (due to decision not work, lack of work, or layoff) (4)

☐   I am a student (5)

☐   I am retired or do not work (6)

Skip To: Q26 If Q23 = I am working my same job I had before but now from my place of residence.

Skip To: Q26 If Q23 = I am retired or do not work

Skip To: Q26 If Q23 = I am a student

Skip To: Q26 If Q23 = I am working my same job I had before but now from new location that is not my place of residence.

Skip To: Q26 If Q23 = I am no longer working (due to decision not work, lack of work, or layoff)

Q24 **Was your job deemed "essential" during shelter-in-place instructions?**

○ Yes (1)

○ No (2)

○ I don't know (3)

Q25 **If were still leaving your residence to go to work during stay-at-home shelter in place instructions (March 16-May 14), did your mode of transportation change?**

○ If Yes, please describe: (1) _____

○ No (2)

Q26 **How many times have you used the following modes of transportation during shelter-in-place for all trips (functional and recreational):**

	In the last 7 days?	In an average week during shelter-in-place (3/16-5/14)
Motor Vehicle (drive alone) (1)	▼ 0 (1 ... >7 (9)	▼ 0 (1 ... >7 (9)
Carpool (drive or ride with others) (2)	▼ 0 (1 ... >7 (9)	▼ 0 (1 ... >7 (9)
Ride-hail (taxi, Uber, Lyft) (3)	▼ 0 (1 ... >7 (9)	▼ 0 (1 ... >7 (9)
Transit (LTD, EmX) (4)	▼ 0 (1 ... >7 (9)	▼ 0 (1 ... >7 (9)
Personally owned bicycle (5)	▼ 0 (1 ... >7 (9)	▼ 0 (1 ... >7 (9)
Peace Health Bikeshare (6)	▼ 0 (1 ... >7 (9)	▼ 0 (1 ... >7 (9)
Walk (7)	▼ 0 (1 ... >7 (9)	▼ 0 (1 ... >7 (9)

Q27 If you bought <u>groceries</u> in the past 7 days, how did you get them? <u>Select all that apply</u>

☐   Motor vehicle (drive alone) (1)

☐   Carpool (drive or ride with others) (2)

☐   Ride-hail (taxi, Uber, Lyft) (3)

☐   Transit (e.g. LTD bus, EmX, Ridesource or paratransit) (4)

☐   Personally owned bicycle (5)

☐   Bikeshare (PeaceHealth rides) (6)

☐   Walk (7)

☐   Delivery (8)

☐   Other (9) _____

☐   Did not buy groceries in the last 7 days (10)

Skip To: Q29 If Q27 = Did not buy groceries in the last 7 days

Q28 How have you obtained groceries in the past 7 days? <u>Select all that apply</u>

☐   At a physical store – you do the shopping (1)

☐   At a store for pick up – online ordering (2)

☐   At a store for pick up– phone ordering (3)

☐   Delivery – online ordering (4)

☐   Delivery – phone ordering (5)

Q29 Comparing the weekly frequencies of you going to a <u>physical store</u> (i.e., not online) for groceries now with same time last year, which of the following assessment applies?

○ I buy groceries in a physical store a lot less frequently (1)

○ I buy groceries in a physical store somewhat less frequently (2)

○ I buy groceries in a physical store about as frequently (3)

○ I buy groceries in a physical store somewhat more frequently (4)

○ I buy groceries in a physical store a lot more frequently (5)

Display This Question:

  If Q28 = Delivery – online ordering

  Or Q28 = Delivery – phone ordering

Q30 Compared to before COVID-19 shelter-in-place instructions, if you are currently using <u>delivery</u> for grocery shopping, how does your weekly frequency of using <u>delivery</u> as a method of grocery shopping compare with the same time last year?

○ I buy groceries via delivery a lot less frequently (1)

○ I buy groceries via delivery somewhat less frequently (2)

○ I buy groceries via delivery about as frequently (3)

○ I buy groceries via delivery somewhat more frequently (4)

○ I buy groceries via delivery a lot more frequently (5)

○ I have never used delivery for grocery shopping (6)

Q31 Compared to before COVID-19 shelter-in-place instructions how often have you been combining trips for multiple purposes (also known as "trip chaining" – e.g., running two separate errands on the same trip, or going to the grocery store on the way home from work)?

○ More frequently (1)

○ About the same (2)

○ Less frequently (3)

Q75
For the next set of questions, we'd like you to answer each question twice: FIRST, for what you have done during an AVERAGE WEEK during COVID-19 shelter-in-place orders (March 16-May 14) and SECOND, for you have done in the PAST 7 DAYS. (If the last 7 days were an average week since shelter-in-place began, then your answers may be the same; however when you answer "for the past 7 days," please respond with your actual behavior in the past 7 days, even if it wasn't a typical or average week.)

Q32 In an average week and over the last 7 days: How many times did you walk, run, hike, or jog <u>in your neighborhood</u> for recreation/physical activity? (e.g., a walk, run, hike, or jog which began and ended at your house)

In an average week during shelter-in-place (3/16-5/14) (1)	▼ 0 (1) ... >7 (9)
The last 7 days: (2)	▼ 0 (1) ... >7 (9)

Q33 In an average week and over the last 7 days: how frequently did you ride a bicycle <u>from your home</u> for recreation/physical activity?

In an average week during shelter-in-place (3/16-5/14) (1)	▼ 0 (1) ... >7 (9)
In the last 7 days: (2)	▼ 0 (1) ... >7 (9)

Q36 Aside from shelter-in-place orders, was there anything unusual in last 7 days that altered your usual patterns (for example: weather, illness, etc.)?

_____

Q77 Did your travel patterns change when Lane County entered Phase 1 of re-opening on May 15? If so, explain how: (skip if not applicable)

_____

Q37 Evaluate how much you have walked and biked during COVID-19 shelter-in-place orders (March 16-May 14**), compared to last year, using the following categories:**

	A lot less now than I did this time last year (1)	Somewhat less now than I did this time last year (2)	About the same now as I did this time last year (4)	Somewhat more now that I did this time last year (5)	A lot more now than I did this time last year (6)
Walk (1)	○	○	○	○	○
Bike (2)	○	○	○	○	○

Q38 Is there anything else you'd like to add or clarify about your answers to the questions in Part 2 about your habits during COVID-19 shelter-in-place?

_____

Q39 Part 3: Beliefs and attitudes

Q40 How accessible are the following places by <u>bicycle</u> for you:

	Don't know (1)	Very inaccessible 1 (2)	2 (3)	3 (4)	4 (5)	5 (6)	6 (7)	Very Accessible 7 (8)
Basic Goods (e.g. groceries) (9)	○	○	○	○	○	○	○	○
Outdoor facilities (e.g. trails, parks, etc.) (10)	○	○	○	○	○	○	○	○

Q78 How accessible are the following places by <u>walking</u> for you:

	Don't know (1)	Very inaccessible 1 (2)	2 (3)	3 (4)	4 (5)	5 (6)	6 (7)	Very Accessible 7 (8)
Basic Goods (e.g. groceries) (9)	○	○	○	○	○	○	○	○
Outdoor facilities (e.g. trails, parks, etc.) (10)	○	○	○	○	○	○	○	○

Q41 Indicate your agreement with the following statements:

	Strongly disagree (1)	Somewhat disagree (2)	Neither agree nor disagree (3)	Somewhat agree (4)	Strongly agree (5)
My neighborhood is a safe place to walk and/or bike. (1)	○	○	○	○	○
My neighborhood's layout (e.g., presence of sidewalks, changes in elevation, etc.) supports walking and biking. (2)	○	○	○	○	○

Q42 Indicate your agreement with the following statements:

	Strongly disagree (1)	Somewhat disagree (3)	Neither agree nor disagree (4)	Somewhat agree (5)	Strongly agree (6)
I believe government should adopt policies to reduce driving. (1)	○	○	○	○	○
I believe it's individuals' responsibility to reduce driving. (2)	○	○	○	○	○
I believe government should adopt policies to increase the share of Electric Vehicles (EVs). (3)	○	○	○	○	○
I believe government should adopt policies to increase the share of electric bicycles. (7)	○	○	○	○	○
I believe people in my community care more about environmental protection than economic growth. (4)	○	○	○	○	○
I believe that people driving less is good for the environment. (5)	○	○	○	○	○
I believe that air quality has improved in many U.S. cities because of the stay-at-home period. (6)	○	○	○	○	○

Q43 Given that many U.S. cities have seen improvements in air quality during stay-at-home orders, to what extent has <u>your support for practices </u>that would reduce driving (e.g. telecommuting, increased public transportation, fees, or bicycle infrastructure) been STRENGTHENED?

○ A great deal (22)

○ A lot (23)

○ A moderate amount (24)

○ A little (25)

○ None at all (26)

Q44 After the pandemic, do you believe people will use the following types of transportation:

	Less (1)	About the Same (2)	More (3)
Drive alone (1)	○	○	○
Carpool (2)	○	○	○
Transit (e.g. bus or EmX (3)	○	○	○
Bicycle (4)	○	○	○
Walk (5)	○	○	○

Display This Question:

If Q19 != I am retired or do not work

Or Q19 != I am a student

Q82 After COVID-19 is no longer a threat, which best describes your likelihood of working from home?

○ I'll work at home more than before COVID (1)

○ I'll work at home the same amount as before COVID (2)

○ I'll work at home less than before COVID (3)

○ Working from home is not relevant to my job (4)

○ I would like to work from home but my employer may not allow it (5)

Q45 Over the last 2 weeks, how often have you been bothered by the following problems?

	Not at all (1)	A few days (5)	Several days (2)	More than half the days (3)	Nearly every day (4)
Feeling nervous, anxious or on edge (1)	○	○	○	○	○
Not being able to stop or control worrying (2)	○	○	○	○	○

Q46 Is there anything else you'd like to add or clarify about your answers on beliefs and attitudes in Part 3?

_____

End of Block: Part 3: Beliefs and Attitudes

Start of Block: Part 4: Demographics

Q47 Part 4: Demographics

✳

Q48 Before COVID-19, how far did you live from your primary work location (in miles)?

_____

Q49 **What is your job sector? (select all that apply)**

☐ Education (Pre-K, K-12, Higher Education) (10)

☐ Financial Activities (Finance, Insurance, Real Estate/Rental) (4)

☐ Information (17)

☐ Leisure and Hospitality (Arts, Entertainment and Recreation; Accommodation and Food Services) (9)

☐ Other Services (16)

☐ Producing Goods (Natural Resources, Mining, Construction, Manufacturing) (18)

☐ Professional and Business Services (Professional, Scientific, Technical Services; Management; or Administration and Support, Waste Management and Remediation) (15)

☐ Public Administration (Local, State, Federal Government) (13)

☐ Retail or Wholesale Trade (1)

☐ Transportation and Utilities (14)

☐ Other (25) _____

☐ Student (26)

✱

Q50 **How many years have you lived in the Eugene/Springfield region?**

_____

Q51 Do you consider the Eugene/Springfield region your permanent home?

○ Yes (1)

○ No (2)

※

Q52 In what year were you born?

_____

Q53 What was your annual household income in 2019?

○ Less than $15,000 (1)

○ $15,000 to $24,999 (2)

○ $25,000 to $34,999 (3)

○ $35,000 to $49,999 (4)

○ $50,000 to $74,999 (5)

○ $75,000 to $149,999 (6)

○ $150,000 to $199,999 (7)

○ $200,000 or more (8)

○ Prefer not to say (9)

Q54 **What is your gender identity?**

○ Female (1)

○ Male (2)

○ Gender non-conforming (3)

○ Prefer not to say (4)

Q55 **What is your race/ethnicity?**

☐ White (2)

☐ Black or African American (3)

☐ Native Hawaiian or Pacific Islander (4)

☐ Latino/Hispanic (5)

☐ American Indian or Alaska Native (6)

☐ Asian (7)

☐ Other (8) _____

☐ Prefer not to say (1)

Q56 **What is your Zip Code?**

_____

Q57 **What is the closest intersection to your house?**

○ Street 1: (e.g., 13th Ave.) (1) _____

○ Street 2: (e.g., Patterson street) (2)
_____

Q58 **What is the highest degree/level of school you have completed?**

○ Less than high school graduate (1)

○ High school graduate (or equivalency) (2)

○ Some college or associate degree (3)

○ Bachelor's degree (4)

○ Graduate degree or higher (5)

Q59 **Indicate the number of people in your household (including yourself) that are:**

○ Age 0-4 (1) _____

○ Age 5-10 (2) _____

○ Age 11-13 (3) _____

○ Age 14-17 (4) _____

○ Age 18 or older (5) _____

Q60 **Do you rent or own the housing unit that you live in currently?**

○ Own (1)

○ Rent (2)

○ Other (3)

Q61 **How many vehicles (automobiles, vans, and trucks) are available for use by members of your household?**

▼ 0 (1) ... >6 (8)

✳

Q62 **How many pets does your household own?**

_____

✳

Q63 **How many people in your household have access to a personal bicycle?**

_____

Q64 **Are you registered to vote?**

○ Yes (1)

○ No (2)

Display This Question:

If Q64 = Yes

Q65 In 2018, did you vote in the midterm general elections?

    ◯ Yes (1)

    ◯ No (2)

Q66 Are you willing to participate in a follow up survey?

    ◯ Yes (1)

    ◯ No (2)

Q67 Would you be interested in participating in a focus group or personal interview? If selected, you will be compensated for your time with a $25 gift card.

    ◯ Yes (1)

    ◯ No (2)

End of Block: Part 4: Demographics

LIKE MOTHER, LIKE CHILD: INTERGENERATIONAL

TRANSMISSION OF MATERNAL EMOTION REGULATION

TO SIX-MONTH INFANTS

by

ANNALIESE ELLIOT

A THESIS

Presented to the Department of Psychology
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Arts

May 2022

An Abstract of the Thesis of

Annaliese Elliot for the degree of Bachelor of Arts
in the Department of Psychology to be taken May 2022

Title: Like mother, like child: Intergenerational transmission of maternal emotion
regulation to six-month infants

Approved: _____*Title and Full Name (or Title at end of Full Name)*_____
Primary Thesis Advisor

   Previous research demonstrates how maternal psychopathology is associated
with negative infant outcomes, however there is minimal research on intergenerational
transmission. Specifically, there is a lack of literature on intergenerational transmission
of emotion dysregulation. Emotion dysregulation has been demonstrated as a precursor
to future psychopathology in childhood and adulthood, therefore infancy is a crucial
time period to develop self-regulatory skills. This study aims to build upon previous
research to further understand how maternal emotion dysregulation predicts poor infant
regulation. This study examines the predictive association among maternal emotional
dysregulation reported prenatally during the third trimester and postnatally at six
months, with the Difficulties in Emotion Regulation Scale (DERS), and observations of
infant self-regulation postnatally (N = 221). Temperament, measured with the Infant
Behavioral Questionnaire (IBQ-R), was controlled to capture the independent
contributions of maternal dysregulation to infant's early indices of emotion regulatory
capacities. Infants' self-regulation and negative affect was measured at 6 months
postpartum with micro-analytic behavioral coding during the Still Face Paradigm (SFP),
a widely used paradigm to examine early relationship patterns between caregivers and

their infants. Although association between prenatal reports of maternal emotion regulation was not significantly related to infants' emerging regulatory capacities, maternal reports of concurrent dysregulation at 6-months postpartum was associated with poorer self-regulation in their infants. This finding suggests that emotion dysregulation can be transmitted across generations by postpartum mother-child interaction influences.

Acknowledgements

I would like to thank Dr. Jennifer Ablow, my Primary Thesis Advisor, for helping me to examine associations between maternal emotion regulation and infant self-regulation in this thesis. I have had the honor to work with her in the Developmental Sociobiology Lab here at the University of Oregon for over a year now. Dr. Ablow's lab has been working on the Prenatal Environment and Child Health (PEACH) Study, from which I obtained the data for this thesis. I have loved being a part of this project and appreciate the contributions of the DSL's research team. The PEACH study is a much larger, longitudinal study based at the Oregon Health and Science University, so I would also like to extend my gratitude to the research team there and, specifically, the primary investigators Dr. Elinor Sullivan and Dr. Hanna Gustafsson.

I would also like to thank the other members of my Thesis Committee including my Second Reader, Dr. Jeffrey Measelle, who helped me greatly in the data analysis process, as well as my CHC Representative, Dr. Carol Paty for her guidance and support in this process. I have been fortunate to receive guidance from them and Dr. Ablow in this strenuous but rewarding process. Lastly, I would like to thank my family for their continuous support and encouragement in all my education. I specifically want to thank my older sister Alexa for editing many thesis drafts and for being someone to turn to during this challenging process.

Table of Contents

List of Tables

Introduction

The inability to regulate or control one's emotions, or emotion dysregulation, is a demonstrated precursor to future psychopathology (Bridgett, Burt, Edwards, & Deater-Deckard, 2015). Development of these important regulatory capacities begins in infancy. Therefore, infancy is crucial time period to examine self-regulatory skills that could have longer, even lifetime implications. Previous research demonstrates how maternal psychopathology is associated with negative infant outcomes, however there is a lack of research on the transmission of maternal emotion dysregulation (Bush et al., 2017; Davis, Glynn, Waffarn, & Sandman, 2011; Korja, Nolvi, Grant, & McMahon, 2017; Thomas et al., 2017). This study aims to build upon previous research to further understand how maternal emotion dysregulation predicts poor infant regulation by examining the predictive association between reported maternal emotional dysregulation with observations of infant self-regulation. Understanding how a pregnant woman's dysregulation relates to her infant's developing regulatory strategies may provide insight into specific mechanisms through which risk for emotion dysregulation is transmitted across generations. From there, we can create targeted interventions to support the development of adequate self-regulation skills in infants.

Emotion regulation

Emotion regulation is characterized by attempts to control, suppress, re-evaluate, or amplify emotions in relation to personal goals such as socially appropriate behavior (Beauchanine, 2015; Crowell, Vlisides-Henry, & Kaliush, 2019; Fernandez, Jazaieri, & Gross, 2016). The ability to regulate one's emotions is adaptive, allowing for effective communication, learning, and the preservation of important social relationships.

Sheppes, Suri, & Gross (2015) outline a theoretical model of emotion regulation as (1) identifying the emotion, (2) selecting a regulatory strategy, (3) implementing the strategy, and (4) monitoring the strategy and adjusting it as needed. Thus, emotion dysregulation is when one has difficulties in any of the outlined steps and fails to adequately regulate their emotions (Cole, Dennis, Martin, & Hall, 2008; Fernandez, Jazaieri, & Gross, 2016). Emotion dysregulation is characterized by patterns of emotional experience and expression that are overly intense, unstable, rigid, or prolonged that ultimately impede appropriate behavior (Crowell, Vlisides-Henry, & Kaliush, 2019). Emotion dysregulation can result in issues such as poor communication that negatively affects one's relationships and overall well-being.

Emotion regulation starts early in life and is a vital skill to develop. However, infant regulation is very different from emotion regulation as an adult. Infants largely learn and develop emotion regulation abilities in the context of a social partner, which is typically in the form of a primary caregiver (Thompson & Goodman, 2010). A caregiver helps an infant co-regulate; for instance, if one's baby is distressed and crying, the caregiver may hold and rock the infant to soothe and calm them. Infants can also self-regulate in order to resolve physiological experiences of stress and distress. One way infants can regulate is through self-soothing strategies such as sucking or self-clasping, which have been found to reduce distress and frustration (Thomas et al., 2017). Additionally, infants can "distract" themselves from their distress by using exploratory strategies or disengaging from stressful stimuli. They continue to learn and develop these self-regulatory skills throughout childhood and adolescence (Beauchanine, 2015). Because emotion regulation is largely socialized, infants rely on

2

caregivers to aid in physiological and behavioral regulation before they can do so independently (Propper et al., 2008). These strategies are beneficial to overall development because self-regulation aids in emotional regulation.

Support from a social partner is crucial in the development of self-regulation skills. For instance, previous studies have demonstrated that poor parenting strategies, such as low warmth, can negatively reinforce self-regulation difficulties in infants which may underlie development of later behavioral or emotional problems (Beauchanine, 2015; Crowell, Vlisides-Henry, & Kaliush, 2019). Infants whose caregivers properly support emotion regulation tend to have better self-regulation abilities. Conversely, infants who have an inattentive caregiver, possibly due to forms of psychopathology such as depression, have less support and are therefore at higher risk of developing poor self-regulation skills (Crowell, Vlisides-Henry, & Kaliush, 2019).

The Still Face Paradigm

The Still Face Paradigm (SFP) is a widely used paradigm that mimics a situation where a caregiver is unavailable to attend to their infant, often resulting in infant distress, and in turn, infant self-regulation. The SFP begins with a baseline of normal play between the mother and infant. Next is the Still Face episode, in which the mother stares at the infant with an unresponsive, blank face. This breaks the infant's expectations of their caregiver being responsive to emotional or communication bids, typically resulting in a negative response from infants and attempts to self-regulate (Mesmen, van Ijzendoom, & Bakermans-Kranenburg, 2009; Tronick, Als, Adamson, Wise, & Brazelton, 1978). Lastly, in the Reunion period the mother becomes responsive

again and engages in normal play. The SFP creates a stressful situation that forces infants to use their coping capabilities when a parent is unavailable to aid in emotional regulation.

Typically, infants use cues like facial expression, crying, tone, and gestures to demonstrate distress to caregivers who are then expected to respond to the infants' needs (Cole, Martin, & Dennis, 2004; Thomas et al., 2017; Tronick, Als, Adamson, Wise, & Brazelton, 1978). Appropriate, sensitive caregiver response to infant cues supports the development of self-regulation skills that are necessary when a caregiver is unavailable, such as during the Still Face portion of the SFP (Conradt & Ablow, 2010). The use of self-regulatory actions, such as self-grasping, sucking, or diverting attention from the stressful stimuli, aids in reducing the infant's expression of negative affect (Maclean et al., 2014; Tronick, Als, Adamson, Wise & Brazelton, 1978). This demonstrates the positive effect of self-regulation on overall emotion regulation. Overall, the Still Face Paradigm has led to extensive research on how infants cope in stressful situations, highlighting the effectiveness of self-regulation strategies (Thomas et al., 2017).

In previous research, the SFP has been used to demonstrate meaningful differences in how infants attempt to manage stress when there is an unresponsive caregiver. One factor that leads to variation in infant responses and regulation is infants' experience with caregiver sensitivity, or caregiver responsiveness (Mesmen, van Ijzendoom, & Bakermans-Kranenburg, 2009). For instance, Kogan & Karter (1996) found that infants of sensitive mothers used more interpersonal regulation during recovery while infants of less sensitive mothers were more likely to use avoidant and

resistant types of regulation during the reunion. This indicates that infants of less responsive mothers are more likely to be self-reliant for emotional regulation. The authors interpreted that these infants may be used to their mothers not responding effectively to their cues of distress. Differently, infants of responsive mothers were found to benefit from positive interactions that strengthen the development of self-regulation skills (Kogan & Karter, 1996).

The SFP has also demonstrated meaningful differences in Respiratory Sinus Arrhythmia (RSA), which indicates emotion regulation abilities by estimating parasympathetic effect on heart rate (Ham & Tronick, 2006). Studies have indicated that infants who are able to recover from the still face episode usually have large increases in RSA during the Reunion episode, demonstrating higher emotion regulation (Kogan & Carter, 1996). In a study using the SFP, it was found that higher maternal sensitivity during the Reunion episode resulted in decreased heart rate and increased RSA, indicating effective emotional regulation (Conradt & Ablow, 2010; Propper et al., 2008). Additionally, infants of less responsive mothers had higher heart rates and lower RSA, which supports the similar effects found in the Mesmen, van Ijzendoom, and Bakermans-Kranenbur's study (2009). Overall, these findings demonstrate the necessity of both self-regulation and the use of caregivers for comfort when infants are in distress.

Intergenerational transmission of emotional regulator difficulties

Intergenerational transmission is the transmission of traits from a parent to their child. It occurs by different pathways, such as a biological transmission in utero or an environmental transmission postnatally through parent-child interactions (Thompson & Goodman, 2010). In past research, associations have been demonstrated between

maternal psychopathology and infant self-regulatory abilities. Findings indicate that

depressed mothers are more likely to have infants who use self-soothing strategies

rather than attentional strategies to self-regulate (Manian & Bornstein, 2009; Warnock

et al., 2016). Additionally, infants of depressed mothers who attempted gaze aversion

were more likely to revert to negative behaviors during the SFP, indicating a failure to

regulate emotionally. Similar findings have been demonstrated in infants of mothers

with high pregnancy anxiety or higher stressful life events and poor self-regulation in

their infants (Bush et al., 2017; Davis, Glynn, Waffarn, & Sandman, 2011; Korja,

Nolvi, Grant, & McMahon, 2017; Thomas et al., 2017). However, Thomas et al. (2017)

found that higher maternal sensitivity mitigated these results, suggesting that both

prenatal exposure and postpartum caregiving influence infant regulatory capacities.

Therefore, interventions supporting maternal warmth and sensitivity could reduce the

poor self-regulation in infants (Conradt & Ablow, 2010).

Furthermore, it has been demonstrated that stress hormones in mothers can

influence infant socioemotional development during infancy (Korja, Nolvi, Grant, &

McMahon, 2017). Bush et al. (2017) found associations among maternal reports of

stressful life events, perceived stress, and self-regulation in their infants. Overall, there

was lower self-regulation and higher RSA during the Still Face Paradigm in infants

whose mothers had higher stress. These findings suggest an intergenerational

transmission of the regulatory effects of adversity through biological mechanisms

(stressful life events preceding giving birth) as well as possible environmental effects by

measuring perceived stress postnatally as well as gestationally.

Binion and Zalewski (2018) studied preschoolers performing the Locked Box Task, which is designed to elicit anger and frustration. They found that maternal emotion dysregulation was associated with less talk, higher distraction, and less problem solving in children, which are all indicators of poor self-regulation that interfered with the task. This indicates an environmental transmission of emotion regulation. However, one issue the study discusses is bidirectionality. Binion and Zalewski note that temperament, or individual differences in behavior that are believed to be biologically based, has been suggested as another contributor to emotion regulation abilities. Therefore, the study was unable to conclude whether mothers' poor emotion regulation impacted children's poor self-regulation, or if children with difficult temperament worsened emotion dysregulation in mothers.

In a study investigating maternal emotion dysregulation in association to infant neurobehavior, Ostlund et al. (2019) found that women who reported higher emotion dysregulation prenatally had newborns with low attention and arousal. The study suggests that the low attention and arousal found in these infants may be for adaptive reasons; low reactivity would allow infants to more easily cope with an unresponsive parent. These findings were the first evidence of an association between maternal prenatal emotion dysregulation and dysregulation in newborns, supporting the idea of intergenerational transmission with potential prenatal routes of transmission, or transmission via biological factors. By examining the predictive association of maternal emotion dysregulation and infant self-regulation, the current study is looking to build upon these findings by Ostlund et al. (2019), aiming to further understand the mechanisms of the intergenerational transmission of emotion regulation. Overall,

infancy is a crucial period for development of emotional health. Studying this time can help us further understand what leads to typical and atypical development and allow for clinicians to identify infants and risk, work on prevention, and help provide interventions (Cole, Martin, & Dennis, 2004; Cole, Dennis, Martin, & Hall, 2008).

Current study

My thesis will be examining the specific question: Does prenatal perceived emotion dysregulation in expectant women predict poor self-regulation capacities and higher expressed negative affect in infants, over and above the effect of infant temperament? Infant temperament is being addressed because it is an additional factor shown to be related to emotion regulation (Binion & Zaleweksi, 2019). The thesis will discuss further implications of the study results by asking: Is emotion dysregulation transferred intergenerationally from mothers to infants? It is hypothesized higher reports of maternal perceived prenatal emotion dysregulation will be associated with infants utilizing less effective regulatory actions during the Still Face portion of the SFP, over and above the effect of maternal reported infant temperament. This study expanding upon the findings of Ostlund et al. (2019), which demonstrated the first evidence of intergenerational transmission of emotion dysregulation to newborns, by examining the association between prenatal and postnatal emotion regulation in mothers and self-regulation in their six-month infants.

Method

Participants

Participants for this study were recruited from Oregon Health & Science University's prenatal clinics and from PEACH study recruitment materials. Medical records were assessed to determine eligibility. All pregnant women between the ages of 18-40 who were in good health were deemed eligible for the study. 50.4% of the infants were female. In total, 266 pregnant women participated in the study with the mean age of 32.82 (SD = 4.13, range = 18.48 - 40.99). Upon entrance to the study, the mean highest school grade of completion for mothers was 16.13 (SD = 2.50, range = 9 - 20). 81.6% of the women were married and living with their spouse. 76.7% of mothers were white or Middle Eastern, 2.3% African American, 5.3% Hispanic, 1.5% American Indian or Alaska Native, 10.9% Asian, 2.3% Native Hawaiian or Pacific Islander, and 3.4% reported themselves as another group. Mother and infant demographics are reported in Table 1. below.

Table 1. *Sample Characteristics*

Infant	%		
Sex, % Female	50.4		
Racial Group / Ethnicity			
Hispanic	1.9		
American Indian / Alaska Native	1.9		
Asian / East Indian	13.2		

Native Hawaiian / Pacific Islander	2.3		
Black / African American	4.5		
White / Middle Eastern	78.9		
Other	4.5		
Mother	**Mean or %**	**SD**	**Range**
Age, years	32.82	4.13	18.48 – 40.99
Highest grade of regular school completed	16.13	2.50	9 – 20
Racial Group / Ethnicity			
Hispanic	5.30		
American Indian / Alaska Native	1.50		
Native Hawaiian / Pacific Islander	2.30		
Black / African American	2.30		
White / Middle Eastern	76.70		
Other	3.40		
Relationship Status			
Never been married	7.1		
Married, living with spouse	81.6		
Married, separated from spouse	0.8		
Divorced	2.6		

Other	2.3		

Procedure

The present study is part of a larger, longitudinal project investigating links

between maternal prenatal health and infant neurodevelopmental outcomes at Oregon

Health & Science University. Because of the inclusion of pregnant women and children

in this study, informed written consent was obtained and all procedures were approved

and overseen by the Institute Review Board of OHSU. Verbal consent was also required

for the virtual study visits (see further explanation below). For the current study,

mothers' emotional regulation was assessed prenatally during their third trimester of

pregnancy with a self-report questionnaire, specifically the Difficulties with Emotion

Regulation Scale (DERS; Gratz & Roemer, 2004) as well as at six months postpartum.

Prior to the onset of the COVID-19 pandemic, participants visited the laboratory located

at OHSU. However, due to the onset of the pandemic, from February 2020 – June 2020,

laboratory visits were adapted to remote data collection with the use of Zoom. Starting

in June 2020, participants were offered the choice of remote data collection or a visit to

the laboratory. Therefore, for the postnatal data collection, in which mothers engaged in

the Still Face Paradigm (SFP; Tronick, Als, Adamson, Wise & Brazelton, 1978) with

their infants at six months postpartum, families were invited to participate in tasks in the

laboratory playroom at OHSU (N = 69) or in their homes remotely over Zoom (N =

152). Video recordings of the SFP interactions were coded at the University of Oregon

(see description below). Participants were compensated for their time.

Measures

Maternal emotion dysregulation. Maternal emotion dysregulation was analyzed using the Difficulties with Emotion Regulation Scale (DERS; Gratz & Roemer, 2004). The DERS is a reliable and valid 36 item scale that asks participants to rate how statements concerning emotions relate to them on a scale from 1-5 in which 1 represents almost never (0%-10% of the time) and 5 indicates almost always (91%-100% of the time). Some of the statements included are "I have difficulty making sense of my feelings" and "When I'm upset, I start to feel very bad about myself." The higher total DERS score represents higher emotional dysregulation; a score of 70 reflects "nonclinical community average" and a score of 96 is "clinically significant emotion dysregulation" (Binion & Zalewski, 2018). 7.8% of the prenatal sample and 4.8% of the postnatal sample was above the clinical cutoff.

Infant temperament. Infant temperament was assessed by mothers completing the Revised Infant Behavior Questionnaire (IBQ-R; Gartstein & Rothbart, 2003) when infants were 6 months old. The IBQ-R is a widely used measure that has 14 total scales. For the present study, individual scales were analyzed and a Negative Affect composite score ($M = 3.00$). was created. The individual scales for the composite were Sadness, Distress to Limitations, or to what extent the infant shows distress in confining situations or when unable to perform a desired action, Fear, and Rate of Recovery. A Self-Regulation composite score ($M = 5.10$) also was used to determine self-regulation. Overall, the questionnaire has 191 items based on concrete actions that mothers rate on a scale of one (never occurs) to seven (always occurs). The scores for each scale are

summed up individually, therefore the higher a score on a specific scale indicates greater temperament in that dimension.

Infant self-regulation and negative affect. Infant self-regulation and negative affect was derived from micro-analytic behavioral codes obtained during the Still Face Paradigm (SFP; Tronick, Als, Adamson, Wise & Brazelton, 1978). The SFP is a well-established procedure that is designed to mimic situations in which caregivers are unable to attend to infants. Typically, the paradigm elicits distress and self-regulation in infants. The SFP starts with a baseline interaction of the mother and infant playing peekaboo for two minutes. The mother then turns away for 15 seconds and returns to her child for the still face episode, in which she maintains a blank face for two minutes. The mother turns away for 15 seconds again, then returns and plays with her infant as she normally would for two minutes in a reunion play episode. Only the still face portion was used for analysis since it typically results in negative reactions that can be soothed with self-regulation. The SFP was recorded and later coded in five second epochs (see description below).

Data coding

Infant emotional expressions and regulation were coded by a team of coders during the Still Face Paradigm, using a modified version of a developed coding scheme further discussed below (Holochwost, Gariepy, Proper, & Mills-Koonce, 2014). All sections of the paradigm, including the baseline, Still Face episode, and Reunion episode were coded, but for this study only the Still Face Episode was examined. All codes were made in time-synchronized five second epochs using Noldus software.

Emotional expressions. Infant emotion expressions were coded using a modified version of a published coding scheme (Holochwost et al., 2014; Moore et al., 2009). Facial, bodily, and vocalization affect were coded in five second epochs.

Facial expression coding options included neutral, facial joy, or facial distress. A neutral face was defined by no clear emotional expression. Facial joy was rated on a scale from one to three, with one being defined as a small smile with no involvement of other facial areas and three being defined as a large smile with stretch lips, bulged cheeks, and crinkled eyes. Facial distress was also rated on a scale from one to three. It includes expressions of sadness, anger, and frustration, with one being defined as distress only in one facial region and a three being defined as strong distress appearing in three regions of the face. If either distress or joy was detected, the expressed emotion was coded above the neutral face no matter how long it lasted during the 5 second epoch.

Bodily codes were determined on presence of bodily distress, presence of bodily joy, or a neutral body. Unlike the facial expressions, bodily codes were only coded for presence and not for intensity on a scale. Neutral body was coded when the body was relaxed and unemotional. Bodily joy was coded when there was a distinct increase in activity level, pointing or reaching at the mother, or actions such as clapping or bouncing. Bodily distress included signs of anger, sadness, or frustration. This included actions such as trying to escape the chair, tensing of the body, or kicking/hitting that was not rhythmic.

Vocal expressions were coded either as positive or distress, with neutral sounds being included in the positive codes. Distress vocalizations were mainly coded in

conjunction with visible facial or bodily distress. Both types of vocalizations were coded on a scale of one to three, with distress one representing short and low intensity protest of whining/whimpering and distress three representing high intensity crying that lasted most of the epoch. For positive vocalizations, a score of one indicated a short, low intensity positive or neutral sound and a score of three representing high intensity talking/laughing for a long duration. The coding levels are presented in Table 2.

Table 2. *Cues Used for Emotional Coding*

Expression	Emotion	Level Cues
Facial Affect	Joy	1. Small smile with no involvement of other facial areas
		2. Medium smile with slight bulge of cheeks and perhaps mouth open and/or crinkling around the eyes
		3. Large smile with stretch lips, bulged cheeks, crinkled eyes, and perhaps mouth open
	Distress	1. Low intensity distress expressed only in one facial region
		2. Distress expressed in two facial regions
		3. Impression of strong distress or distress appearing in three regions of the face
Bodily Affect	Presence of Joy	0. No sign of bodily joy

		1. Distinct increase in activity level, pointing or reaching at the mother, or actions such as clapping or bouncing
	Presence of Distress	0. No sign of bodily distress
		1. Signs of anger, sadness, or frustration. Trying to escape the chair, tensing of the body, or kicking/hitting
Vocalization	Positive	1. Short, low intensity positive or neutral sound
		2. Intermittent giggling, talking, or babbling. Lower intensity and duration
		3. High intensity talking/laughing for a long duration
	Negative	1. Short and low intensity protest of whining/whimpering
		2. Definite protest of moderate intensity
		3. High intensity crying for a long duration

Regulatory actions. Regulatory actions were coded using the same developed system. Regulatory behavior also was coded in five second epochs. If more than one action occurred in one epoch, the one with a longer duration was coded. The actions were:

Self-regulation (SR 1): evidence of infant behaviors to self soothe, such as actively mouthing or touching an object with an averted gaze. This could include sucking on a body part or object, pulling on clothes, rhythmic movements, or wringing hands together.

Exploration (SR 2): infants are actively touching and gazing at an object, such as their own body part or the chair.

Attention seeking (SR 3): infant is trying to get the caregiver's attention. The infant's gaze must be on the caregiver's face and there must be an exaggerated expression, movement, or vocalization.

Interrater agreement. There were four different coders for the SFP. Coders were trained to accomplish 80% agreement with a master coder. Once trained, continued interrater reliability was assessed throughout all videos coded to avoid coding drift and was determined on 25% randomly assigned videos. Percent agreement was calculated and resulted in averages of 84.95% agreement for facial expression, 84.95% agreement for bodily expression, 81.78% agreement for vocalization, and 82.20% for self-regulation. This demonstrates excellent agreement for all categories.

Missing Data. Because the task elicits distress, some infants were unable to complete the entire SFP. Additionally, some mothers did not follow the task protocol and took their infants out of the chair, causing the task to end early. There were eight videos with a short SF period and 20 videos with a recovery period significantly shorter than the normal two minutes. There were 12 videos missing the turnaround 2 section and two videos with no recovery. All of these videos were still used in the analyses.

Prior to conducting the central analyses for this thesis, families with no missing data were compared with families with any missing data using analysis of variance procedures (ANOVA). On the study's primary demographic, infant, and maternal data, there were no statistically significant mean level differences between families with and without missing data (F's ranged from 0.001 to 1.72, p's ranged from 0.97 to 0.19, respectively). 69.5% of the original sample was used in the final analyses (N = 185).

Data analysis plan

Statistical analyses were conducted using SPSS Version 28. Means, standard deviations, and proportions were used to characterize the sample for descriptive purposes. Bivariate correlations were used to study associations between the variables and to limit the number of covariates included in the final analysis models. Multiple regression analyses were used for the main analysis of the data. Regressions were run with predictors entered hierarchically. For each form of self-regulation, three models of increasing complexity were tested. Model 1 included the IBQ Negative Composite scores, the IBQ Regulation Composite scores, and infants' Total Distress score during the Still Face. The prenatal DERS completed during the participant's third trimester of pregnancy was added into Model 2, and the postnatal DERS completed at 6 months postpartum was added into Model 3. For each model. Unstandardized estimates (β), their standard errors (SE β), t values, and 95% confidence intervals are displayed to help analyze the unique contributions of each predictor.

Results

Preliminary analysis

Descriptive statistics are presented in Table 3. For the infants, this included

IBQ-R scores for the Negative Affect composite ($M = 3.00$, $SD = 0.59$) as well as the

Self-Regulation composite ($M = 5.10$, $SD = 0.51$). Additionally, descriptive statistics of

data collected during the SFP, including negative distress, broken down by the type of

distress shown (e.g., bodily, vocal, and facial) as well as the different types of self-

regulation exhibited by infants, are displayed in Table 3. In terms of infant's distress

during the SFP, the data in Table 3 suggests that a majority of infants showed distress in

all three modalities. Similarly, it was found that 98.5% of infants used at least some

form of self-regulation during this period. Mothers' prenatal DERS ($M = 65.69$, $SD =$

17.53) and postnatal DERS ($M = 62.82$, $SD = 17.96$) are reported, as well as the

percentage of mothers above a clinical cutoff for emotion dysregulation, which was

7.8% during the third trimester and 4.8% at six months postpartum. The variation within

the descriptive statistics necessitated further testing to decide what was important for

final analyses.

Intercorrelations were tested before the regression models to examine the

associations between predictor and outcome variables, which are presented in Table 4.

Expected significant negative associations were found between Self-Regulation (SR) 0

(no visible regulation) and SR 1 (self-soothing regulation), $r(221) = -0.53$, $p < 0.001$,

SR 0 and SR 2 (exploratory regulation), $r(221) = -0.52$, $p < 0.001$, and between SR 0

and SR 3 (attention seeking regulation), $r(221) = -0.15$, $p = 0.02$. All of these were

negative associations, indicating that as infants spent more time not regulating, other

forms of regulation decreased. It is interesting to note that the correlation coefficient (r) was much smaller between SR 0 and SR 3 in comparison to the other types of regulation, meaning it was a weak association. Similarly, there were statistically significant negative correlations between SR 1 (self-soothing regulation) and SR 2 (exploratory regulation), $r(221) = -0.33$, $p < 0.001$ as well as SR 1 and SR 3, $r(221) = -0.22$, $p < 0.001$, however both associations were fairly small. This means as self-soothing regulation increased, other forms of regulation decreased. There was a significant negative association found between IBQ-R composite scores for Negative Affect and Regulation, $r(262) = -0.25$, $p < 0.001$, indicating that higher scores of Negative Affect were correlated with lower scores of Self-Regulation.

There were also associations found with the DERS, or scores for maternal emotion regulation, at both time points. At the third trimester, the DERS was associated with the IBQ-R for Negative Affect, $r(232) = 0.23$, $p < 0.00$, meaning higher scores of maternal dysregulation were associated with higher Negative Affect scores in their children. The prenatal DERS was negatively associated with the IBQ-R for Self-Regulation, $r(232) = -0.16$, $p = 0.02$, which indicates that higher scores of maternal dysregulation were associated with lower scores for infant Self-Regulation. However, the DERS at the third trimester was not associated with any types of self-regulation during the Still Face. The postnatal DERS at six months was similarly associated with the IBQ-R for Negative Affect, $r(249) = 0.33$, $p < 0.001$ and Regulation, $r(249) = -0.15$, $p = 0.02$. While the postnatal DERS scores increased as Negative Affect scores increased, the DERS decreased as Regulation scores increased. In contrast to the lack of association found between the prenatal DERS and observations of infants' emotion

regulation during the SFP, the postnatal DERS was associated with observed SR 0 (no visible regulation), $r(208) = 0.21$, $p = 0.003$, suggesting that as maternal scores for emotion regulation increase, infants spend less time regulating during the SFP. There was also a small but significant negative correlation between SR 1 (self-soothing regulation) and the postnatal DERS, $r(208) = -0.15$, $p = 0.03$, indicating that higher scores of maternal dysregulation were associated with infants spending less time regulating by self-soothing tactics during the SFP. As expected, the prenatal and postnatal DERS were positively associated, $r(222) = 0.72$, $p = < 0.001$.

To explore possible sample related sources of individual differences in both infant self-regulation and maternal dysregulation, associations with demographics were examined. This included infant ethnicity (Hispanic or not), mother's highest year of regular schooling (Mother School in Table 4), and infant sex. A small but significant association was found between mother's highest year of education and infant ethnicity, $r(228) = 0.14$, $p = 0.04$. However, neither variable was significantly associated with the main variables in question (i.e., infant regulation during the SFP, prenatal and postnatal DERS, etc.), as such, they were not included as covariates in the primary analyses. Child sex was also included in the preliminary analysis and yielded no significant correlations. Given the significant associations found among maternal reports of dysregulation, infant temperament, and infant's observed self-regulation, further testing of regression models was appropriate.

Table 3. *Descriptive Statistics*

Infant	M or %	SD	Range
IBQ-R			

Negative Affect Composite	3.00	0.59	1.44 – 5.02	
Self-Regulation Composite	5.10	0.51	3.50 – 6.47	
Sadness	3.33	0.84	1.13 – 6.00	
Distress to Limitations	3.36	0.75	1.50 – 5.50	
Fear	2.41	0.88	1.14 – 5.81	
Rate of Recovery	4.01	0.99	1.08 – 7.00	
Self-Regulation				
% used no SR	1.50			
0 - no self-regulation visible	0.38	0.26	0.00 – 1.00	
1 - self regulation	0.35	0.24	0.00 – 1.00	
2 - exploratory	0.19	0.23	0.00 – 1.00	
3 - attention seeking	0.08	0.12	0.00 – 0.50	
% used any SR	98.50			
Facial Distress				
% no facial distress visible	38.40			
1 - low intensity	0.09	0.11	0.00 – 0.67	
2 - medium intensity	0.08	0.15	0.00 – 0.78	
3 - high intensity	0.04	0.12	0.00 – 0.83	
% showed any facial distress	61.60			

Bodily Distress				
No bodily distress visible	22.60			
% showed any bodily distress	77.40			
Vocal Distress				
No vocal distress visible	5.00			
1 - low intensity	0.20	0.23	0.00 – 1.00	
2 - medium intensity	0.16	0.23	0.00 – 1.00	
3 - high intensity	0.02	0.06	0.00 – 0.52	
% showed any distress vocalization	95.00			
Mother	M sum	SD	Range	% above clinical
DERS 3rd trimester	65.69	17.53	36.00 – 121.00	7.80
DERS at 6 months postpartum	62.82	17.96	36.00 – 125.00	4.80

Table 4. *Intercorrelations*

	IBQ Neg	IBQ Reg	SR 0	SR 1	SR 2	SR 3	Total Distres	Prenatal DERS	Postnatal DERS	Baby Ethnicit	Mother School	Child Sex
IBQ Neg	-											
IBQ Reg	- 0.25	-										

	1	2	3	4	5	6	7	8	9	10	11	12
SR 0	0.13	0.02	-									
SR 1	-0.01	0.03	-0.53	-								
SR 2	-0.14	-0.10	-0.52	-0.33	-							
SR 3	0.01	0.10	-0.15	-0.22	-0.10	-						
SF Total Distress	0.05	0.10	0.13	0.07	0.00	-0.03	-					
Prenatal DERS	0.23**	-0.16	0.08	-0.07	-0.01	-0.02	-0.01	-				
Postnatal DERS	0.33**	-0.15	0.21**	-0.15	-0.07	-0.01	0.02	0.72**	-			
Baby Ethnicit	-0.01	-0.07	0.12	-0.09	0.03	-0.14	-0.11	-.04	-0.09	-		
Mother School	-0.01	-0.14	0.04	-0.06	0.02	-0.01	-0.05	0.09	0.01	0.14*	-	
Child Sex	0.08	0.30	-0.01	0.03	-0.08	0.10	-0.05	0.02	0.07	-0.13	0.02	-

* p < 0.05

** p < 0.01

*** p < 0.01

Three regression models of increasing complexity were used to test the proposed associations across each level of self-regulation during the Still Face, which can be seen in Table 5. Model 1 consisted of the possible covariates, which were the IBQ-R for Negative Affect, Self-Regulation, and Total Infant Distress during the SFP. Self-regulation (SR) 0, which was when infants showed no sign of regulation, was borderline significantly associated with the IBQ-R for Negative Affect ($\beta = 0.06$, $t = 1.19$, $p <$

0.10) and Total Distress ($\beta = 0.02$, $t = 1.84$, $p < 0.10$) in Model 1. In Model 2, we added the Prenatal DERS during the third trimester. In this model, SR 0 (no regulation) was only borderline significantly associated with Total Distress during the SFP ($\beta = 0.03$, $t = 1.88$, $p < 0.10$). Lastly, Model 3 added the Postnatal DERS taken at six months. Again, a borderline significant association was found between SR 0 and Total Distress ($\beta = 0.03$, $t = 1.80$ $p < 0.10$). Most interesting was the significant association found between SR 0 and the postnatal DERS ($\beta = 0.00$, $t = 2.38$, $p < 0.05$), meaning that infants who spent more time not regulating had mothers with higher DERS scores at six months, above and beyond the effects of IBQ-R scores, Total Distress, and the prenatal DERS.

For Self-Regulation 1, which is regulation through self-soothing strategies such as self-grasping or sucking, there were no significant associations found in Model 1 or Model 2. However, in Model 3, a significant negative association was found with the Postnatal DERS ($\beta = -0.00$, $t = -2.26$, $p < 0.05$). This indicates that infants who spent less time using self-soothing regulatory strategies was associated with higher maternal postnatal DERS scores, when controlling for IBQ-R scores, Total Distress during the SFP, and the Prenatal DERS.

Outcomes for Self-Regulation 2, regulation by exploration, and Self-Regulation 3, attention seeking, for each model can be found in Table 5. There were no significant associations found between SR 2 or SR 3 with the Prenatal or Postnatal DERS.

Table 5. Regression Models

MODELS	Self-Regulation 0				Self-Regulation 1				Self-Regulation 2				Self-Regulation 3			
	B	(SE B)	t	[95% CI]	B	(SE B)	t	[95% CI]	B	(SE B)	t	[95% CI]	B	(SE B)	t	[95% CI]
Model 1																
IBQ Neg	0.06	(0.05)	1.15#	[-0.00, 0.12]	-0.01	(0.03)	-0.16	[-0.06, 0.05]	-0.06	(0.03)	-2.34*	[-0.11, -0.01]	0.01	(0.01)	0.46	[-0.02, 0.05]
IBQ Reg	0.02	(0.04)	0.61	[-0.05, 0.09]	0.01	(0.03)	0.22	[-0.06, 0.07]	-0.05	(0.03)	-1.79#	[-0.12, 0.01]	0.03	(0.02)	1.55	[-0.01, 0.06]
SF Total Distress	0.02	(0.01)	1.9##	[-0.00, 0.05]	0.01	(0.01)	1.06	[-0.01, 0.04]	-0.03	(0.01)	-2.93**	[-0.05, -0.01]	-0.00	(0.01)	-0.64	[-0.02, 0.01]
	$F(3,219)=2.59, r^2=0.04, p=0.05$				$F(3,219)=0.40, r^2=0.01, p=0.05$				$F(3,219)=5.91, r^2=0.08, p=0.05$				$F(3,219)=0.81, r^2=0.01, p=0.05$			
Model 2																
IBQ Neg	0.05	(0.04)	1.54	[-0.02, 0.12]	-0.02	(0.03)	-0.73	[-0.09, 0.04]	-0.04	(0.03)	-1.95	[-0.10, 0.02]	0.01	(0.02)	0.63	[-0.02, 0.04]
IBQ Reg	0.01	(0.04)	0.9#	[-0.07, 0.09]	-0.02	(0.04)	-0.41	[-0.09, 0.06]	-0.03	(0.04)	-0.87	[-0.10, 0.04]	0.03	(0.02)	1.69	[-0.01, 0.07]
SF Total Distress	0.03	(0.01)	1.8#	[-0.00, 0.05]	0.01	(0.01)	0.90	[-0.01, 0.04]	-0.03	(0.01)	-2.74**	[-0.05, -0.01]	-0.01	(0.01)	-0.77#	[-0.02, 0.01]
Prenatal DERS	0.00	(0.00)	0.74	[-0.00, 0.00]	-0.00	(0.00)	-0.84	[-0.00, 0.00]	0.00	(0.00)	0.09	[-0.00, 0.00]	-0.00	(0.00)	-0.10	[-0.00, 0.00]
	$F(4,186)=1.97, r^2=0.04, p=0.05$				$F(4,186)=0.57, r^2=0.01, p=0.05$				$F(4,186)=2.62, r^2=0.05, p=0.05$				$F(4,186)=0.84, r^2=0.02, p=0.05$			
Model 3																
IBQ Neg	0.04	(0.04)	0.98	[-0.04, 0.11]	-0.00	(0.03)	-0.09	[-0.07, 0.06]	-0.04	(0.03)	-1.20	[-0.10, 0.03]	0.01	(0.02)	0.91	[-0.02, 0.04]
IBQ Reg	0.03	(0.04)	0.77	[-0.05, 0.12]	-0.03	(0.04)	-0.72	[-0.10, 0.05]	-0.04	(0.04)	-1.11	[-0.11, 0.03]	0.04	(0.02)	1.85#	[-0.00, 0.07]
SF Total Distress	0.03	(0.01)	1.80#	[-0.00, 0.05]	0.01	(0.01)	0.79	[-0.02, 0.04]	-0.03	(0.01)	-2.58*	[-0.05, -0.01]	-0.00	(0.01)	-0.60	[-0.02, 0.01]
Prenatal DERS	-0.00	(0.00)	-0.91	[-0.00, 0.00]	0.00	(0.00)	0.85	[-0.00, 0.00]	0.00	(0.00)	0.09	[-0.00, 0.00]	0.00	(0.00)	0.14	[-0.00, 0.00]
Postnatal DERS	0.00	(0.00)	2.3#*	[0.00, 0.01]	-0.00	(0.00)	-2.36*	[-0.01, 0.00]	0.00	(0.00)	0.24	[-0.00, 0.00]	0.00	(0.00)	-0.27	[-0.00, 0.00]
	$F(5,179)=2.77, r^2=0.07, p=0.05$				$F(5,179)=1.98, r^2=0.4, p=0.05$				$F(5,179)=2.02, r^2=0.05, p=0.05$				$F(5,179)=0.80, r^2=0.02, p=0.05$			

** correlation is significant at the 0.01 level

* correlation is significant at the 0.05 level

correlation is significant at the 0.10 level

26

Discussion

This study sought to examine associations between maternal reported emotional dysregulation and the self-regulation abilities of their six-month-old infants. It was hypothesized that high levels of emotion dysregulation, reported by mothers during their third trimester of pregnancy, would be associated with lower levels of observed infant self-regulation abilities at six months during the Still Face Paradigm. The study found no association between any type of self-regulation and mothers' reported emotion dysregulation at their third trimester of pregnancy. As such, this study was unable to support the findings of Ostlund et al. (2019) which found that newborns of mothers reporting high dysregulation prenatally had low attention and arousal, indicating poor self-regulation. The previous study suggested a prenatal pathway of emotion regulation transmission, which the current study findings did not support. Instead, an association was found between maternal reported emotion dysregulation postnatally at 6 months and infant observed self-regulation during the Still Face.

Specifically, as mothers reported greater dysregulation, the absence of self-regulation in infants (self-regulation 0) also increased. This supports the idea that a more dysregulated mother appears to have a baby that is poorer at self-regulation. This conclusion is further supported by the negative association found between the maternal dysregulation and self-soothing regulation in infants, meaning that mothers who report high levels of dysregulation tend to have babies that engage in very little self-soothing when distressed during the SFP. These findings support the results from Binion and Zalewski (2018), which found that more dysregulated mothers were more likely to have preschoolers with poor self-regulation skills. Additionally, the current findings align

with those of Thomas et al. (2017), which found that high maternal sensitivity, in spite of psychopathology that results in poor emotion regulation, to be a mitigating factor in relation to infant self-regulation. This demonstrates how caregiving skills, such as being responsive and attentive to one's child, can influence infant development of self-regulation. All of these studies as a whole support the idea that environmental factors, such as caregiving skills or mother-infant interaction, impact infant self-regulation abilities. Overall, the present study supports the previous literature demonstrating a relationship between maternal dysregulation and infant self-regulation.

Previous studies have demonstrated that infants rely on their caregiver to support the development of their self-regulation capacities (Thompson & Goodman, 2010). This study further supports that relationship by demonstrating how mothers' emotion dysregulation is associated with infant's lack of self-regulation. When an infant depends upon a dysregulated caregiver, they are unable to properly co-regulate with their caregiver, resulting in poor regulation. This study illustrates how important adequate emotion regulation is in mothers to aid their infants' self-regulation capacities.

The multi-method approach to data collection was a strength of this study. While maternal data was collected by a questionnaire, infant self-regulation was assessed through observed behavior during the Still Face Paradigm, a paradigm designed to elicit negative affect and self-regulation in infants. In addition to the behavior analysis, mothers were able to report infant temperament through the IBQ-R. This multimethod approach of data collection decreases possible inaccuracies influencing the data due to self-report. For instance, if a mother has poor emotion regulation, maybe she is more likely to see her infant's temperament in a negative light.

Using behavioral observation helped mitigate these possible biases. Moreover, microcoding in epochs of five seconds, rather than a more generalized global coding system, allowed for the detection of more subtle or nuanced forms of self-regulation. An additional strength was the study sample, which was fairly large and included mothers with a wide range of emotion dysregulation (see Table 3).

A strength in the study design was the use of a prenatal measure, which allowed for potential predictive associations to be studied. Collecting mothers' perceived dysregulation both prenatally and postnatally allowed us to examine whether associations between maternal dysregulation and infants emerging regulatory capacities were biological, possible due to prenatal programming, or environmental and occurring due to socialization. In this study, mothers' prenatal reports of their dysregulation, however, was not associated with postnatal infant regulatory behavior, whereas their postnatal reports were. Given the association between mothers' reports of postnatal dysregulation and their infants observed regulatory capacities, bidirectionality must be considered. It cannot be assumed that maternal dysregulation results in an infant's poor regulation, because there is also the possibility that a dysregulated infant causes higher stress in a mother and worsens her regulation skills (Binion & Zalewski, 2019). Future studies could further dissect these associations by including earlier prenatal maternal measures, such as during the first trimester, to test a predictive association. Similarly, future studies could get an earlier measure of infant regulation, for instance, Ostlund et al. (2019) measured maternal emotion dysregulation after the 25th week of pregnancy and measured infants 24 hours after birth at minimum. Although it is difficult to measure self-regulation in newborns, any associations found would demonstrate a

biological transmission rather than impact of environmental factors. Lastly, the issue could also be addressed by creating an intervention that targets mothers with high emotion dysregulation. The study could include pre- and post-intervention analyses to determine whether improving mothers' emotion regulation skills also improved their child's self-regulation capacities.

A limitation of the study is the lack of demographic diversity within the sample. As shown in Table 1., almost 80% of the study participants were Caucasian and more than 80% were married and living with their spouse. Therefore, the study could be missing possible sources of individual differences that could influence emotion regulation. Moreover, this makes the study findings less generalizable to a more diverse population. A more generalizable analysis should include more diversity in order to better detect possible covariates and individual differences. A final limitation of this study is the possible influences of self-report bias. As previously mentioned, using a multimethod approach for measuring infant regulation helped avoid mothers' bias when reporting on their child's behavior. However, only self-report was used to measure maternal emotion regulation. Self-serving reports could lead to an inaccurate depiction of their regulatory abilities. To address this problem, future studies could include behavioral measures and observations of maternal emotion regulation or clinical ratings.

In conclusion, the findings of this study support the previous literature that demonstrates maternal emotion dysregulation as a risk factor for their children's self-regulation abilities. We found that infants of more dysregulated mothers tend to have infants who self-regulate less in a stressful situation. Because this association was found with mothers and infants postnatally, this suggests that emotion dysregulation may be

transmitted through external, environmental factors. Overall, this study supports the

idea of intergenerational transmission of emotion dysregulation from mothers to their

children. This demonstrates a need for interventions targeted at emotion regulation of

new mothers.

Bibliography

Beauchaine, T.P. (2015). Future directions in emotion dysregulation and youth psychopathology. *Journal of Clinical Child and Adolescent Psychology, 44*, 875–896.

Beauchaine, T.P., & Zisner, A. (2017). Motivation, emotion regulation, and the latent structure of psychopathology: An integrative and convergent historical perspective. *International Journal of Psychophysiology, 117*, 108–111.

Binion, G., & Zalewski, M. (2018). Maternal emotion dysregulation and the functional organization of preschoolers' emotional expressions and regulatory behaviors. *Emotion, 18*(3), 386–399. https://doi.org/10.1037/emo0000319

Bush, N., Jones-Mason, K., Coccia, M., Caron, Z., Alkon, A., Thomas, M., . . . Epel, E. (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.

Bridgett D.J., Burt N.M., Edwards E.S., Deater-Deckard K. (2015). Intergenerational transmission of self-regulation: A multidisciplinary review and integrative conceptual framework. *Psychology Bulletin, 141*(3), 602-654.

Cole, P. M., Dennis, T. A., Martin, S. E., & Hall, S. E. (2008). Emotion regulation and the early development of psychopathology. In M. Vandekerckhove, C. von Scheve, S. Ismer, S. Jung, & S. Kronast (Eds.), Regulating emotions: Culture, social necessity, and biological inheritance (pp. 171–188). Blackwell Publishing. https://doi.org/10.1002/9781444301786.ch7

Cole, P., Martin, S., & Dennis, T. (2004). Emotion regulation as a scientific construct: methodological challenges and directions for child development research. *Child Development, 75(2*), 317-333.

Conradt, E., & Ablow, J.C. (2010). Infant physiological response to the still-face paradigm: Contributions of maternal sensitivity and infants' early regulatory behavior. *Infant Behavior and Development, 33*, 251-265.

Davis, E. P., Glynn, L. M., Waffarn, F., & Sandman, C. A. (2011). Prenatal maternal stress programs infant stress regulation. *Journal of child psychology and psychiatry, and allied disciplines, 52*(2), 119–129. https://doi.org/10.1111/j.1469-7610.2010.02314.x

Crowell, S. E., Vlisides-Henry, R. D., & Kaliush, P. R. (2019). Emotion generation, regulation, and dysregulation as multilevel transdiagnostic constructs. *The Oxford Handbook of Emotion Dysregulation*, 84–97. https://doi.org/10.1093/oxfordhb/9780190689285.013.7

Fernandez, K.C., Jazaieri, H., & Gross, J.J. (2016). Emotion regulation: A transdiagnostic perspective on a new RDoC domain. *Cognitive Therapy and Research, 40*, 426–440.

Gartstein, M.A., Rothbart, M.K., 2003. Sudying infant temperament via the revised infant behavior questionnaire. *Infant Behavior Development, 16*, 64-86.

Gratz, K. L., & Roemer, L. (2004). Multidimensional assessment of emo- tion regulation and dysregulation: Development, factor structure, and initial validation of the difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment, 26*, 41–54. http://dx.doi.org/10.1023/B:JOBA.0000007455.08539.94

Ham, J., & Tronick, E. (2006). Infant resilience to the stress of the still-face: Infant and maternal psychophysiology are related. *Annals New York Academy of Science, 1094*, 297-302.

Holochwost, S., Gariepy, J., Propper, C., & Mills-Koonce, W. (2014). Parenting behaviors and vagal tone at six months predict attachment disorganization at twelve months. *Developmental Psychobiology*, 56. 10.1002/dev.21221.

Kogen, N., & Carter, A.S. (1996). Mother-infant reengagement following the still-face: The role of maternal emotional availability in infant affect regulation. Infant *Behavior and Development, 19*, 359-370.

Korja, R., Nolvi, S., Grant, K., & McMahon, C. (2017). The Relations Between Maternal Prenatal Anxiety or Stress and Child's Early Negative Reactivity or Self-Regulation: A Systematic Review. *Child Psychiatry and Human Development, 48*(6), 851-869.

MacLean, P. C., Rynes, K. N., Aragón, C., Caprihan, A., Phillips, J. P., & Lowe, J. R. (2014). Mother–infant mutual eye gaze supports emotion regulation in infancy during the Still-Face paradigm. *Infant Behavior and Development, 37*(4), 512–522. https://doi.org/10.1016/j.infbeh.2014.06.008

Manian, N., & Bornstein, M. H. (2009). Dynamics of emotion regulation in infants of clinically depressed and nondepressed mothers. *Journal of Child Psychology and Psychiatry, 50*(11), 1410–1418. https://doi.org/10.1111/j.1469-7610.2009.02166.x

Mesmen, 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*, 120- 162.

Moore, G.A., Hill-Soderlund, A.L., Propper, C.B., Calkins, S.D., Mills-Koonce, W.R., Cox, M.J., (2009). Mother–infant vagal regulation in the face-to-face still-face

paradigm is moderated by maternal sensitivity. *Child Development, 80*, 209–223.

Ostlund, B. D., Vlisides-Henry, R. D., Crowell, S. E., Raby, K. L., Terrell, S., Brown, M. A., … 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

Propper, C., Moore, G.A., Mills-Koonce, W.R., Halpern, C.T., Hill-Soderlund, A.L., Calkins, S.D., Carbone, M.A., & Cox, M. (2008). Gene– Environment Contributions to the Development of Infant Vagal Reactivity: The Interaction of Dopamine and Maternal Sensitivity. *Child Development, 79*(5), 1377-1394.

Sheppes, G., Suri, G., & Gross, J. J. (2015). Emotion regulation and psychopathology. Annual review of clinical psychology, 11, 379–405. https://doi.org/10.1146/annurev-clinpsy-032814-112739

Thomas, J. C., Letourneau, N., Campbell, T. S., Tomfohr-Madsen, L., & Giesbrecht, G. F. (2017). *Developmental origins of infant emotion regulation: Mediation by temperamental negativity and moderation by maternal sensitivity. Developmental Psychology, 53*(4), 611–628. https://doi.org/10.1037/dev0000279

Thompson, R. A., & Goodman, M. (2010). Development of emotion regulation: More than meets the eye. In A. M. Kring & D. M. Sloan (Eds.), Emotion regulation and psychopathology: A transdiagnostic approach to etiology and treatment (pp. 38–58). The Guilford Press.

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 American Academy of Child Psychiatry, 17*(1), 1–13.

Warnock, F., Craig, K., Bakeman, R., Castral, T., & Mirlashari, J. (2016). The relationship of prenatal maternal depression or anxiety to maternal caregiving behavior and infant behavior self-regulation during infant heel lance: An ethological time-based study of behavior. *BMC Pregnancy and Childbirth, 16*(1), 264.

DOES LOW-DOSE THIAMINE SUPPLEMENTATION AFFECT

MOTHERS' SUPPORT FOR INFANT SECONDARY

ENGAGEMENT?

by

BRIDGET JOHNS

University of Oregon

Acknowledgements

There are so many people who have supported and guided me through my thesis that I would like to thank. Firstly, I would like to thank Dr. Dare Baldwin for all her support. It has been such an honor to work under her for the Cambodia Project; she is such an inspiration and had led myself and many, many other students through our research and theses, and always makes time for each of us. Her kindness, patience, and intelligence have been invaluable to me throughout this process. Dare has been a role model for me since I joined her lab due not only to her distinguished background in research but also through her strength and kindness in character. I am so grateful to have gotten the opportunity to work with her and contribute to her research with the Cambodia Project – I will carry the skills she taught me for a lifetime. I would also like to thank Diana DeWald, who has spent many Wednesday afternoons working with me on research as well as guiding me through my thesis. I have learned so much from her and would in no way be here without her. Diana is as strong as she is kind and driven, and to have her as my mentor, I definitely hit the jackpot. She has been there with me every step through this process, from creating the coding manual, coding the data, writing up results, and creating posters and presentations. Thank you so much for always being there for me, Diana, you are irreplaceable! I would also like to thank Dr. Ulrich Mayr for his instruction of our Honors in Psychology section this year. His direction and supervision during this school year has been such a great help and another person of support during the brainstorming, data analyzation, and writing of all of our research projects.

Finally, I would like to thank my mom and dad for their unwavering support while I have been in college, for being just as excited about my research as I am, and for their patience with me when I felt stressed or overwhelmed with the process. I would like to thank my Uncle Mike

and Aunt Tracy as well for being my rocks while I've been at school in Oregon, and for always providing a quiet and encouraging environment for me to work on my thesis when the craziness of campus got overwhelming. I would also like to thank Anna Sanchirico for helping me get involved in this research lab, and also Karabeth Schini, Joleen Winings, and Savannah Wease for all of their help coding the behavioral data for my thesis. I could not have done it without you, thank you!

Table of Contents

List of Figures

List of Tables

Abstract

Thiamine, or Vitamin B1, is a crucial nutrient for the human health and neurocognitive development, yet millions are affected by thiamine deficiency in regions such as Southeast Asia, due to cultural reliance on thiamine-poor polished white rice as a dietary staple. This thesis investigated the possibility that thiamine status may affect a key aspect of mother-infant interaction that is known to support infants' neurocognitive development, called secondary engagement. Secondary engagement is a process of shared attention to an external referent, and it serves as an important context for infants' learning across domains. I hypothesized that supplemental thiamine might benefit mothers' mothers' propensity and ability to engage with infants in shared attention regarding a novel object. The research took place as part of a larger, randomized, controlled trial investigating possible benefits of low-dose maternal thiamine supplementation for breastfed babies' neurocognitive development. To examine possible effects of thiamine on maternal joint engagement efforts, we developed a a novel task in which mothers were asked to engage infants in joint attention on a novel object through five epochs involving gradual addition, and then removal, of social cues such as gaze/facial expression/voice and gestures. We coded maternal secondary engagement efforts in this task using a five-point Likert scale indexing four dimensions: presentation of object, joint engagement efforts, contingent responding, and affective tone. As predicted, ratings on each of these dimensions displayed a quadratic trend across the five epochs of the task, indicating mothers implemented the task as instructed. However, maternal thiamine supplementation group did not significantly influence maternal secondary engagement efforts in the task, suggesting that mothers' ability to support infant secondary engagement is robust in the face of risk for thiamine deficiency.

Introduction

For infants and caregivers around the world, thiamine, or vitamin B1, is an essential micronutrient that promotes a variety of cognitive, behavioral, and health outcomes. In Southeast Asia, thiamine deficiency is all too common due to a cultural dietary reliance on polished white rice. This deficiency contributes to infant mortality via the disease Beriberi. Even when infants survive thiamine deficiency, they experience long-term negative consequences for development. For example, research by Fattal-Valevski and colleagues (2009) followed a group of Israeli infants who had been fed soy-formula inadvertently lacking thiamine. Their findings indicated that even subclinical thiamine deficiency undercut infants' neurocognitive development.

Most Western countries have thiamine incorporated into their foods, so consequences from its deficiency are rarely seen in these areas. However, there are still many regions where thiamine is not naturally found in common foods, so this issue remains a global concern. On the other hand, recent evidence indicates that supplementing lactating mothers with thiamine helps protect infants' neurocognitive development, and especially their language development (Measelle, et al. 2021; Baldwin, et al. under review). My research took place in the context of a larger, randomized, controlled trial in Cambodia, with a specific focus on investigating the extent to which maternal thiamine supplementation may benefit a precursor to language learning; specifically, mothers' propensity and ability to engage with infants in joint attention regarding a novel object.

What is joint attention, and how does it contribute to language development?

Joint attention (JA) – referred to as secondary engagement (SE) throughout the remainder of this thesis – is of particular interest as it is a major locus within which infant learning occurs,

and helps to construct infants' language development and knowledge acquisition (Baldwin, 2000; Tomasello, 1999). SE represents a dyadic process of shared attention to an external referent; it is an important context scaffolding infants' learning and social development. In the present study, SE was instantiated within a semi-structured task in which mother-infant dyads engaged with a toy that was novel to infants. Quantity and quality of early SE interactions predict many aspects of cognitive development (e.g., Smith, Adamson, & Bakeman, 1988). Some examples of this include learning about object functions and, in particular, language, but SE interactions can also be predictive of developmental disorders such as autism (Adamson, Bakeman, & Suma, 2019).

In terms of object functions, evidence indicates that SE interactions can be particularly helpful in aiding infants' learning and development. One study by Woodward found that infants respond systematically to a caretaker's gaze regarding a novel object during SE interactions, particularly when grasping for an object is involved (Woodward, 2003). Infants have been shown to understand grasping in terms of a relation between a caretaker and an object and are able to understand gaze as a link between a caretaker and an object as well. This is further developed the older an infant becomes (Woodward, 2003). These events are an indication of SE interactions further supporting an infants' development, specifically with drawing connections between gaze, caretaker, and object.

SE has been particularly well-documented in regard to language development, and the ability to participate in SE interactions between caretakers and infants appears to be a precursor to language development for infants (e.g., Baldwin, 1991). This is in part due to the activity of labeling objects, as language learning can be facilitated with a caretaker's utterance about an object if the caretaker and infant both are engaging in SE towards it (Baldwin, 1991; Tomasello

& Farrar, 1986). So, as a result of SE, infants are able to interpret labels for novel objects, and are able to further build their world of language through these SE interactions. (Baldwin, 1991). Moreover, SE is particularly useful in understanding language development in the way that it helps infants to form generalizations about how words and language are used in learning about the world as they mature (Ahktar & Tomasello, 2000).

Deficits in SE interactions have been shown to undercut infants' language development, and these same deficits in SE can be an indication of autism (Baron-Cohen, Baldwin, & Crowson, 1997). Research indicates that autism can be typified by impairments in initiating SE, and this impairment is hypothesized to be one of the earliest behavioral markers of atypical neural connectivity in autism for infants (Mundy & Newell, 2007). Thus, SE is a powerful tool in learning about infants' neurocognitive development, in regard to learning as well as potential manifestations of disorders such as autism.

Overall, there is considerable evidence indicating that SE interactions are crucial for infants' learning and development. Across multiple domains, frequency and quality of SE interactions can be useful in infant neurocognitive development, as seen in research regarding infant learning about language and objects, but also useful in terms of low-quality SE interactions serving as a potential early diagnostic of autism.

How might maternal thiamine supplementation benefit SE interactions?

Thiamine deficiency can lead to many health and cognitive issues. As thiamine is critical basic metabolic processes, individuals deficient in thiamine can become lethargic and chronically fatigued. There are also problems regarding concentration, as this deficiency can negatively

impact cognitive functions (Whitfield et al. 2018). Thiamine supplementation thus holds

potential to prevent or reverse such negative implications for of thiamine deficiency.

Thiamine deficiency, for which both mothers and infants in Southeast Asia are at risk,

may inhibit energy and skill-intensive SE interactions. For mothers in particular, a deficiency in

thiamine might plausibly undercut their ability to create and sustain SE interactions with infants.

In infants' first year of life, thiamine supplementation for lactating mothers may be useful in

promoting effective SE interactions for mothers, but is useful as well as for breastfeeding infants

who are at risk of developing thiamine deficiency. While my research takes place in the context

of a larger, randomized, controlled trial investigating the possible benefits of low-dose maternal

thiamine supplementation for breastfed babies' neurocognitive development, my research

focused in particular on mothers' ability to support infants in secondary engagement; specifically,

mothers' propensity and ability to engage with infants in joint attention regarding a novel object.

This research included neurocognitive assessments as well as video recordings of mothers and

infants engaging in a novel, semi-structured SE task.

Broadening information about SE across cultures

This study is also of importance because, while SE has been studied in Western cultures

such as the United States, there is relatively little research examining mother-infant SE

interactions in non-Western populations. The literature pertaining to SE in Western cultures

indicates that it is indeed a powerful tool in infant learning and development (Barton &

Tomasello, 1991). Some literature indicates that when race is taken into account in a Western

sample, there is no difference between races in how well an SE task is completed (Saxon &

Reilly, 2006). However, the dimensions investigated with SE may be different cross-culturally,

so conducting research on this task with a wider array of non-Western cultures is of the utmost importance to gain a better understanding of how it can impact neurocognitive development. Secondary engagement is also of interest as a potential clinical assay of the quality of mother-infant interaction *vis à vis* infants' neurocognitive development.

SE interactions have proven to be incredibly useful in examining an infants' progress in their neurocognitive development; however, these interactions are rarely included in clinical assessments of infants' neurocognitive development, and we are not aware of any research that includes SE as an assessment in the context of investigating possible benefits of micro-nutrient supplementation. But, it would make sense to include SE assessments when examining infants' neurocognitive development for a myriad of reasons; it is a precursor to language learning, a potential indication of developmental disorders such as autism, a useful way for infants to learn about objects, and more. If SE tasks are included in clinical assessments of infants' neurocognitive development, it is possible that we may better understand just how well an infant is learning and developing in their first year of life.

We developed a new SE task designed for this purpose as well as a novel maternal coding system to assess mother's SE behaviors within the task. Caregivers were told the overarching goal of the task was to establish and sustain infants' interest in a novel object over the course of the interaction. As the task progressed across five 30-second "epochs," mothers were prompted to add and then subsequently remove cues to secondary engagement. Videos of these interactions at two timepoints – 12 and 24 weeks postnatal – were coded via five-point Likert scales index four dimensions within mothers' behavior: presentation of object (PO), joint engagement efforts (JEE), contingent responding (CR), and affective tone (AT).

We hypothesized that maternal thiamine supplementation would enhance mothers' efforts to support secondary engagement in interaction with their infants. We had two specific predictions: 1) ratings on each of the four dimensions would display a quadratic trend cross the five epochs of the secondary engagement task, and 2) higher levels of thiamine supplementation be related to higher scores on all four dimensions of maternal secondary engagement efforts.

Methods

Participants

The sample was comprised of 335 mother-infant dyads from the Kampong Thom province in Cambodia. Dyads were recruited through antenatal care visits at eight different health centers around the Kampong Thom province, and joined the study two weeks after giving birth. Mothers ranged in age from 18 to 45 years old (*M* = 28.1, *SD* = 6.2; See Table 1), and had healthy, exclusively breastfed, newborns. Most of the mothers (69%) had more than one child and the vast majority were married (98%). Roughly half of the infants were female (48%). Written informed consent was obtained from mothers for both themselves and their infants prior to the study. The full study attrition rate averaged across the four thiamine supplementation groups was 9.05% over the course of data collection (see Figure 1). Additionally, individual dyad data were excluded from our partial dataset due to 1) infant fussiness that interrupted the flow of the secondary engagement task, 2) videos ending prior to completion of the task, and 3) no clear task structure observed in a given video. The partial dataset reported here comprised 159 dyads at 12 weeks, and 139 dyads at 24 weeks postnatal.

Table 1: Infant and Maternal Characteristics by Treatment Group at Baseline

	TOTAL *N*=335	0 mg *n*=83	1.2 mg *n*=86	2.4 mg *n*=81	10 mg *n*=85
Infant					
Sex, *female*	161 (48%)	43 (52%)	43 (50%)	33 (41%)	42 (49%)
Length-for-age (Z-score) at 2 wks	-.62 (1.02)	-0.52 (0.98)	-0.66 (1.11)	-0.69(1.01)	-0.63 (1.01)
Mother					
Age, *years*	28.1 (6.2)	28.3 (6.1)	27.9 (6.7)	28.1 (6.1)	28.1 (5.9)
Parity, *multiparous*	57 (69%)	54 (65%)	54 (63%)	58 (72%)	64 (75%)
Ethnicity, *Khmer*	335 (100%)	83 (100%)	86 (100%)	81 (100%)	85 (100%)
Marital status					
Married	330 (98%)	79 (95%)	86 (100%)	81 (100%)	84 (99%)
Divorced/Separated/Widowed	5 (<1%)	4 (5%)	0 (0%)	0 (0%)	1 (1%)
Milk total thiamine concentrations (µg/L) at 2 wks	129.1 (74.4)	135.5 (77.7)	129.3 (71.4)	126.3 (77.3)	125.4 (72.3)
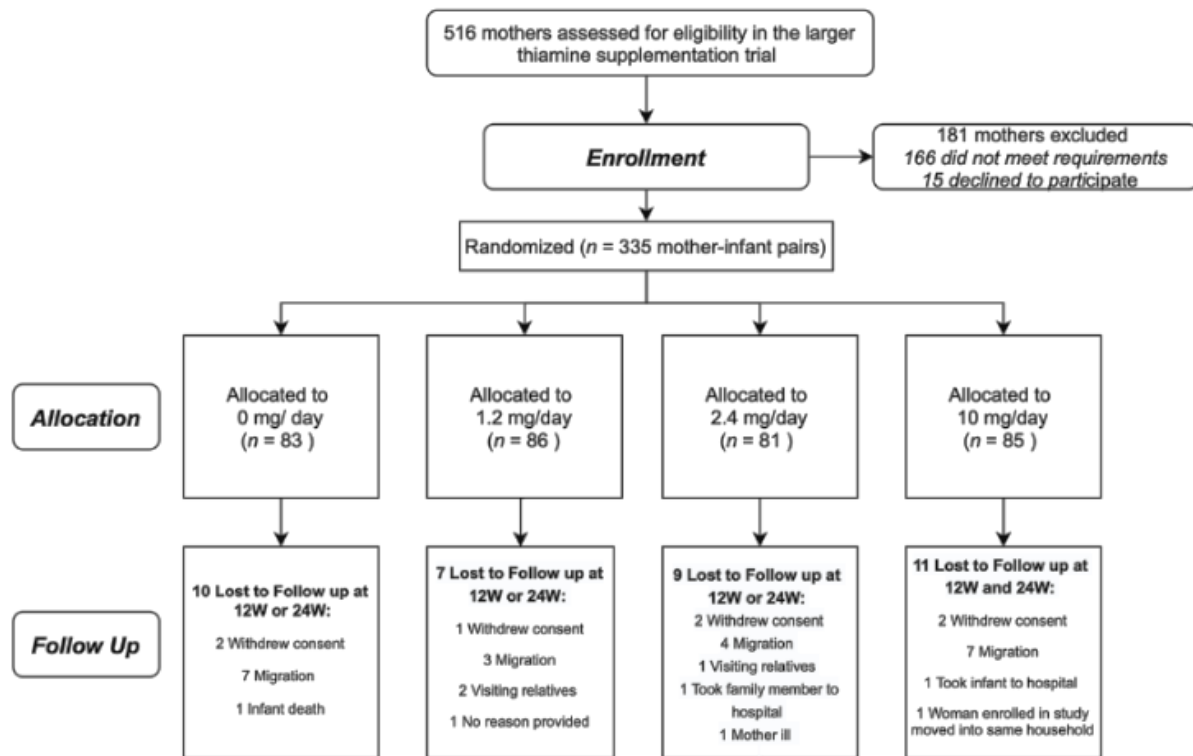
Figure 1. Participant attrition within the full dataset

Procedure

The study was a double-blind, four-parallel-arm randomized, repeated measures controlled trial, with data collection taking place between September 2018 and December 2019. The caregivers participating in the study were not involved in any extra nutrition programs and had not consumed thiamine supplements in the four months before their babies were born. Two weeks postpartum, caregivers were randomly selected for one of four thiamine supplementation groups and given fortified salt tablets. They were either in the placebo dose group, receiving 0 mg of thiamine, the estimated average requirement (EAR) at 1.2 mg of thiamine per dose, double the EAR receiving 2.4 mg of thiamine, or the positive control group of 10 mg of thiamine per

dose. Mothers were asked to take one capsule, every day, between 2 and 24 weeks postpartum. Their thiamine intake was recorded via weekly check-ins by local Khmer-speaking researchers.

Caregivers also participated in a secondary engagement task (SET) when infants were 12 and 24 weeks old, in which they were told the overarching goal of the task was to establish and sustain infants' interest in a novel object over the course of the interaction. The task was video-recorded. Mothers were given task instructions by Khmer-speaking researchers, and asked to bring their babies into a calm, focused state prior to starting the task (i.e. infant feeding and changing needs were addressed prior to the SET, and the infant was not sleepy). As the task progressed across five 30-second "epochs," mothers were prompted to add and then subsequently remove cues to secondary engagement (e.g., line-of-regard, voice and gesture, see Figure 2). Specifically in epoch 1, caregivers were asked to engage their infant in looking at a toy, while they themselves did not look at the infant or toy. In epoch 2, they were asked to remain silent but turn towards the infant and toy in visually-directed joint attention. In epoch 3, the caregiver was asked to add voice and gesture to support joint attention. In epoch 4, similar to epoch 2, caregivers removed vocal cues but attended to the toy visually with their infant. Finally, in epoch 5, as epoch 1, caregivers silently turned away from the infant and toy, while the infant continued interacting with the toy. The rationale for having mothers first add, then remove social cues to secondary engagement over the course of the SET was to provide multiple opportunities to assess how effectively mothers utilized various "tools" (such as gaze, facial expression, voice, and gesture) to engage and sustain secondary engagement with infants, as well as to assess (in research undertaken by others that was beyond the scope of this thesis) infants' responsiveness to

various cues to secondary engagement that mothers provided.

PHASES	Epoch 1	Epoch 2	Epoch 3	Epoch 4	Epoch 5
	30 sec	30 sec	30 sec	30 sec	30 sec
SECONDARY ENGAGEMENT BEHAVIORS	Infant views novel object; caregiver looking away from both infant and object	Silent caregiver turns head and eyes to jointly attend to novel object visually	Caregiver talks and gestures about novel object during joint visual attention	Caregiver ceases vocal talk and continues to jointly attend to novel object visually	Silent caregiver turns visual attention away from infant and novel object; infant may continue to view object
Mother engages in both visually-and-verbally-directed joint attention with infant and toy	X	X	✓	X	X
Mother looks at infant and toy in visually-directed joint attention	X	✓	✓	✓	X

Figure 2. Structure of the Secondary Engagement task

Measures

Maternal Thiamine Status

Several thiamine biomarkers were recorded for mothers at 2 and 24 weeks postnatal to assess their total thiamine diphosphate concentration (ThDP). First, blood samples were assayed to detect erythrocyte transketolase (ETKac) concentration–an enzyme essential for metabolizing thiamine, with lower ETKac indicating greater thiamine status. Second, breast milk samples were collected and analyzed via E-max dose response curves (Gallant et al., 2020; Whitfield et al., 2019). Milk total thiamine concentrations are reported by maternal treatment group in Table 1. While confirmatory hypotheses were used to assess the relation between treatment group and maternal joint attention, analysis of biomarker thiamine data described was beyond the scope of this thesis.

Maternal SES and Self-reported Well-being

Self-reported well-being was assessed via four survey questions at 3 timepoints: 2 weeks, 12 weeks, and 24 weeks postnatal. Mothers reported the quality of their sleep and their infants'

sleep on a 5-point Likert scale. Additionally, a previously validated (e.g. Gjerdingen et al., 2009) two-item Post Health Questionnaire (PHQ-2) was translated from English for the Khmer-speaking sample to identify postpartum depression. Together, the sleep and PHQ-2 data were standardized to create a 'maternal well-being composite' discussed at greater length in the results section under "exploratory analyses." Further, a wealth equity index was created based upon self-reported survey questions regarding education level, income, and household size, plus researcher observations regarding home and toilet condition, fuel sources, transportation, etc. (See Table 2 for details).

Table 2: Maternal Education and Household/SES Information

	TOTAL *N*=335	0 mg *n*=83	1.2 mg *n*=86	2.4 mg *n*=81	10 mg *n*=85
Education					
None	40 (12%)	10 (12%)	8 (9%)	13 (16%)	9 (11%)
Primary (1-6 years)	161 (48%)	43 (52%)	37 (43%)	40 (49%)	41 (48%)
Lower Secondary (7-9 years)	83 (25%)	16 (19%)	29 (34%)	19 (24%)	19 (22%)
Upper Secondary (10-12 years)	43 (13%)	12 (15%)	9 (11%)	8 (10%)	14 (17%)
Higher education	8 (2%)	2 (2%)	3 (3%)	1 (1%)	2 (2%)
Milk total thiamine concentrations (µg/L) at 2 wks	129.1 (74.4)	135.5 (77.7)	129.3 (71.4)	126.3 (77.3)	125.4 (72.3)
Household					
Husband education					
None	38 (11%)	10 (12%)	9 (10%)	9 (11%)	10 (12%)
Primary (1-6 years)	151 (45%)	42 (51%)	37 (43%)	39 (48%)	33 (39%)
Lower Secondary (7-9 years)	97 (29%)	21 (25%)	24 (28%)	23 (28%)	29 (34%)
Upper Secondary (10-12 years)	34 (10%)	5 (6%)	13 (15%)	8 (10%)	8 (9%)
Higher education	15 (4%)	5 (6%)	3 (3%)	2 (3%)	5 (6%)
Household size, *number of people*	3.9 (1.9)	3.7 (1.7)	3.6 (1.8)	4.0 (2.1)	4.1 (2.0)
Median Annual household income, *US$ (IQR)*	1620 (950-3500)	1800 (950-3000)	2050 (963-3500)	1600 (1000-3000)	2000 (1200-3500)
Wealth Index Score*					
Poorest	81 (24%)	22 (27%)	12 (15%)	21 (26%)	25 (29%)
Second	69 (21%)	16 (19%)	14 (16%)	20 (25%)	19 (22%)
Middle	108 (32%)	26 (31%)	31 (36%)	24 (30%)	27 (32%)
Fourth	54 (16%)	14 (17%)	20 (23%)	11 (13%)	9 (11%)
Wealthiest	23 (7%)	5 (6%)	8 (10%)	5 (6%)	5 (6%)

Data are mean (SD) or n (%), except household income, shown as median (IQR). Percentages may not add to 100% due to rounding.
* Wealth equity index (WEI) quintiles calculated based on the Demographic Health Survey Program guidelines (USAID); Cambodian WEI developed using 2014 DHS data.

Secondary Engagement Task—Maternal Behavior

To validate the SET task, we created a maternal behavior code based upon previous work by Cox, Crnic, Mills-Koonce, Fiori-Cowley, Murray, and Gunning (2003) to assess mothers' secondary engagement interactions within the task. A team of 5 coders who were blind to maternal treatment group rated mothers' behavior in recorded videos via Likert scales ranging from 1 (behavior not exhibited) to 5 (behavior strongly present throughout 30-second epoch) on four dimensions: affective tone, contingent responding, joint engagement, and object presentation. Coders recorded both a "main" maternal score per dimension per epoch, meaning that for the majority of the epoch, the caregiver exhibited a certain score for more than half of the time, and a "high" score, indicating the highest possible score a caregiver exhibited during an epoch. If a caregiver briefly presented what would be an outstanding score, but did not consistently display joint attention efforts for the majority of the epoch, a lower score would be given for the "main" and a higher one would be given for the "high" score. Additional details regarding the maternal coding system can be found in Appendix A.

Results

The overarching goal of this project was to investigate the extent to which maternal thiamine supplementation affects mothers' ability to support secondary engagement with infants at 12 and 24 weeks postnatal. Video recordings of mothers and babies interacting together in the secondary engagement task were coded for four dimensions of maternal support for secondary engagement: presentation of object (PO), joint engagement efforts (JEE), contingent responding (CR), and affective tone (AT). We approached analysis of these data by means of a mixed design MANOVA (treatment group X timepoint X epoch) which included all four dimensions of maternal behavior. From this analysis, we first report findings related to mothers' support for secondary engagement, including the effect of timepoint (12 versus 24 weeks), with an eye to discovering the degree to which we saw patterns predicted for the secondary engagement task. We followed that by reporting findings related to the extent to which thiamine influenced these patterns.

Validation of Secondary Engagement Task

Firstly, we predicted that maternal support for secondary engagement would display a quadratic pattern of increasing and decreasing support over the course of the secondary engagement task. We also examined whether maternal support for secondary engagement would display a linear decrease over the course of the task. The MANOVA provided a test of these predictions by means of the main effect of epoch and also by polynomial contrasts on the epoch variable. We expected maternal secondary engagement skills to display a quadratic trend across the five task epochs. The five (epoch) X two (timepoint: 12 vs. 24 weeks) mixed-design multivariate ANOVA including all four maternal SE variables revealed a significant main effect

of epoch, $F(16, 108) = 75.848$, $p = .000$, partial eta-squared $= .918$. Confirming our prediction, all four variables displayed a significant quadratic trend across epochs (univariate F's > 180.967, p's $< .000$) (Figure 3). This finding provided strong confirmation that, overall, mothers implemented the SET as they'd been instructed. Regarding linear trends that might occur in addition to the quadratic trends, joint engagement efforts (JEE) (but not the three other maternal behavior dimensions) revealed a significant linear trend across epochs, $F(1, 123) = 4.452$, $p = .037$, partial eta squared $= .035$, where mothers displayed lower rates of joint engagement effort in epoch relative to epoch 1. These findings indicate that, on the one hand, mothers' joint engagement efforts showed an overall decline over the course of the SET, whereas in contrast, mothers sustained their other secondary engagement efforts at a level comparable to what they had displayed at the outset of the SET.
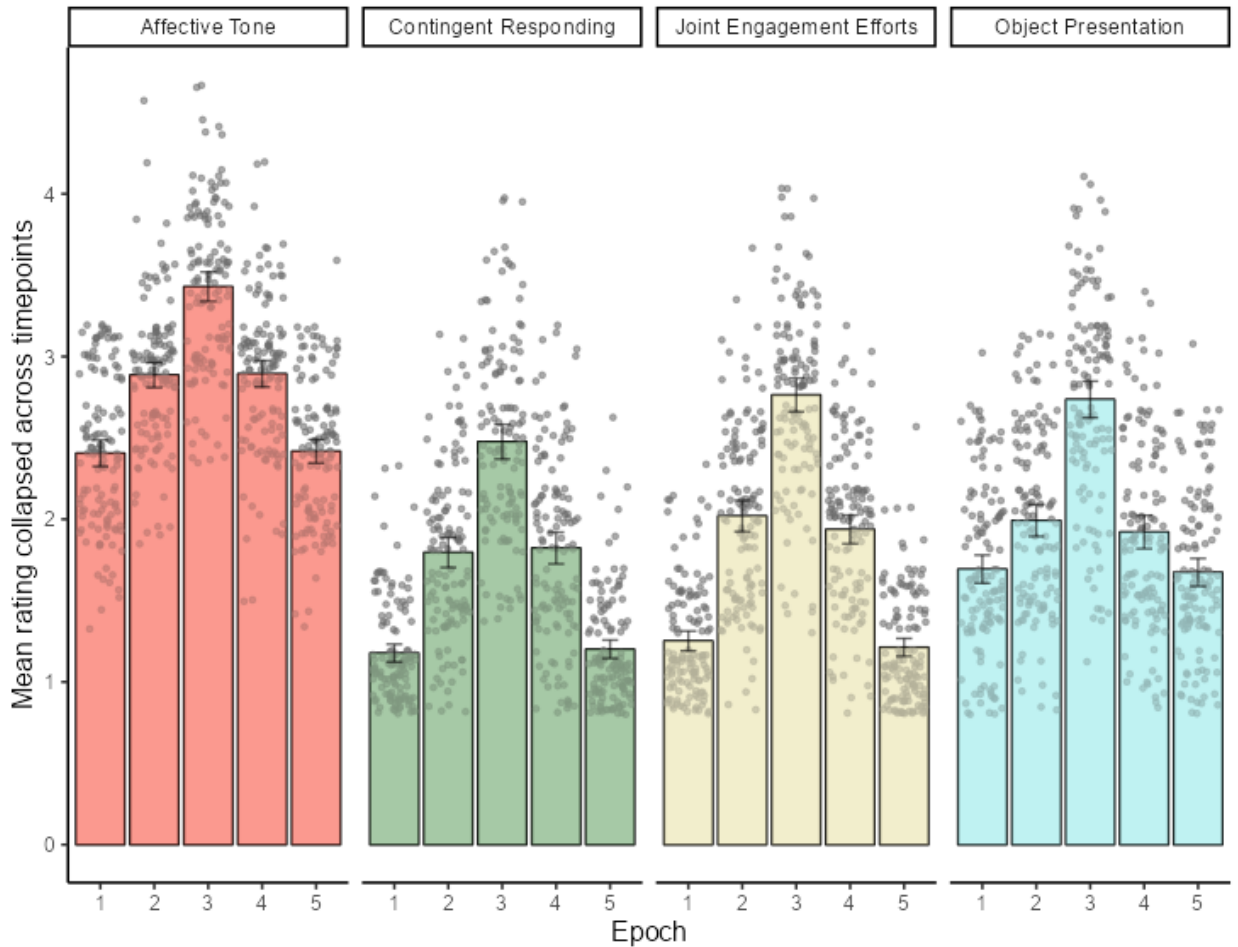
Figure 3. Mean rating of maternal secondary engagement behaviors on four dimensions across the five 30-second epochs of the SE task (collapsed across 12- and 24-week timepoints). Error bars represent +/- 1 standard error of the mean.

Timepoint Differences in the Secondary Engagement Task Validation

This analysis also revealed a significant main effect of timepoint, $F(1, 123) = 12.458$, $p = .001$, partial eta squared = .092, with higher overall mean levels for three maternal codes: object presentation, joint engagement efforts, and affective tone (but not contingent responding) when infants were 12 relative to 24 weeks (Figure 4). This may suggest that mothers scaled down joint attention efforts over time to accommodate infants' increasing joint attention abilities. No significant epoch X timepoint interaction emerged in the analysis.
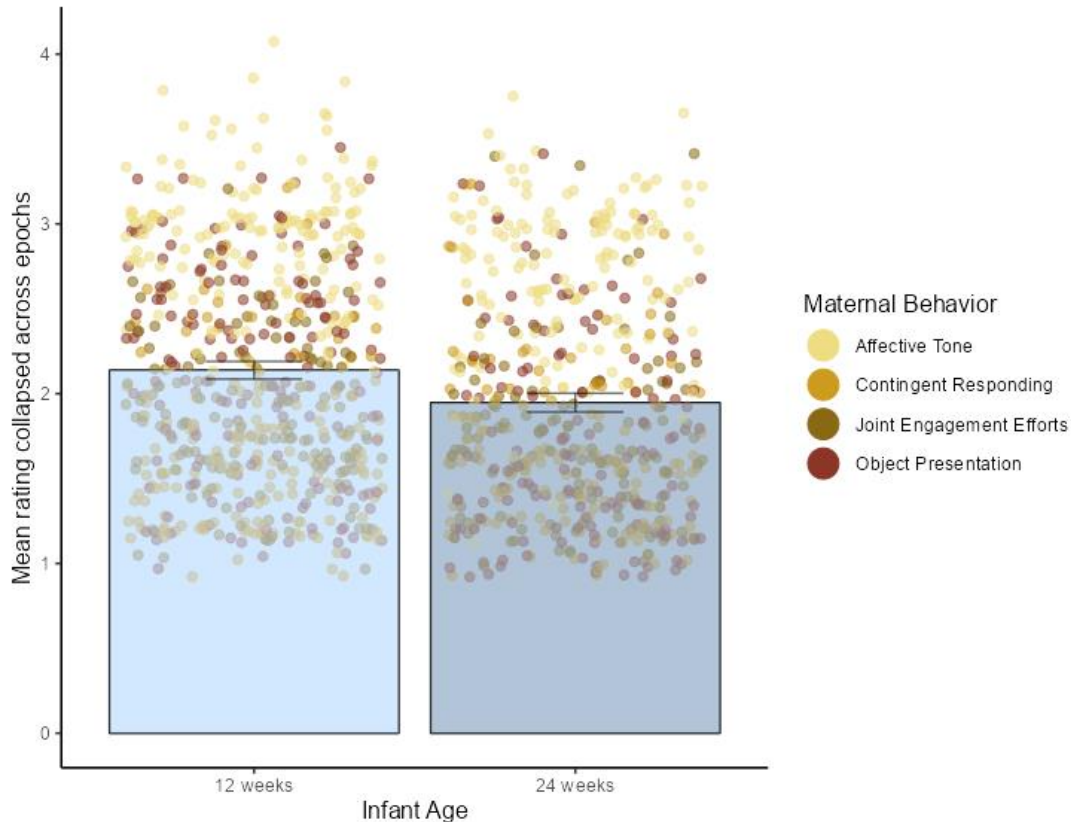
Figure 4. Mean rating of maternal secondary engagement behaviors across two timepoints of infant age and collapsed across five 30-second epochs of the SE task (behavior dimensions distinguished by point color). Error bars represent +/- 1 standard error of the mean.

Possible Effects of Maternal Thiamine Supplementation in Supporting Secondary Engagement

In addition to task validation, we predicted a main effect of maternal thiamine supplementation group, specifically with higher levels of thiamine supplementation being associated with higher levels of maternal support for secondary engagement. We also predicted that thiamine supplementation would be associated with a more pronounced quadratic pattern in mothers' support for secondary engagement across epochs. In the MANOVA, this latter prediction would be tested by an interaction between the maternal thiamine supplementation group and the polynomial contrasts for a quadratic trend.

Against prediction, this analysis revealed no significant main effect of maternal thiamine supplementation dose on maternal behavior, $F(3, 123) = .550$, $p = .649$, partial eta-squared $= .0013$ (Figures 5 & 6 display this result by timepoint). Both the maternal thiamine supplementation X epoch interaction and the timepoint X maternal thiamine supplementation interaction were non-significant [$F(12, 366) = .752$, $p = .700$, partial eta-squared $= .024$; $F(3,123) = 1.236$, $p = .300$, partial eta-squared $= .029$, respectively]. Lastly, there was also no significant three-way interaction between epoch, maternal thiamine supplementation, and timepoint, $F(12, 366) = .876$, $p = .572$, partial eta-squared $= .028$.
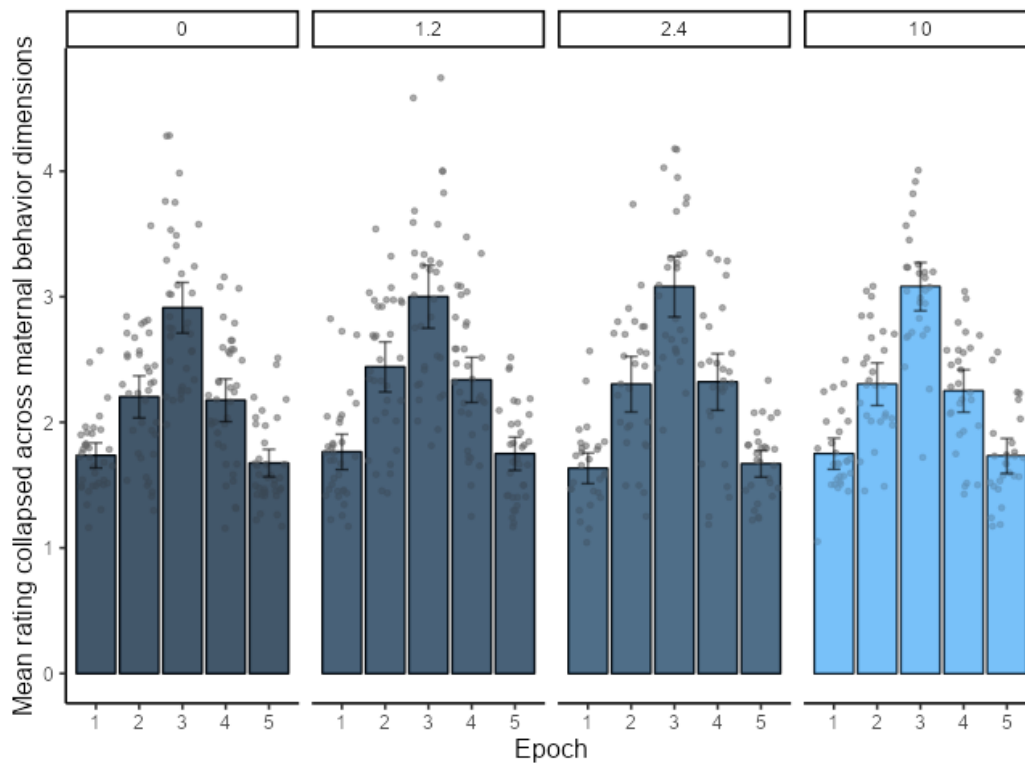
Figure 5. Mean rating of maternal secondary engagement behaviors for the **12-week timepoint** collapsed across the SE task epochs compared in relation to maternal thiamine supplementation group. Error bars represent +/- 1 standard error of the mean.
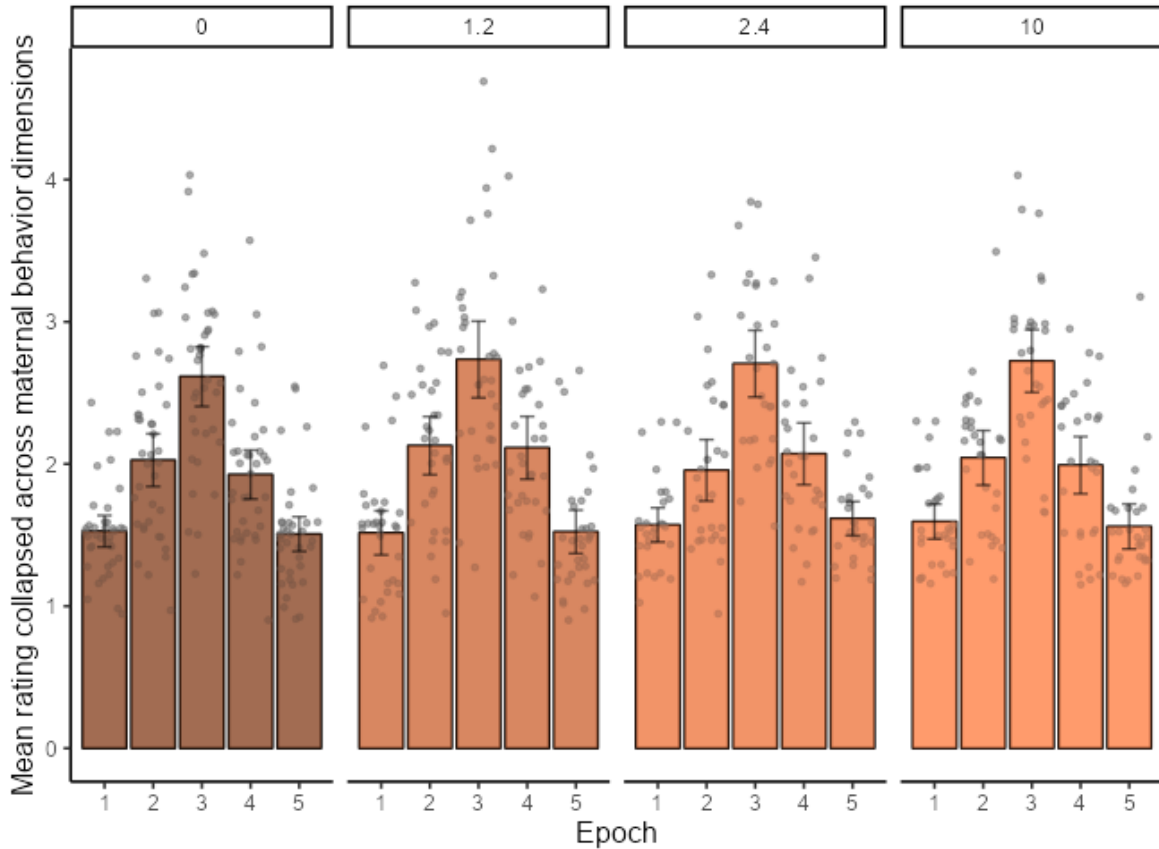
Figure 6. Mean rating of maternal secondary engagement behaviors for the **24-week timepoint** collapsed across the SE task epochs in relation to maternal thiamine supplementation group. Error bars represent +/- 1 standard error of the mean.

Effect of Maternal Thiamine Supplementation with binary grouping

We ran an additional MANOVA (timepoint X epoch X treatment category) to assess the effect of *any* maternal thiamine supplementation, treating thiamine treatment group as a binary variable (i.e. placebo (0mg/dose) versus all other dose groups (1.2, 2.4, 10mg/dose). This again revealed no significant main effect of maternal thiamine treatment group on mothers' secondary engagement efforts, $F(4, 121) = .848$, $p = 0.490$, partial eta squared = .027 (see Figure 7), and no interaction effects.
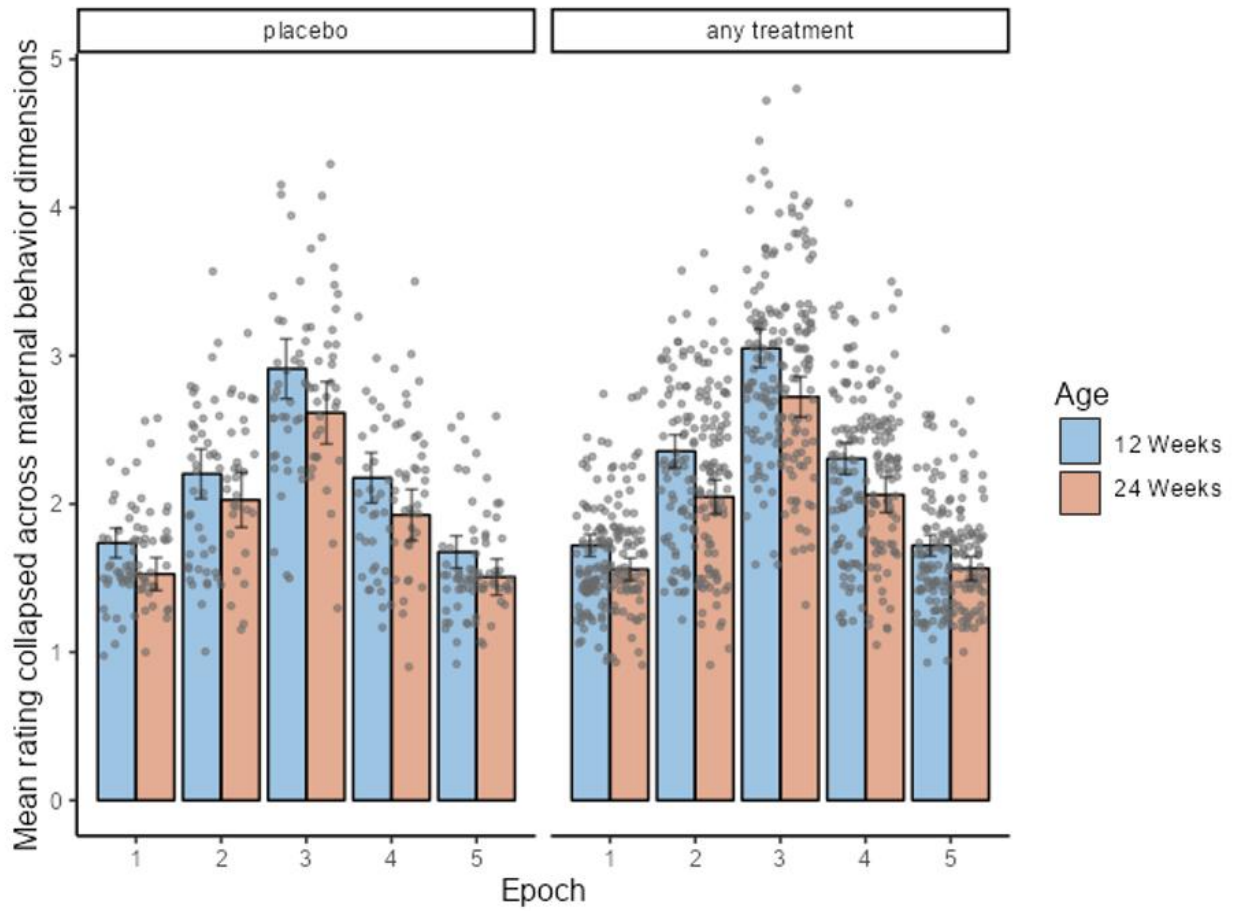
Figure 7. Mean rating of maternal secondary engagement (collapsed across four maternal behavior dimensions) shown by age group, task epoch and treatment category (i.e., placebo versus thiamine supplementation). Error bars represent +/- 1 standard error of the mean.

Discussion

The purpose of this study was to examine the extent to which maternal thiamine supplementation influenced mothers' support for secondary engagement with their infants, both at 12 and 24 weeks postnatal. We also aimed to determine whether maternal support for secondary engagement would display a quadratic trend of increasing and decreasing over the support of the task, as we had predicted. Indeed, this predicted quadratic trend emerged clearly across all measured secondary engagement dimensions, indicating that mothers were able to perform the task as instructed. This provides an important source of validation that the SET can be implemented in a rural field setting as conceptualized. Other analyses revealed that, contrary to prediction, maternal thiamine supplementation exerted no systematic influence mothers' secondary engagement efforts in the SET. We predicted that it might, given the important role that thiamine plays in metabolic activity combined with secondary engagement requiring significant investment of psychological energy on mothers' part, but it did not. Lastly, an interesting and unexpected finding arose in relation to the timepoint at which infants participated in the SET with their mothers: mothers displayed higher levels of secondary engagement effort across all dimensions when interacting with their 12-week infants, relative to when infants were 24 weeks old. One possible interpretation of this finding is that infants' growing ability to actively participate in secondary engagement interactions between 12 and 24 weeks left mothers able to scale back their need to effortfully support such engagement at the older timepoint. At the very least, these findings point to Cambodian mothers systematically adapting their secondary engagement efforts in relation to their infants' age.

Broader Implications

There are a few implications to consider in regard to these results. First is the validation of the SE task. The present findings indicate that the task holds potential as an important tool in assessing infants' neurocognitive and social development in the future. SE abilities on an infants' part are exceptionally influential in infant neurocognitive development; it is useful for tracking neurocognitive progress across many different dimensions such as language learning, social skills, learning about emotions, and can be utilized in tracking the onset of developmental disorders such as autism. Thus validation of the SET in the present findings sets the stage for valuable opportunities to use this task in the future in a variety of other possible field settings.

The second implication to consider is that of the lack of thiamine influence on maternal support for secondary engagement. On the one hand, it is important to recognize that the supplementation mothers received in the study was only at low doses of thiamine, so perhaps higher doses may have had an effect. Additionally, mothers in this study were not supplemented with thiamine until after their babies' births. They also only received six months of thiamine supplementation total. It is at least possible that, if supplementation had started earlier, and with higher doses, there might have been an observed effect of thiamine on maternal support for secondary engagement. However, if future research confirms the present finding – that thiamine supplementation indeed has no effect on maternal support for SE interactions regardless of timeframe or dosage level, then our results are potentially encouraging. That is, such a result would imply that mothers' efforts to support secondary engagement with their infants are relatively robust, even in contexts where mothers are at risk for thiamine deficiency.

Limitations

As with any study, this thesis was presented with a variety of challenges and limitations that are crucial to consider when viewing its outcomes. One limitation we encountered was that generalization of the SE task is not possible yet, because it is currently confined to this relatively small, Cambodian-specific sample. We are planning to apply this task in a North American context as well, but currently, this task has only been undertaken within our study's sample. It is also important to consider the context in which the instructions (and research itself) for this task were implemented. There are a few different dimensions to consider here. Firstly the SE task is not fully naturalistic. It involves coaching in order to execute the step-up step-down process. This begs the question of whether the handling of the task or instructions in this way disrupted the natural way mothers might engage in secondary engagement. That is to say, there is some speculation on how mothers would perform in SE interactions in a more natural, unstructured environment. Since the task was implemented in such a specific manner, there is also the question of whether or not the task taxed the infants. Moreover, we do not fully know the extent to which the task instructions disrupted Cambodian mothers given cultural propensities.

There is also the question of statistical power in regard to our research. In particular, it is unclear whether the sample size was indeed large enough to reliably detect possible effects of thiamine on mothers' secondary engagement efforts.

Lastly, it is important to note that this thesis reports on only a preliminary subsample of the set of mothers and infants' who participated in the SET task within the larger randomized, controlled trial. Thus all findings must be regarded as preliminary, and interpreted with caution.

Future Directions

One clear future direction of this study is implementing the SE task in a western, higher income population, within a more controlled setting. In doing so, the cross-cultural validity of this task can be determined. There is also the question of thiamine dosage level. Another direction to take after this study would be to adjust thiamine dosage level to be higher than what we utilized, as it is possible the level of thiamine dosage was too low to have a systematic effect on mothers' secondary engagement efforts (or on their behavior more generally). And, thiamine dosage duration could be examined if replicating this study. Perhaps mothers could receive higher doses for longer periods of time when assessing their ability to support infants in secondary engagement. Furthermore, it may be useful for future researchers to follow infants for longer than 24 weeks as their SE abilities blossom. Past research indicates that infants do not usually have awareness of SE until roughly 9-12 months of age. We chose to assess at 12 and 24 weeks postnatal in order to focus on the mothers' abilities to support SE interactions. Many mothers do not attempt to engage in SE until their infant seems to have some awareness of it, and our research was trying to catch these interactions as well as how mothers differ in SE before infants fully develop these skills. Ideally, we would like to follow mothers through this time as infants' skills for SE further develop and blossom; future researchers could continue to assess these interactions for a longer span of time to examine this more closely. It should be noted, though, that a follow-up assessment was done in the context of the larger study in which my thesis took place. This follow-up took place at 52 weeks, and mothers at this time were no longer being supplemented with thiamine. Many tasks from our overarching study were redone with our sample in which mothers and infants were observed playing with toys in an unstructured manner. Although the SE task was not specifically redone at this time, it would be a point of interest to

observe how mothers perform in the SE task after thiamine supplementation had ended, which

future researchers could investigate if this study is to be replicated in the future.

Conclusion

The goal of my research was to determine if our novel secondary engagement task would be validated within our sample – that is to say, we aimed to determine if the codes for mothers' secondary engagement efforts would correspond to the structure of the task as we asked mothers to implement it. Our findings clearly indicated that mothers were able to implement the task as instructed, providing initial validation of the SET as a task for assessing mother-infant secondary engagement. Our second question concerned the extent to which higher doses of thiamine might enhance mothers' secondary engagement efforts. According to our analyses, thiamine supplementation did not have a significant effect on mothers' secondary engagement efforts in the SET. If these results are to be taken at face value, they suggest that Cambodian mothers' skills and propensities for engaging jointly with their infants are robust even in the face of collective risk for thiamine deficiency.

Appendix A

Maternal Secondary Engagement Coding Manual

General Decision Principles for Coders:

1. If something unusual or noteworthy occurs within an epoch, please describe this in the 'general comments' column (e.g, mother fails to follow instructions, baby fusses throughout, lighting too dark to code effectively, video provides excellent example of maternal SE, etc.). Please also provide the time-code for when the unusual, comment-worthy issue or event arose, highlight video title column in the coding assignment spreadsheet in green if you have a question on video, highlight in red if the video is noteworthy/good example of SE.
2. If the behaviors displayed in the video are ambiguous (for example, caregiver is facing away from camera), take a conservative stance and adopt the lower maternal code that is consistent with the behaviors observed. Be sure to mention in the comments that it was ambiguous, and what the other code you considered was, and why.
3. Please note the highest and main maternal codes for each dimension (i.e. the highest observed code and the code that persisted for the greater part of an epoch). Note that these codes may sometimes be the same.
4. If there is disruption during the epoch lasting at least 15 seconds (e.g. mother answers a phone call, baby feeds or defecates), or ambiguity affecting your code, please place asterisks in the main and high columns for each dimension, and write an explanation in the 'general comments' column. Additionally, highlight the video title column (column "A) based upon the key provided in the Coding Assignments document.

Epoch Time Segmentation Guidelines

Segmentation Procedures:

Marking Epoch Onset: An epoch begins when mother displays that she has begun to initiate a change in her behavior in her response to the new instructions, i.e. the very first sign that she is changing behavior in response to what she's been instructed to do during a new epoch; this may occur before the tester has completed verbalizing instructions. Record the minute and second of onset in the "Epoch ___ time" columns.

    a. *Note: If you are having trouble identifying the beginning of the task (especially if the mother is interacting with the infant during the baseline period), it may be easier to identify the onset of Epoch II and then work backwards from there. The onset of the first epoch should be roughly 30 seconds before the onset of the second.*

- *Epoch Length:* While these videos will have some error due to the nature of human researchers, each epoch should be roughly 30 seconds in length, for an overall task time of roughly 3 minutes. If, after coding, you find that any epochs are less than 25

seconds or more than 35 seconds, it is important to review those epochs. Sometimes an epoch will be unavoidably too long or too short, but this is somewhat rare. Epochs are usually close to 30 seconds in length.

Maternal Secondary Engagement → 4 Dimensions rated on a 5pt. scale

Note: we have 4 types of general behavior we are coding for, but within those larger dimensions, there are multiple behavioral dimensions related to each behavior type. For example, mothers can exhibit multiple 'presentation' behaviors for the "Presentation of Object" dimension, such as shaking a toy, talking with it or about it, rolling it along the ground, etc.

1. **Presentation of Object (PO)**

Presenting object to infant in a way to engage his/her interest in it (e.g., holding it where infants can see it, squeezing it, shaking it, pointing to it, presenting/giving it to infants, etc) (rate on 1-5 scale of effectiveness + effort)

1—No sign or almost no sign of engaging object presentation (for example: holds the object extremely still, does not hold the object or does not gesture with it for the greater part of an epoch).

2—Minor, inconsistent, and/or largely ineffective signs of engaging object presentation (for example: holds the object with little movement, holds or gestures with the object in a way that doesn't capture infant's attention (for example, out of the infant's line of vision). Can also apply when PO is moderately effective but there is little sign of engaging object presentation.

3—Moderately sized, fairly consistent, and/or moderately effective signs of engaging object presentation (for example: holds or gestures with the object in an interesting way but does not demonstrate more than one "presentation" behavior (squeezing, shaking, talking about object, etc.)).

4—In general, there are clear, consistent, and/or effective signs of engaging object presentation (for example: intentionally presents object to catch an infant's interest and demonstrates multiple "presentation" behaviors but does not necessarily make a concerted effort to keep infant's interest on the object throughout an epoch).

5—Strong, clear, highly consistent, and/or highly effective signs of engaging object presentation (for example: makes the object "come alive" for the infant and consistently attempts to direct infant's attention to target object for the majority of epoch using a variety of "presentation" behaviors).

NOTE: In 24 weeks, there may be less presentation of object because 6 month infants are more able to hold object than 3 month infants; but a caregiver could receive any level of this dimension regardless of whether or not they were holding the toy or the infant was holding the toy (for example, infant is holding toy but caregiver holds it with them or taps it--this might

result in a higher code. If infant is holding the toy and caregiver remains passive throughout epoch--this might result in a lower code).

2. Joint Engagement Efforts (JEE)

Describes efforts to bring infant into joint attention (e.g., holding it in line of sight between infant and caregiver, gaining eye contact with infant before directing their attention to the object, looking back and forth between infants' eyes/face and object, act on object and then offer infant the chance to act on object; manipulate infants' hand so as to act on object at the same time, (rate on 1-5 scale of such efforts)

1—No sign or almost no sign of efforts to bring infant into joint attention (for example: although they may be holding the object, caregiver exhibits no efforts to interact with the object jointly in a given epoch. Caregiver may also appear distracted or uninterested in the task).

2—Minor, inconsistent, and/or largely ineffective efforts to bring infant into joint attention (for example: caregiver mostly engages with infant in a primary way, but shows one or two attempts at joint engagement with the object).

3—Moderately clear, fairly consistent, and/or moderately effective efforts to bring infant into joint attention (for example: makes multiple attempts at engaging jointly (e.g. holding the object in the line of sight between themselves and infant, acting on object then giving infant a chance to do so), but only does so for about half of an epoch).

4—In general, there are clear, consistent, and/or effective efforts to bring infant into joint attention (for example: caregiver demonstrates the ability to elicit joint attention consistently throughout an epoch by talking at an object and pointing to it, looking back and forth between object and infant, etc. but at times interacts only with the object or only with the infant).

5—Strong, clear, highly consistent, and/or highly effective efforts to bring infant into joint attention (for example: throughout an epoch, caregiver scaffolds joint attention skillfully using the behaviors in the description above (or others)).

NOTE: this dimension focuses specifically on caregiver attempts to engage infant in secondary engagement (e.g. sharing joint attention with respect to the novel object) rather than to engage with the infant in a primary sense (e.g. sharing direct, one-to-one exchange of attention, emotion, or language). In other words, secondary engagement has an inherent 'aboutness' to it; the interaction between caregiver and infant is about the object.

NOTE for codes 2-3: If it is largely the baby rendering caregiver's efforts ineffective, describe this in general comments for the video. Take the degree of effectiveness into account in your rating, but make note of whether this is due to the infant's behavior or not.

NOTE: infant's attention must either start out on the object or on the mother for joint efforts to begin. That is, primary attention on the object or caregiver must be present prior to joint

attention. For example, when infant's attention is on the object, a caregiver might use intonation or vocal melody as a tool to scaffold joint attention.

3. Contingent Responding (CR)

Contingent responding to shifts in infant's engagement with (interest in) the object (rate on 1-5 scale of contingency)

1—No sign or almost no sign of contingent responding to shifts in infant's engagement with (interest in) the object (for example: caregiver does not change behavior or respond when infant shows signs that they are not engaged with object (e.g. infant looks away, plays with another person or object), or infant shows signs that they are frustrated/overstimulated by attempts at engagement (e.g. squirming, crying)). Conversely, caregiver does not change behavior when infant shows increasing interest in the object (they do not gesture or speak more animatedly, for instance).

2—Minor, inconsistent, and/or largely ineffective contingent responding to shifts in infant's engagement with (interest in) the object (for example: caregiver makes small shifts in responses to infant when he/she displays the signals above, but does not persistently redirect the infant toward the object throughout an epoch or gives up after one or two attempts.

3—Moderately clear, fairly consistent, and/or moderately effective contingent responding to shifts in infant's engagement with (interest in) the object (for example: caregiver responds contingently to infant's shifts in engagement with an object about half of the time). A caregiver might gesture more animatedly with the object after infant smiles during their interaction but continues to be animated after the infant gives signals that they cannot follow the motion of the object or are no longer engaging with the object due to the large gestures.

4—In general, there are clear, consistent, and/or effective contingent responding to shifts in infant's engagement with (interest in) the object (for example: caregiver demonstrates consistent awareness of infant's engagement with object and adjusts behavior when infant the gives signals of distress or enjoyment, with small exceptions (e.g. infant begins crying when object is held too close to their face and it takes ~5 seconds for a mother to move it away and catch the infant's interest in the object in a more positive way)).

5—Strong, clear, highly consistent, and/or highly effective contingent responding to shifts in infant's engagement with (interest in) the object (for example: throughout an entire epoch, caregiver shows an exemplary ability to pick up infant's signals of interest in an object and changes behavior quickly (~1-2 seconds) whenever an infant's interest in the object wanes or is no longer jointly held).

4. Affective Tone (AT)

A caregiver's warm affective tone and degree of enthusiasm toward infant persists throughout an epoch. Caregiver shows an interest in the task and the interaction with their infant, and they do not appear distracted or confused while interacting with their infant (on a 1-5 scale).

1—No sign or almost no sign of a warm affective tone and enthusiasm toward infant, possibly even hostility toward infant (for example: caregiver does not relate to their infant in a positive and focused way (e.g., looks away or around at others or talks to others, spends an epoch re-adjusting their infant, appears frustrated when infant struggles through an epoch).

2—Minor, inconsistent, and/or largely ineffective warm affective tone and enthusiasm toward infant (for example: caregiver talks very quietly and gestures with object very little in 3rd epoch, does not appear to necessarily enjoy the interaction with their infant, needs reassurance from researcher).

3—Moderately clear, fairly consistent, and/or moderately effective warm affective tone and enthusiasm toward infant (for example: caregiver shows a positive, but downplayed, interest in the object and infant throughout a task and is generally focused and calm).

4—In general, there are clear, consistent, and/or effective warm affective tone and enthusiasm toward infant (for example: caregiver has full attention on the infant and object throughout an epoch and demonstrates a positive affect, but only demonstrates a limited range of behaviors with positive affect (e.g. just voice and smiling)).

5—Strong, clear, highly consistent, and/or highly effective warm affective tone and enthusiasm toward infant (for example: in addition to 4, demonstrates a wide range of warm or positive affective responses).

Bibliography

Adamson, L. B., Bakeman, R., Suma, K., & Robins, D. L. (2019). An expanded view of joint attention: Skill, engagement, and language in typical development and autism. *Child Development*, 90(1).

Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. *Becoming a Word Learner*, 115–135.

Baldwin, D. A. (1991). Infants' contribution to the achievement of Joint Reference. *Child Development*, 62(5), 875–890.

Baldwin, D. A. (2000). Interpersonal understanding fuels knowledge acquisition. Current Directions in Psychological Science, 9, 40-45.

Baron-Cohen, S., Baldwin, D. A., & Crowson, M. (1997). Do children with autism use the speaker's direction of gaze strategy to crack the code of language? *Child Development*, 68(1), 48.

Barton, M. E., & Tomasello, M. (1991). Joint attention and conversation in mother-infant-sibling triads. *Child Development*, 62(3), 517-529.

Chan, K., Gallant, J., Leemaqz, S., Baldwin, D. A., Borath, M., Kroeun, H., ... & Whitfield, K. C. (2021). Assessment of salt intake to consider salt as a fortification vehicle for thiamine in Cambodia. *Annals of the New York Academy of Sciences*, 1498(1), 85-95.

Family Life Project Parent Coding Manual. Martha J. Cox, Ph.D. and Keith Crnic, Ph.D. (2003), University of North Carolina, Chapel Hill.

Fattal-Valevski, A., Azouri-Fattal, I., Greenstein, Y.J., Guindy, M., Blau, A., & Zelnik, N.

(2009). Delayed language development due to infantile thiamine deficiency. *Developmental Medicine & Child Neurology*, 51(8), 629–634.

Gallant, J., Chan, K., Green, T. J., Wieringa, F. T., Leemaqz, S., Ngik, R., ... & Whitfield, K. C. (2021). Low-dose thiamine supplementation of lactating Cambodian mothers improves human milk thiamine concentrations: a randomized controlled trial. *The American Journal of Clinical Nutrition*, 114(1), 90-100.

Measelle, J. R., Baldwin, D. A., Gallant, J., Chan, K., Green, T. J., Wieringa, F. T., ... & Whitfield, K. C. (2021). Thiamine supplementation holds neurocognitive benefits for breastfed infants during the first year of life. *Annals of the New York Academy of Sciences,* 1498(1), 116-132.

Mundy, P., & Newell, L. (2007). Attention, joint attention, and social cognition. *Current Directions in Psychological Science*, 16(5), 269–274.

Smith, C. B., Adamson, L. B., & Bakeman, R. (1988). Interactional predictors of early language. *First Language*, 8(23), 143-156.

Tomasello, M., & Farrar, M. J. (1986). Joint attention and early language. Child development, 1454-1463.

Whitfield, K. C., Bourassa, M. W., Adamolekun, B., Bergeron, G., Bettendorff, L., Brown, K. H., Cox, L., Fattal-Valevski, A., Fischer, P. R., Frank, E. L., Hiffler, L., Hlaing, L. M., Jefferds, M. E., Kapner, H., Kounnavong, S., Mousavi, M. P. S., Roth, D. E., Tsaloglou, M. N., Wieringa, F., & Combs, G. F. (2018). Thiamine deficiency disorders: Diagnosis, prevalence, and a roadmap for Global Control Programs. *Annals of the New York Academy of Sciences*, 1430(1), 3–43.

Whitfield, K. C., Kroeun, H., Green, T., Wieringa, F. T., Borath, M., Sophonneary, P., ... &

Gallant, J. (2019). Thiamine dose response in human milk with supplementation among lactating women in Cambodia: study protocol for a double-blind, four-parallel arm randomised controlled trial. *BMJ open*, 9(7).

Woodward, A. L. (2003). Infants' developing understanding of the link between Looker and Object. *Developmental Science*, 6(3), 297–311.

HOW TO QUANTIFY "GOOD SLEEP":

A SPECTRAL ANALYSIS OF SLEEP SPINDLE MORPHOLOGY IN HEALTHY ADULT

EEG AND THE ROLE OF SLEEP SPINDLES IN AGING AND NEURODEGENERATION

by

ELYRIA KABASENCHE

A THESIS

Presented to the Department of Psychology
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

May 2022

An Abstract of the Thesis of

Elyria Kabasenche for the degree of Bachelor of Science
in the Department of Psychology to be taken June 2022

Title: How to Quantify "Good Sleep":
A Spectral Analysis of Sleep Morphology in Healthy Adult EEG and the Role of Sleep
Spindles in Aging and Neurodegeneration

Approved: _David M. Condon, Ph.D._
Primary Thesis Advisor

The importance of good sleep cannot be overstated. What makes sleep "good", productive, and beneficial are of interest to any sleep researcher. Studying morphology of sleep features can provide insight about what differentiates healthy and unhealthy sleep and create benchmarks for recognizing instances when characteristics such as aging and disease may be impacting sleep quality. The purpose of this study was to examine an N2 sleep feature termed a sleep spindle and conduct an analysis of morphology on a sample of healthy adult EEG using recently validated and created sleep spindle detection algorithm to create a baseline measurement for spindle presence. The effect of age on spindles was of particular interest and was found to be related to a decrease in spindle length. The possible reason for this effect is discussed, as well as future applications for use of this algorithm and spindle analysis

Acknowledgements

I would like to thank Dr. David Condon for serving as my primary advisor and for his unfailing support in the development and execution of this thesis. It would not have been possible without him. Thank you as well to Dr. Don Tucker, for access to the data analyzed and his enormous amount of knowledge about and passion for EEG science, and for providing the Brain Electrophysiology Lab as a wonderful undergraduate research experience. Thank you to Dr. Liska Chan for serving on my Clark Honors College thesis committee. And finally, thank you to the CHC and the Department of Psychology for fostering such a wonderful educational experience.

The support and encouragement of my family, particularly Bill and Eliason Kabasenche, and my friends throughout the pandemic, the stress of college, and the culmination of my education in the form of this thesis have proven invaluable. Thank you all for the love and encouragement.

The creators of this template (CHC Librarian Miriam Rigby & CHC Academic & Thesis Programs Manager Miriam Jordan) would like to thank Reed College for providing their Thesis Template for the inspiration of many elements of this template.

Table of Contents

List of Figures

Figure 1. Spindle Duration (seconds) as an Effect of Participant Age (years)
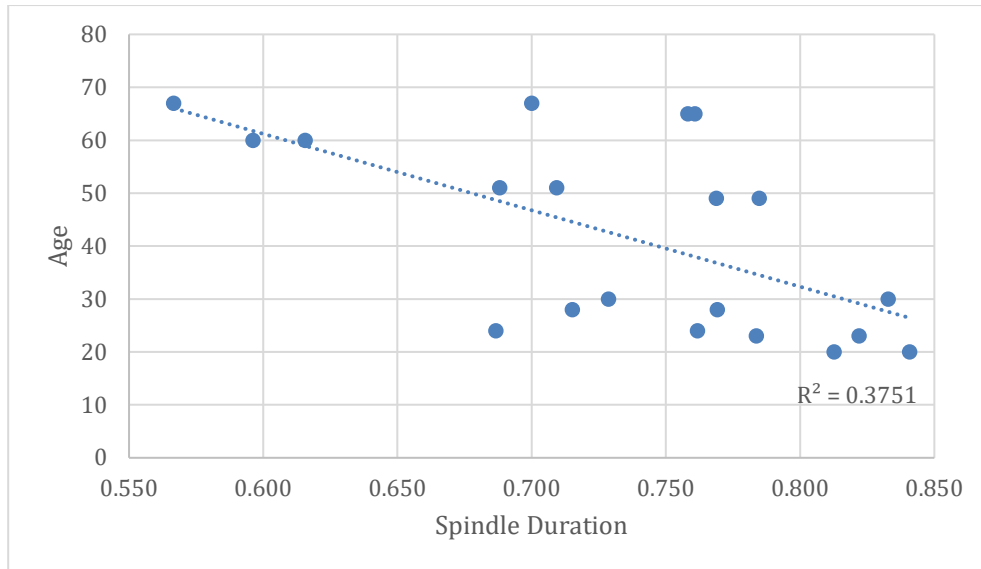
List of Tables

Regression Statistics	
Multiple R	0.612452455
R Square	0.375098009
Adjusted R Square	0.340381232
Standard Error	14.70924537
Observations	20

ANOVA					
	df	SS	MS	F	Significance F
Regression	1	2337.685813	2337.685813	10.80451697	0.004096976
Residual	18	3894.514187	216.3618993		
Total	19	6232.2			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%	Lower 95.0%	Upper 95.0%
Intercept	147.845253	32.45929062	4.55479002	0.000245494	79.65081391	216.0396921	79.65081391	216.0396921
X Variable 1	-144.4006835	43.93054294	-3.28702251	0.004096976	-236.6953294	-52.1060376	-236.695329	-52.10603757

Introduction

One third of the human life is spent sleeping, and the quality of that sleep is very important for physical and mental function. The importance of "good sleep" is widely alluded to, by physicians and parents alike, but recordings such as an electroencephalogram (EEG) allows researchers to view what the brain is doing during sleep. This information can give insight to questions such as "How does sleep help memory?' and "How much REM sleep is normal?" Establishing a baseline for "normal" sleep features, including duration, allows for the use of sleep recordings as a non-invasive way to characterize abnormal sleep patterns, including departures from normative sleep patterns at the population level (i.e., between people) or departures from normal patterns within an individual.

Many factors can contribute to changes in sleep quality: aging, mood disorders, and neurodegenerative diseases are just a few examples. However, each of these factors changes sleep in a distinctive way prompting different targets for treatment intervention. Certain features of sleep, like REM staging and slow waves, have been studied extensively. Others, such as sleep spindles (described below), remain relatively unexamined. Thus, much more work is needed to understand the relative importance and prevalence of these features in normal adult sleep. One of the best ways to examine sleep comes from recording the brain's aggregated electrical activity throughout the night, in the form of an electroencephalogram (EEG). Studying this EEG requires people who have been trained in identifying relevant sleep features or using machine learning to create an algorithm that is as efficient at detecting sleep features as a human

would be. Lacourse et al. (2019) created a sleep spindle detection algorithm comparable to highly trained human raters; the goal is to exceed that algorithms capability.

Why is analyzing sleep features important? Creating a study where a specific type of brain activity during sleep can be linked to changes in memory, performance, or disease is immensely valuable both for understanding brain function and improving sleep quality. For example, at least one link between aging and memory decline has been proposed based on analysis of sleep characteristics; Scullin (2012) found that older adults who perform worse on memory tasks have a corresponding decrease in slow wave sleep. Implications of this finding are that if the duration of slow wave sleep could be increased in older adults, perhaps memory decline would not be as dramatic and devastating. While slow wave sleep has been linked to memory, there are other sleep features that have not been studied as much and their effect and purpose remains relatively unknown.

Sleep spindles, which occur primarily in non-rapid eye movement stage 2 of sleep (N2 and N3) are one of those such sleep features. Spindles are smooth sinusoidal waves that occur during sleep, in relatively short (greater than 0.5 seconds) bursts of 11-16 Hz activity. They were first observed 80 years ago in early EEG sleep observations, but it wasn't until 30 years ago that it was determined sleep spindle activity is driven by thalamocortical loops in the forebrain (Fernandez and Lüthi, 2020). Spindle activity and slow wave activity are inversely related in periods of sleep, suggesting that they play different roles in sleep homeostasis. In addition to the regulatory role of spindles, they have implications in cognitive functions. They could become an increasingly important measure of sleep quality as aging occurs, due to the previously stated decrement in slow

wave activity also corresponding to aging. The decision to examine sleep spindles comes from a desire to understand function, establish baseline measurements for shape and frequency of the spindles, and ensure that there in an algorithm that is able to identify and note those features with at least the same accuracy as an expert sleep rater. One potential application for spindle detection centers around the understanding that key characteristics of spindles change when neurological diseases are present. Understanding how and why they change in these instances can offer insight into the mechanism of disease and potential treatment or sleep therapy.

Spindles are typically detected by use of electroencephalography (EEG) recordings of sleep. This results in 8-hour recordings for each individual assessed. Manually going through those recordings to mark and measure features is time consuming and introduces human error into the process. The use of an algorithm in this study ensures that machine learning can be used to eliminate time constraints and introduce higher levels of precision into data recording and analysis. Creating an algorithm that compares well to gold standard EEG raters would mean that labs or hospitals would not need to allocate as many resources or personnel to EEG analysis.

The goal of the study will be to use this algorithm to examine differences in spindle density across participants. Because all participants are healthy adults with no sleep complications, the main effect to examine will be how age shapes sleep spindle morphology and what potential applications that has for cognitive function.

Proposed Investigation

In sleep scoring, expert raters are considered the gold standard for identifying sleep features such as sleep spindles. However, manually examining an 8-hour EEG recording is time consuming; doing so for an entire study's worth of participants is a massive demand. Computer learning is a useful alternative assuming sufficient sensitivity and discretion can be taught.

Lacourse et al. (2019) present an algorithm referred to as A7. While many other algorithms exist, the A7 algorithm is open source, allowing for continual testing and improvement. The aim of this study is to use this algorithm to analyze previously collected data to measure sleep spindle density and plot how this changes as age of participants increases.

Existing Literature

This study attempts to build on previous literature and algorithms created to analyze sleep spindles and examine the possible applications for detecting abnormalities in spindle quantity and density, specifically related to aging. Sleep spindles are related to aspects of cognition such as learning and memory and seem to change when neurological diseases are present.

The primary paper used to direct the study was Lacourse et al (2019), as this work introduced the A7 algorithm used for spindle detection and analysis. It was created to reduce the time and expense of using human raters and to show that an algorithm could reach the same amount of agreement as human raters. The six automated spindle detection algorithms previously created did not show high validation with human scorers; algorithm #7 (A7) set out to be different by comparing performance to human raters and four other detection algorithms. With detection parameters of absolute sigma (the wavelength of spindle waves) power, relative sigma power, sigma covariance, and sigma correlation, the A7 identified a spindle when all those parameters exceed their threshold. In tests, A7 received the highest F score, 0.17 points above the next detection algorithm and comparable to human raters, who were 0.03 points below A7. While each method of spindle detection is yields different advantages, this was a high level of spindle detection that agreed well with gold standard rated epochs. Because this algorithm has already been validated, further use will assume that validation holds.

Fernandez and Lüthi (2020) offer a brief summary of spindle characteristics, location, and implications for cognitive activity. Fernandez and Lüthi describe how

spindles play a role in memory consolidation by supporting synaptic plasticity, and how the interplay between the thalamus and the cortex (essential for cognition) is the same circuit that underlies spindles. Neurodevelopmental disorders such as attention disorders and autism that have a genetic component often express those genes in the thalamic nucleus, so differences in spindle expression are typically present in individuals with those disorders. Epilepsy, Alzheimer's, schizophrenia, Parkinson's, and many more neurodegenerative/developmental disorders also show marked differences in spindle density or amplitude, suggesting that spindles could serve as some sort of predictive biomarker for cognitive abnormality, disease, and decline.

Latrielle et al. (2015) examined in more detail the link between spindle characteristics and individuals with Parkinson's disease specifically. EEG sleep recordings were taken from a sample of Parkinson's patients without dementia and a healthy control sample, and slow wave and spindle activity was measured. Then, 4.5 years later, the same tests were given to examine the effects of aging and disease progression. About 30 percent of the Parkinson's sample had developed dementia in that time, and spindle density decrease was noted in those individuals, both compared to baseline and to the healthy population. While slow wave sleep was decreased compared to controls, there was no difference between Parkinson's patients who developed dementia and those who did not. This suggests that lower spindle density and frequency in Parkinson's patients may be predictive of decline into dementia.

Christensen et al. (2015) also examined sleep spindles in Parkinson's patients but chose to look specifically at characterizing morphological differences. A Parkinson's group and a control group were compared, and significant differences were

found between the groups, giving an idea of morphological differences that may affect or result from neurodegeneration. Spindle density was decreased, spindles themselves were longer and at a lower frequency, and peak to peak amplitude was higher compared to controls, which were age and sex matched. This work suggested that Parkinson's disease may affect (either directly or not) the area of the brain responsible for spindle generation (thalamus), although more research would need to be done to understand this unique effect as opposed to other neurodegenerative diseases, as well as to ensure this effect is not related to Levodopa and other similar drugs. Christensen concluded by stating the importance of making sure algorithms exist to detect abnormal spindle activity, because spindle detectors are likely to be used in analyzing EEG of neurodegenerative disease patients.

Astori et al. (2013) review the physiology of sleep spindles and their pace making in order to offer some insight as to how manipulating spindles may affect neural function. In humans, fast spindles occur over parietal and central areas of the cortex and slower ones are localized to the frontal cortex; Astori et al. found that fast spindles appear to couple with slow waves in memory consolidation related events between the hippocampus and the cortex, while slow spindles then recruit frontal areas as memory storage space. Another proposed function of sleep spindles relates to the thalamus as the relay station for sensory signals. Spindles may serve a protective function for sleep by filtering out excess stimuli. fMRI studies showed auditory cortex activation to noise during NREM sleep, but that activation was absent during spindle events. The review also references ways in which model systems may be used in conjunction with

optogenetics to explore the thalamocortical loops and cortical feedback underlying

spindle instances.

Methods

Previously collected EEG data from a sleep study at the Brain Electrophysiology Lab was used for analysis. Of the 10 subjects used, 5 were female (*M=41.70, SD=18.6).* Participants were partially a convenience sample of family and friends and partially recruited older adults with no health complications. The original study spanned three nights, but only the second and third nights of data were used for analysis, which included a randomized night of neuromodulation to investigate its effect on slow wave sleep. Neuromodulation entailed a block of stimulation at slow wave frequency (0.5-4 Hz) at the first detected sign of N2 sleep (either a sleep spindle or a K-complex, another marker of N2 sleep) in order to try and induce or enhance slow wave activity. This EEG had previously been down sampled, filtered, marked with artifact on bad channels (which were replaced), had N2 and N3 segments extracted and parsed into 15 second segments, and had each segments marked as artifactual or not. This resulted in the files which were used for analysis, after being converted into a .mat file format. This meant the EEG was clean and artifact free, and segments of interest for this analysis had been extracted. A script was run to transform each original .mff file into numerical representations for each of the parameters the A7 algorithm uses to detect spindles. These comprise of the sleep stage, if there is artifact in the segment, and the signal data from the EEG itself.

After the files had been properly configured, they were run through the A7 algorithm in MATLAB, which resulted in spindle detection and information such as spindle duration and if spindles were expected to be found in that segment of sleep. The A7 algorithm uses a central channel towards the top of the head and a reference channel

on one of the mastoids. These were C3 and M2 on a 10-20 array EEG net. Comparable channels were found to be channel 59 and the right mastoid channel on the EGI dense array 256 electrode EEG nets.

Running the algorithm on segmented data from nights 2 and 3 from each of the 10 participants resulted in detection of 7148 sleep spindles. These spindles were characterized with where in the segment the spindles started and ended (seconds), the duration of the spindle (seconds), the sleep stage in which the spindle was found, and the number of the segment, so the spindle could be located in the original EEG segment. Spindles were randomly checked to ensure the numerical output aligned with visual detection of a spindle waveform activity. Of the 10 spindles examined, the algorithm detection corresponded with visual detection. This is to be expected, because the A7 algorithm has already been validated, but the additional verification was reassuring.

Results

Spindle measures were output into a text file (.txt), which was converted to an excel spreadsheet. A profile was created for each participant, where the N2 and N3 spindles recorded for each of the nights slept were averaged. N3 data was much less prevalent, with some nights only yielding detection of 1 or 2 spindles. For this reason, N3 data was excluded from analysis. Additionally, N2 is the stage of sleep where spindles are a benchmark sleep feature and most literature focuses on that stage alone, so this study chose to follow that trend and only report basic measures of N3 sleep spindles, which across all participants, resulted in (M=0.726, SD= 0.172) across 339 spindles. N2 data was much more prevalent, with 6809 spindles detected (M=0.735, SD= 0.077).

The primary question was what effect, if any, age had on spindle morphology. A regression analysis was run to examine the relationship between age and spindle density (see tables 1, 2, and 3). A negative correlation was found (R^2=0.375, $F(1,19)$= 10.80, p= 0.004) (see figure 1). Additional T-tests were run to compare the conditions of gender and the effect of neuromodulation (if one exists). There did not appear to be an effect of electrical stimulation (M=0.758,SD=0.069) on sleep spindles compared to uninterrupted sleep (M=0.792,SD=0.081) ($t(18)$= 1.353, p=0.193). There was a significant effect of gender ($t(18)$=2.483, p= 0.023) with females (M=0.773,SD=0.048) having significantly longer spindles compared to males (M=0.697,SD=0.084).

Discussion

The results were expected, with the exception of the effect (or lack thereof) of neuromodulation on sleep spindles. The intention of the neuromodulation in the original study was to target and enhance slow oscillations in N2 sleep. While it does seem that slow waves are linked to sleep spindles, their relationship has not been studied in much detail and any attempt to alter slow oscillations does not appear to have subsequently affected sleep spindles in this analysis. Future directions in this realm could include pairing this algorithm with a slow oscillation detector and attempting to examine the pairing of spindles with slow waves. Both are markers of N2 and have implications for cognitive function. Yordanova et al. (2017) examine the temporal coordination of the different types of sleep spindles (fast and slow) with slow waves and the implications this has for memory consolidation based on a pre-sleep task. A similar investigation could be carried out with methodology much like this study; reliable detection of spindles would make investigating the link between them and slow waves more feasible.

An effect of age was expected, with a negative correlation being found. This relationship was consistent with existing literature about spindle morphology. As participants age, spindle duration decreases. Nicolas et al. (2001) conducted an extensive investigation of sleep spindle morphology over different age groups and found that density and duration both decrease in older adults. They noted that most of the changes seem to happen before the age of 40, and then spindles are relatively consistent until more dramatic effects of aging become apparent around the age of 70 or with onset of neurodegenerative disease. The "long maturation of the central nervous

system" is attributed to this, in addition to age-related changes of the thalamo-cortical pacemaker responsible for generating spindles. Nicolas et al. (2001) propose impairment in neural recruitment of the pacemaker or a desynchronization of neurons also in the chain of command as being partially responsible for this decrease in function.

There was a significant effect of gender found across N2 sleep spindles. Franco et al. (2020), when looking at sleep and gender-based development, found that there seem to be gender-based differences in sleep during development. Namely, they found both sleep spindles and slow waves (previously discussed to be linked in some way) have more density in females during development. This trend would seem to continue into adulthood as seen in this sample.

Limitations of this study include the relatively small sample size (10 participants, 2 nights each analyzed) and the lack of direction in terms of investigating spindles, as this was merely a baseline analysis. The A7 algorithm was only validated on N2 epochs, so including N3 epochs was more experimental, and the lack of comparable spindle detections suggests that there may be morphological differences in N2 and N3 spindles or sleep data such that the algorithm would need to be modified for use in studies involving N3 sleep. Data collection was cut short due to COVID-19, so being able to continue analyzing sleep data as it is collected and adding to the understanding of spindles will allow for more reliability in results.

Additionally, testing the spindle-slow wave pairing that has been previously reported in literature would be a good application of the A7 algorithm. The spindle detection could be conducted entirely with the algorithm, and that should orient to

paired slow waves. A different detection method might need to be utilized to cross reference slow waves. Another analysis to run would be comparing these results to a population of adults with a neurodegenerative disease, such as Parkinson's. Those sleep spindles would be expected to be diminished, and creating a profile for spindle differences could be used to characterize severity of disease onset.

To conclude, there was a significant effect of age found in the sample of data examined. There was no effect of gender or neuromodulation in this sample. The A7 algorithm was able to successfully detect sleep spindles and give morphological information about them, with implications for future research involving sleep spindles and perhaps examining them in conjunction with slow waves to look at temporal pairing and cognitive function.

Glossary

Artifact: When EEG recordings pick up some activity other than action potentials, such as when the participant moves, resulting in inconsistent, abnormal, and unreadable EEG

Electroencephalogram (EEG): An electrical recording that aggregates brain activity by recording electrical charges caused by action potentials via a web of channels over the skull

Morphology: Changes of structure (in the case of EEG, amplitude, frequency, waves shape, and density are examples of morphological characteristics)

Parkinson's Disease: A neurodegenerative disease that begins destroying dopaminergic neurons in the substantia nigra and results in pathological motor changes and deterioration

Sigma: A waveform that oscillates at 11-15 Hz, found in NREM sleep

Sleep Spindle: A burst of sigma activity, less than 0.5 seconds and in the 11-15 Hz range

Thalamus: A brain structure located above the brain stem, which processes sensory and motor information and relays signals to the cerebral cortex

Bibliography

Astori, S., Wimmer, R. D., & Lüthi, A. (2013). Manipulating sleep spindles – expanding views on sleep, memory, and disease. *Trends in Neurosciences*, *36*(12), 738–748. https://doi.org/10.1016/j.tins.2013.10.001

Christensen, J. A. E., Nikolic, M., Warby, S. C., Koch, H., Zoetmulder, M., Frandsen, R., Moghadam, K. K., Sorensen, H. B. D., Mignot, E., & Jennum, P. J. (2015). Sleep spindle alterations in patients with Parkinson's disease. *Frontiers in Human Neuroscience*, *9*. https://www.frontiersin.org/article/10.3389/fnhum.2015.00233

Ellenbogen, J. M. (2005). Cognitive benefits of sleep and their loss due to sleep deprivation. *Neurology*, *64*(7), E25–E27. https://doi.org/10.1212/01.wnl.0000164850.68115.81

Fernandez, L. M. J., & Lüthi, A. (2020). Sleep Spindles: Mechanisms and Functions. *Physiological Reviews*, *100*(2), 805–868. https://doi.org/10.1152/physrev.00042.2018

Franco, P., Putois, B., Guyon, A., Raoux, A., Papadopoulou, M., Guignard-Perret, A., Bat-Pitault, F., Hartley, S., & Plancoulaine, S. (2020). Sleep during development: Sex and gender differences. *Sleep Medicine Reviews*, *51*, 101276. https://doi.org/10.1016/j.smrv.2020.101276

Helfrich, R. F., Mander, B. A., Jagust, W. J., Knight, R. T., & Walker, M. P. (2018). Old Brains Come Uncoupled in Sleep: Slow Wave-Spindle Synchrony, Brain Atrophy, and Forgetting. *Neuron*, *97*(1), 221-230.e4. https://doi.org/10.1016/j.neuron.2017.11.020

Lacourse, K., Delfrate, J., Beaudry, J., Peppard, P., & Warby, S. C. (2019). A sleep spindle detection algorithm that emulates human expert spindle scoring. *Journal*

of Neuroscience Methods, *316*, 3–11.
https://doi.org/10.1016/j.jneumeth.2018.08.014

Latreille, V., Carrier, J., Lafortune, M., Postuma, R. B., Bertrand, J.-A., Panisset, M.,
   Chouinard, S., & Gagnon, J.-F. (2015). Sleep spindles in Parkinson's disease
   may predict the development of dementia. *Neurobiology of Aging*, *36*(2), 1083–
   1090. https://doi.org/10.1016/j.neurobiolaging.2014.09.009

Le Bon, O. (2020). Relationships between REM and NREM in the NREM-REM sleep
   cycle: A review on competing concepts. *Sleep Medicine*, *70*, 6–16.
   https://doi.org/10.1016/j.sleep.2020.02.004

Nicolas, A., Petit, D., Rompré, S., & Montplaisir, J. (2001). Sleep spindle
   characteristics in healthy subjects of different age groups. *Clinical
   Neurophysiology*, *112*(3), 521–527. https://doi.org/10.1016/S1388-
   2457(00)00556-3

Sejnowski, T. J., & Destexhe, A. (2000). Why do we sleep?11Published on the World
   Wide Web on 7 November 2000. *Brain Research*, *886*(1), 208–223.
   https://doi.org/10.1016/S0006-8993(00)03007-9

Yordanova, J., Kirov, R., Verleger, R., & Kolev, V. (2017). Dynamic coupling between
   slow waves and sleep spindles during slow wave sleep in humans is modulated
   by functional pre-sleep activation. *Scientific Reports*, *7*(1), 14496.
   https://doi.org/10.1038/s41598-017-15195-x

CULTURAL CONSIDERATIONS

IN THE FILMING INTERACTIONS TO NURTURE

DEVELOPMENT INTERVENTION

by

ALVIN LENGKONG

A THESIS

Presented to the Department of Psychology
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

An Abstract of the Thesis of

Alvin Lengkong for the degree of Bachelor of Science
in the Department of Psychology to be taken Spring 2022

Title: Cultural Considerations in the Filming Interactions to Nurture Development

Intervention

Approved: _____*Title and Full Name (or Title at end of Full Name)*_____

Primary Thesis Advisor

The Filming Interactions to Nurture Development (FIND) intervention is a strengths-based video coaching program designed to disrupt the consequences of early exposure to toxic stress and promote naturally occurring, developmentally supportive interactions between infants and their caregivers. While the FIND intervention has been shown to be effective in improving certain child and parent outcomes, no prior study has explored the importance of cultural factors when implementing the FIND intervention. The focus of this paper is to better understand the cultural differences between English-speaking and Spanish-speaking families, and to identify if these differences influence the intervention's effects on parent outcomes (i.e., parent sense of competence, parental stress, and parental self-efficacy). Data was derived from a randomized controlled trial using a pretest-posttest design. Families were randomly assigned to an active control or the FIND intervention group. Results indicated no significant interactions between preferred administration language and condition. Caregivers in the FIND condition were found to improve significantly more in their parent-reported sense of competence from pre- to post-intervention compared to those in the active control condition.

Additionally, Spanish-speaking caregivers saw a significantly greater decrease in parent-reported stress from pre- to post-intervention than English-speaking families. Implications and limitations of the FIND intervention's cultural considerations will be discussed.

Acknowledgements

I would first like to thank graduate student Andrea Imhof, who has always been someone I could rely on. I express my sincerest gratitude for having the privilege of working with someone like Andrea. She worked tirelessly to guide and mentor me throughout this entire process, regardless of if my needs were related to my project or not. Without her constant support, this thesis would not exist. Her investment of time, effort, and energy to work with and encourage me is evident in every aspect of my thesis.

I would like to thank my primary thesis advisor Dr. Philip Fisher for the opportunity to work within the Stress Neurobiology and Prevention lab to complete my thesis. I am very grateful for you allowing me to be a part of such an amazing project and team.

I would like to thank my Clark Honors College representative and advisor Dr. Dare Baldwin for teaching me so much during my four years in the honors college. Whether it was as an advisor, a professor, or a part of my thesis committee, she has allowed me to have a college experience that I will always cherish.

I would like to thank the University of Oregon, the Robert D. Clark Honors College, the College of Education, and the Department of Psychology for giving me the opportunity to grow as a student and a person.

I would also like to thank my family in Eugene, Oregon and in Jakarta, Indonesia. They have always been with me to encourage and pick me up regardless of the circumstances. We have been through so much together these past few years, and I

just want to express how much I love you all, and that I could never imagine myself getting to this point without your constant love and support.

Lastly, I would like to thank my mama. Your unconditional love has no bounds and pushes me to be the best student and man I can be. Everything I have become and everything I have done is only possible because of what you have sacrificed for me, Tata, and Nana. You are the glue that holds our family together, and you motivate me to excel in everything I do. I am here because of you. Love You Always.

Table of Contents

Introduction

The activation of the hypothalamic-pituitary-adrenocortical (HPA) axis and the ensuing increase in production of cortisol is the major endocrine mechanism behind the mammalian reaction to stress (Aguilera, 1994; Tarullo & Gunnar, 2006). Early life stress (e.g., non-responsive caregiver, multiple caregiver transitions, neglect) has been found to have a significant impact on the future functioning of a child's stress response system and especially of the HPA axis. Dysfunction of the HPA axis has been linked to an increased likelihood of a variety of poor physical and physiological health outcomes in adolescence and adulthood (Bruce et al., 2013; Halligan et al., 2007, Sánchez et al., 2001; Tarullo & Gunnar, 2006).

There is significant evidence that HPA axis dysregulation is associated with a variety of poor outcomes later in life (Guerry & Hastings, 2011; Guilliams & Edwards, 2010). Research has shown that some interventions can disrupt this cycle and even reverse these effects (e.g., Fisher et al., 2007). More specifically, research has shown that a strong, supportive relationship between a child and caregiver can be a significant protective factor that has long-term benefits for both the caregiver and the child (Cicchetti et al., 1991; Gunnar et al., 1996). A strong, established relationship between an infant and a caregiver can act as a "buffer" that regulates and dampens the infant's physiological response to stress (Flannery et al., 2017; Hostinar et al., 2014). For example, parents who were higher in sensitivity and responsiveness when their children were 2, 4, and 6 months old had children who exhibited significantly smaller cortisol responses to a stressful experience in infancy than parents who were lower in those variables (Gunnar et al., 1996). This protective effect has also been replicated in a variety of different stressful environments (Gunnar et al., 1996). These findings have significant implications in a

clinical setting, and many interventions now seek to establish a strong, supportive attachment between caregivers and their children.

FIND Intervention

Among these interventions is the Filming Interactions to Nurture Development (FIND) intervention, a strengths-based video coaching program designed to disrupt the consequences of early exposure to toxic stress. The FIND intervention is a brief program for caregivers and is designed to promote naturally occurring, developmentally supportive interactions between infants and their caregivers (Fisher et al., 2016). These interactions, often labeled "serve and return" interactions, occur when an infant "serves" by presenting a signal or gesture and a caregiver "returns" that serve by responding in an appropriate or nurturing manner (Fisher et al., 2016; Schindler et al., 2017; Shonkoff & Bales, 2011). Even the smallest back-and-forth interaction can be considered a serve and return interaction. Serve and return interactions are crucial for infant brain development, influencing later life outcomes such as academic performance and mental health (Schindler et al., 2017; Shonkoff & Bales, 2011). Young children naturally seek interaction with their environment, and the way that caregivers respond to these cues has a significant impact on the child's development.

FIND is flexible, brief, and appropriate for use across a variety of contexts. The FIND intervention has been implemented in a variety of different settings including in homeless shelters and primary care settings. The intervention is derived from microsocial interaction research at the Oregon Social Learning Center (Patterson et al., 1992) and the Marte Meo video coaching intervention (Axberg et al., 2006; Fisher et al., 2016) and focuses on supportive "serve and return" interactions. These interactions are broken down into five core elements that are

described in behavioral terms to make it easier for caregivers to understand and implement (Fisher et al., 2016). The first element is *Sharing the Focus,* which occurs when the caregiver notices and shows interest in the stimuli that the child is focused on. The second element is *Supporting and Encouraging*, which includes the caregiver responding by offering help/support, acknowledging the infant, or praising the infant. The third element is *Naming*, which occurs when the caregiver verbally labels what the infant is seeing, doing, or feeling. The fourth element is *Back and Forth,* which is an extension of the serve and return process that takes place when a caregiver "returns" an infant's "serve" and follows that by waiting for the infant's further initiations. The fifth and final element is *Endings and Beginnings*, which includes the transitions from one back and forth interaction to another (Fisher et al., 2016).

There are several aspects of the FIND intervention that makes it unique. First, the FIND intervention presents caregivers the opportunity to view their interactions with their infant from a different perspective (i.e., through videos of themselves rather than actors). Using recorded videos also allows implementation of these interventions to be highly flexible, which is especially important when working with parents. Additionally, the FIND intervention utilizes strictly a strengths-based approach. Rather than focusing on what the caregiver did wrong or could improve on, the coach uses clips of the caregiver successfully using a core element and builds on these examples. The FIND intervention is specifically implemented in this way to increase caregiver self-efficacy (Fisher et al., 2016; Liu et al., 2021) – the degree to which a parent feels competent and confident in raising their children and handling problems (Johnston & Mash, 1980).

Relevance

The importance of the child-caregiver relationship and parental self-efficacy may be especially relevant for families living in poverty, whose environmental stressors often force their way into all aspects of life. There are a number of environmental risk factors associated with childhood poverty: increased family turmoil, family instability, less social support, and poorer schools/childcare (Evans, 2004). Without intervention, these early life stressors can lead to poor long-term outcomes for both children and caregivers. Caregivers living in poverty are also more likely to resort to negative parenting styles such as corporal punishment and reduced parent-child interactions (Albright & Tamis-LeMonda, 2002; Bluestone & Tamis-LeMonda, 1999; Sheely-Moore & Bratton, 2010). Living in poverty is the harsh reality that a growing number of families are forced to face. Alongside the beginning of the coronavirus pandemic, the United States saw an increase in its poverty rate from 10.5% in 2019 to 11.4% in 2020. Perhaps even more surprising, a staggering 17% of Hispanic individuals lived in poverty compared to just 8.2% of non-Hispanic Whites (Shrider et al., 2021). The Hispanic population continues to be disproportionately represented below the United States poverty line (Quillian, 2012). The impact of this overrepresentation is even more pronounced for immigrant Hispanic families. Families of Hispanic immigrants have been found to experience greater economic strain than nonimmigrant Hispanic or White families (Lubotsky, 2011).

Economic stressors, however, are far from the only significant stressors that Hispanic families have to overcome. In the United States, Hispanic families are met with an oftentimes contradictory patchwork of welcoming or stigmatizing sociopolitical environments that vary greatly depending on the context (Smokowski et al., 2008). For most families, acculturative stress also underlies the way that they react to all of these environments. Acculturative stress is

4

reductions in the health status of individuals that are related in a systematic way to known features of the acculturation process (Berry et al., 1987). The dual stressors of parenting and acculturation can lead to heightened levels of stress for both the caregiver and the child (Kim, 2008). High levels of acculturation stress are also associated with fewer positive parenting practices and compromised family functioning (Boruszak-Kiziukiewicz & Kmita, 2003; Lorenzo-Blanco et al., 2016; Miao et al., 2018). Cultural differences in rules, beliefs, preferences, codes of communication, and standards of competence have significant implications for the parenting of children (Calzada., 2010).

Early Head Start (EHS) Programs

The FIND intervention is designed specifically for at-risk families, which includes families living in poverty. However, there are also a number of other programs and services that are readily available for this population. Among these services, and perhaps the most well-known of them, are the Early Head Start (EHS) programs. These programs are designed to be flexible in their approach to meeting the needs of their communities, and specifically, low-income families of children from birth to age three. This approach includes providing support to caregivers and promoting positive child development in areas such as language, literacy, and social and emotional development (Love et al., 2005). To be eligible for EHS programs, families must have an income below the federally set poverty line. Additionally, all children in foster care, homeless children, and children from families receiving public assistance are eligible for EHS services regardless of income.

Present Study

Though past studies have shown that the FIND intervention is effective in improving parent and child outcomes (Giuliani et al., 2019; Liu et al., 2021), none of these studies have explored group intervention differences based on culture or language. The present study offers a novel framework for considering how cultural factors impact the efficacy of clinical interventions applied to those that comprise the majority of families living in poverty. To investigate if the FIND intervention is effective in a Hispanic population, the present study will use the caregivers' preferred administration language to divide the sample into two groups of families. Preferred administration language was chosen as an independent variable because language assimilation is one of the most commonly used and effective measures of acculturation (Kang, 2006; Thomson et al., 2009; Kirkman-Liff, 1991; Lee et al., 2009). Additionally, preliminary analyses ran by Benito-Gomez & Rojas (2021) found evidence of different relationships between parental stress and reciprocity between English-speaking and Spanish-speaking families. The present study will explore if the effectiveness of the FIND intervention on parental sense of competence, parental stress, and parental self-efficacy is influenced by whether families receiving the intervention considered English or Spanish as their preferred language. We hypothesize that there will be a significant interaction between the preferred administration language of the intervention and the intervention's effects on parental sense of competence, stress, and self-efficacy.

Methods

Procedures

The present sample included families recruited from EHS programs in the Denver metropolitan area. Recruitment strategies included a research coordinator speaking with enrolled caregivers about the study opportunity, disseminating recruitment materials, attending parent-focused events, and participants' word-of-mouth referral.

To be eligible for the current study, families needed to (1) be eligible to receive EHS services, (2) have children aged between 4 and 36 months old, and (3) be fluent in English or Spanish. A total of 202 families enrolled for the initial screening, during which additional eligibility criteria were assessed: eligible children (1) must not have any developmental delays and (2) must not regularly use any medications that interfere with cortisol assays.[1] After screening, 138 families were eligible and interested in participating in the study. Eligible families participated in a pre-intervention assessment session in which self-report surveys were used to collect data on specific caregiver and child variables. Caregiver-infant dyads were then randomly assigned into either the EHS-control condition (n=50) or the FIND-intervention condition (n=88). The uneven sample sizes in the active control and intervention groups were a deliberate design decision to include maximum access to active intervention in these vulnerable populations. Those in the control condition continued receiving EHS services for the 10-week period. Those in the intervention condition completed a 10-week span of FIND intervention sessions while also having access to EHS services.

---

[1] Data on caregiver cortisol levels was also collected but is not used in the present study.

After completing the intervention, families were assessed using the same measures a second time (for a full flow-chart of participants, see *Figure 1*). During the study procedure, contact with the families was maintained through weekly meetings with FIND coaches (for intervention group families) or weekly check-in phone calls/texts (for control group families). Families received $50 incentives per research visit.

Participants

Nearly all caregivers were mothers ($n = 88$), with only one father. The mean age of caregivers was 32.75 years ($SD = 6.18$), and the mean age of children was 23.33 months ($SD = 9.88$). There were 50 male children (56.2%) and 39 female children (43.8%) included in the sample. Most caregivers identified as Hispanic/Latino(a) (78.7%; $n = 70$), and 83.15% of caregivers ($n = 74$) identified their race as Caucasian/White. Caregivers also indicated their preferred language (either English or Spanish). Both assessments and the following intervention sessions were administered using their preferred language. There were 68.5% ($n = 61$) families completing the assessments in Spanish, and the rest in English ($n = 28$). The average annual family income in the sample was $30,425.65 ($SD = $21,947.58$).

FIND Intervention Implementation

The FIND intervention spanned 10 weekly home visit meetings that alternated between video recording sessions and coaching sessions. Throughout the 10 weeks, coaches worked closely with the caregiver to improve on specific aspects of positive parent-child interactions. During filming sessions, coaches recorded a 10-minute film of the caregiver and their infant interacting during routine, everyday activities. Between recording and coaching sessions, videos

were edited either by the coach or a trained editor to show brief clips in which the caregiver uses one or more elements of the previously listed five elements. These clips were chosen to facilitate learning and optimize engagement. While watching the clip, the coach provided their analysis and praises for the caregiver. Coaches were trained to maximize the salience of the different elements, reduce the reliance on professional expertise, and facilitate consistent implementation of the elements. The entirety of the recording and review processes were repeated for every element, meaning the coach spent two weeks on each of the five elements. Measures

PSOC

The PSOC is a self-report scale developed by Gibaud-Wallston and Wandersman (1978) that is commonly used to assess parenting self-esteem (Ohan et al., 2000). The measure divides these items into two scales: *Satisfaction* (liking of the parenting role) and *Efficacy* (perceived competence in the parenting role; Johnston & Mash, 1789). Each item is answered using a six-point Likert-scale that ranges from 1 (strongly agree to 6 (strongly disagree). Scoring for some items are reversed so that higher scores indicate greater caregiver self-esteem for all items. The reported alpha coefficients for the satisfaction and efficacy scales are .82 and .70, respectively. Six-week test-retest correlations were reported to have ranged from .46 to .82 (Gibaud-Wallston & Wandersman, 1978).

PSI

The short form of the PSI was used. The PSI short form is a questionnaire derived from the original 120-item PSI measure. Like the long form, the short form PSI is designed to measure stress in the parent-child relationship and identify families that may need further services (Abidin, 1995). These items are divided into three subscales: *Parental Distress* (level of distress related to conflicts with a partner, social support, and stresses resulting from life restrictions due

9

to child rearing), *Parent-Child Dysfunctional Interaction* (dissatisfaction about interaction with their infant and perception of their infant compared to other infants/children), and *Difficult Child* (perception of their infant's self-regulatory abilities such as temperament or defiance; Reitman et al., 2002). Each subscale consists of items that are rated on a five-point Likert-scale from 1 (strongly disagree) to 5 (strongly agree). Higher scores on the scale indicate greater levels of caregiver stress. Abidin (1995) reported alpha coefficients of .91 for the total scale, .87 for the *Parental Distress* subscale, .80 for the *Parent-Child Dysfunctional Interaction* subscale, and .85 for the *Difficult Child* subscale, with tests-retest correlations after 6 months ranging from .68 to .85.

SEPTI-TS

The SEPTI-TS Short Form is a 26-item questionnaire derived from the original SEPTI that included 53 items. Both measures were designed to assess parental self-efficacy in the caregivers of toddlers (Meunier & Roskam, 2009). These items are divided into four subscales: *Nurturance* (expressing loving and caring feelings towards the child and responding empathically, *Discipline* (setting limits for a child), *Play* (getting involved in child's play), and *Routine* (establishing structure and routine in a child's daily activities, eating, and sleeping) (Rijen et al., 2013). The items are all rated on a six-point Likert-scale with response categories varying from 1 (strongly disagree) to 6 (strongly agree). Higher scores indicate greater parental self-efficacy. Rijen et al. (2013) reported alpha coefficients of .88 for the total scale, .82 for the *Nurturance* subscale, .79 for the *Discipline* subscale, .80 for the *Play* subscale, and .79 for the *Routine* subscale. Test-retest correlations were not reported in the study. Data AnalysisTwo-way ANOVAs were conducted to examine the pre- and post-intervention differences in caregivers' parenting sense of competence (PSOC), parental stress (PSI), and parental self-efficacy (SEPTI)

and if these differences were significantly influenced by the condition (i.e., FIND-intervention or active control) or preferred administration language (i.e., English or Spanish). The two-way ANOVAs were also used to examine if there was a significant interaction between these two variables. Residual analysis was performed to test for the assumptions of the two-way ANOVA. Outliers were assessed by inspection of a boxplot, normality was assessed using Shapiro-Wilk's normality test, and homogeneity of variances was assessed by Levene's test.

Results

PSOC Analysis

A two-way ANOVA was conducted to examine the effects of the condition (FIND or control) and primary language spoken in the household (English or Spanish) on the change in PSOC scores from pre- to post-intervention. Figure 2 shows that families in the FIND condition significantly increased PSOC scores across the intervention period compared to control families. The interaction between administration language (english/spanish) and condition (FIND/control) was not significant. There was one outlier, as assessed as being greater than 3 times the interquartile range from the edge of the box in a boxplot. However, the outlier was not excluded from the analysis as it did not have an effect on the significance of the results. Residuals were normally distributed ($W = .989$, $p = .652$) and there was homogeneity of variances ($F(3, 84) = .361$, $p = .781$).

The interaction between the condition and primary language spoken in the household on the change in PSOC scores from pre- to post-intervention was non-significant, $F(1, 84) = 1.171$, $p = .282$, partial $\eta^2 = .013$ (See *Figure 2*). There was a greater change in PSOC scores from pre- to post-intervention in families in the FIND condition ($M = 2.04$, $SD = 4.20$) compared to families in the control condition ($M = .028$, $SD = 3.97$), a statistically significant difference ($M_{diff} = 1.972$, $F(1, 84)=6.22$, $p = 0.015$). However, there was no significant difference in change in PSOC scores for families that primarily spoke English ($M = .963$, $SD = 4.31$) and families that primarily spoke Spanish ($M = 1.33$, $SD = 4.18$; $M_{diff} = .367$, $F(1,84) = 0.34$, $p = .563$).

PSI Analysis

A two-way ANOVA was conducted to examine the effects of the condition (FIND or control) and primary language spoken in the household (English or Spanish) on the change in PSI scores from pre- to post-intervention. Figure 3 shows that Spanish-speaking families significantly decreased PSI scores across the intervention period compared to English-speaking families, without considering condition. The interaction between administration language (english/spanish) and condition (FIND/control). There was one outlier, as assessed as being greater than 3 times the interquartile range from the edge of the box in a boxplot. Residuals were normally distributed ($W = .976$, $p = .102$) and there was homogeneity of variances ($F(3, 83) = 1.52$, $p = .216$).

The interaction between the condition and primary language spoken in the household on the change in PSI scores from pre- to post-intervention was found to be non-significant, $F(1, 83) = 2.01$, $p = .160$, partial $\eta^2 = .022$ (See *Table 1*). There was no significant difference in the change in PSI scores from pre- to post-intervention in families in the FIND condition ($M = -6.39$, $SD = 15.87$) compared to families in the control condition ($M = -7.82$, $SD = 16.75$; $M_{diff} = 1.44$, $F(1,83) = .953$). However, there was a statistically significant difference ($M_{diff} = -9.50$, $F(1, 83) = 5.36$, $p = .023$) between the decrease in PSI scores for families that primarily spoke English ($M = -.786$, $SD = 13.71$) and families that primarily spoke Spanish ($M = -10.29$, $SD = 16.65$),

SEPTI-TS Analysis

A two-way ANOVA was conducted to examine the effects of the condition (FIND or EHS-control) and primary language spoken in the household (English or Spanish) on the change in SEPTI-TS scores from pre- to post-intervention. Marginal means of caregivers' change in PSOC scores are presented in Figure 4. There were no outliers, and all assumptions were met.

13

Residuals were normally distributed ($W = .99$, $p = .863$) and there was homogeneity of variances ($F(3, 63) = .85$, $p = .469$).

The interaction between the condition and primary language spoken in the household on the change in SEPTI-TS scores from pre- to post-intervention was found to be non-significant, $F(1, 63) = 1.43$, $p = .236$, partial $\eta^2 = .022$. Though this effect was non-significant, it seems that the FIND intervention may have had a different effect on the change in SEPTI-TS scores for English- and Spanish-speaking families in the FIND intervention condition. In this condition, we saw a non-significant but greater increase in SEPTI-TS scores for Spanish-speaking families, $M = 13.73$, 95% CI = [6.04,21.42], than English-speaking families, $M = 2.67$, 95% CI = [-8.65, 13.98]. See Figure 4.

There was no significant difference in the change in SEPTI-TS scores from pre- to post-intervention in families in the FIND condition ($M = 10.24$, $SD = 18.26$) compared to families in the control condition ($M = 6.138$, $SD = 21.49$; $M_{diff} = 4.10$, $F(1,63) = .118$, $p = .732$). There was also no significant difference in change in SEPTI-TS scores for families that primarily spoke English ($M = 4.57$, $SD = 18.99$) and families that primarily spoke Spanish ($M = 10.24$, $SD = 19.93$; $M_{diff} = 5.67$, $F(1, 63) = .856$, $p = .358$).

Discussion

FIND is a brief and strength-based video coaching intervention program designed to promote developmentally supportive caregiving (Fisher et al., 2016). The intervention has previously been found to improve caregiver and child outcomes (Liu et al., 2021; Joseph et al., 2022; Nese et al., 2016), however, no study has examined if these intervention effects are influenced by cultural factors. This study looked to fill this gap by exploring if the effectiveness of the FIND intervention on parental sense of competence, parental stress, and parental self-efficacy is influenced by whether the intervention is administered to Spanish-speaking or English-speaking families.

Contrary to our prediction, analyses showed that FIND did not have a significant main effect on PSI or SEPTI-TS scores, but did have a significant effect on PSOC scores. The increase in PSOC scores from families in the FIND condition was significantly greater compared to families in the control condition. Our findings suggest that the FIND intervention was not as effective in improving all of the target outcomes for this particular sample, especially when compared to what past studies have found. This also could suggest that cultural factors may play an impactful role in the effectiveness of the FIND intervention, and that more work needs to be done to explore this possibility. This finding could be attributed to the predominantly Spanish-speaking and Hispanic sample used in the study. If there are differences in FIND intervention results when using a predominantly Hispanic sample, adjustments must be made to ensure that the FIND intervention is not only effective for White and/or acculturated populations.

Preferred administration language did not have a significant main effect on PSOC or SEPT-TS scores, but did have a significant effect on PSI scores. The decrease in PSI scores was significantly greater for Spanish-speaking families than English-speaking families. This could

suggest that cultural factors play an important role in parental outcomes, but it is also possible that this this importance is something we would see with all intervention outcomes. Contrary to our hypothesis, we did not find an interaction between condition and preferred administration language for the PSOC, PSI, or SEPTI-TS measures. This could mean that any form of services or interventions provided to Hispanic families can drastically improve caregiver stress. This may also just be a result of two effective programs (EHS and FIND) in targeting at-risk families and caregivers.

Limitations & Future Directions

This study had several limitations. First, preferred administration language is not a holistic measure of acculturation, and does not capture all cultural factors that may influence the FIND intervention effects. The present study used existing data that was collected as a part of a larger study that was not focused on cultural factors. Future studies should focus specifically on looking at cultural factors and include more measures of acculturation and cultural differences in Hispanic vs. non-Hispanic families (i.e., Marin & Gamba, 1996). Second, all measures of caregiver outcomes were self-report surveys, which may generate reporter biases. Future studies should use observational measures perhaps in addition to self-report measures on parenting practices. Third, the sample sizes included were unbalanced, with significantly more Spanish-speaking families than English-speaking families participating in the study. Future studies should continue to use randomized controlled trials but with balanced, larger sample sizes. Lastly, the current sample had a relatively high attrition rate. Future studies should try to prevent a high attrition rate perhaps by increasing incentives or reducing the workload of the participants. Additionally, future studies could intentionally collect data on attrition rates and reasons for

families dropping out of the intervention. This may also provide insight into whether certain cultural groups respond better to the intervention more than others.

Future studies should also expand to include participants of different cultural minorities that are commonly seen where the FIND intervention would be most readily used. Though there were no significant interactions found in the present study, future studies should use a more robust measure of acculturation or cultural differences to ensure that the FIND intervention is effective for families of any race or culture.

The present study provides perhaps presents more questions than it answers. Though we did not find what we expected, our findings hint that cultural considerations are important to consider when implementing the FIND intervention. The FIND intervention has already been found to be associated with a variety of positive outcomes. However, it is important that we ensure that families of all cultures would benefit from the implementation of the FIND intervention. As the United States population continues to grow more racially and ethnically diverse (Jensen et al., 2021), it will be more important than ever to better understand the relationship between cultural factors and the FIND intervention's effectiveness.

List of Figures

Figure 1

Flowchart of Participants and Procedures
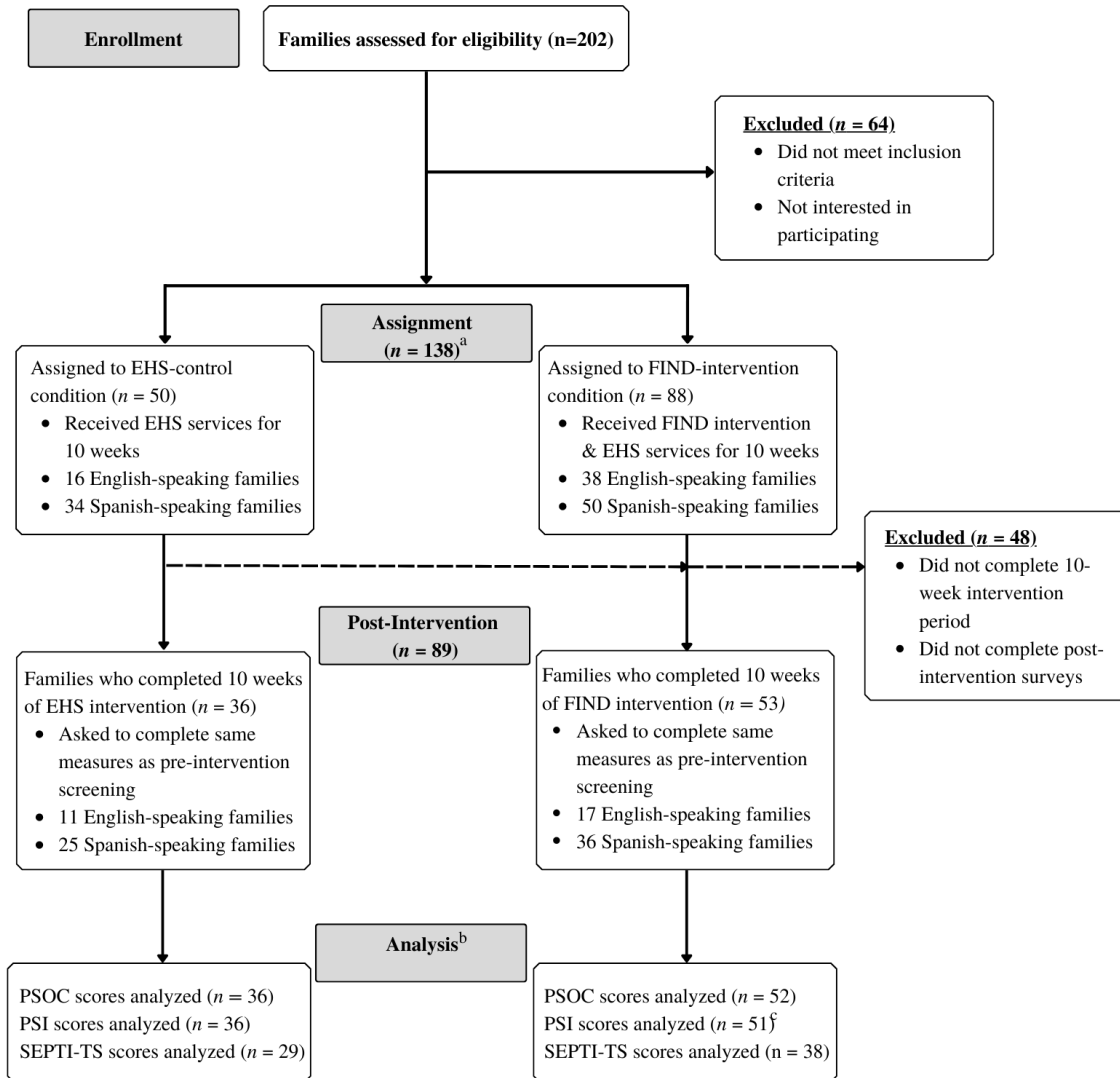
Note. [a]All participating caregivers completed self-report surveys including the PSOC, PSI, and SEPTI-TS both before allocation and after the ten-week intervention period. [b]One caregiver did not complete the PSI at post-intervention, one caregiver did not complete the PSOC at post-intervention, and twenty-one caregivers did not complete the SEPTI-TS assessment either at screening, post-intervention, or both. [c]One outlier was identified and later removed from the analysis of PSI Scores (see Methods section).

Figure 2

Estimated Marginal Means of Change in PSOC Scores
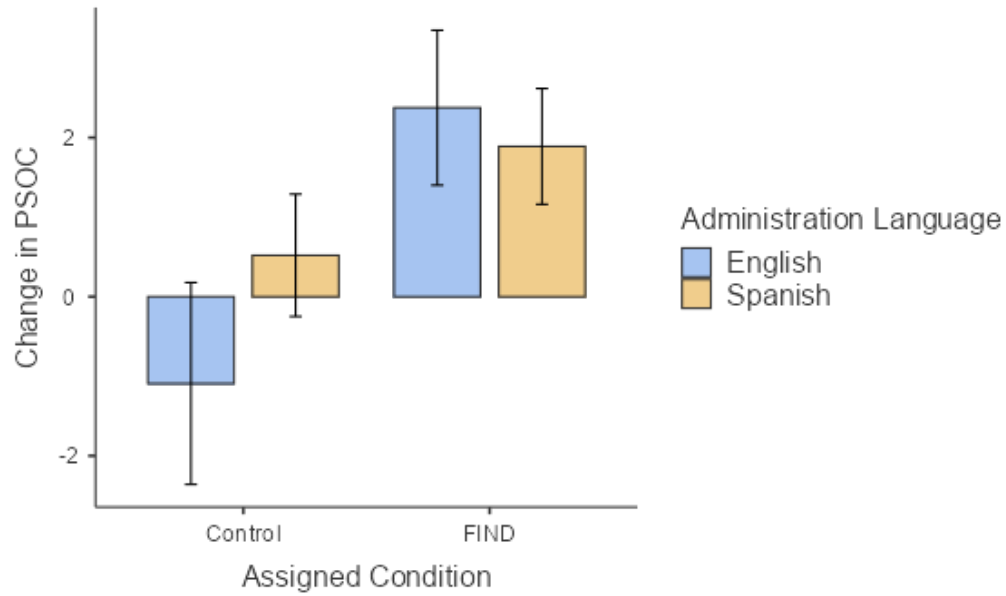
Note. Error bars indicate the 95% confidence interval. Among families in the FIND condition (right two bars), analyses showed a significantly greater increase in PSOC scores compared to families in the control condition (left two bars). No significant interaction effect was found.

Figure 3

Estimated Marginal Means of Change in PSI Scores
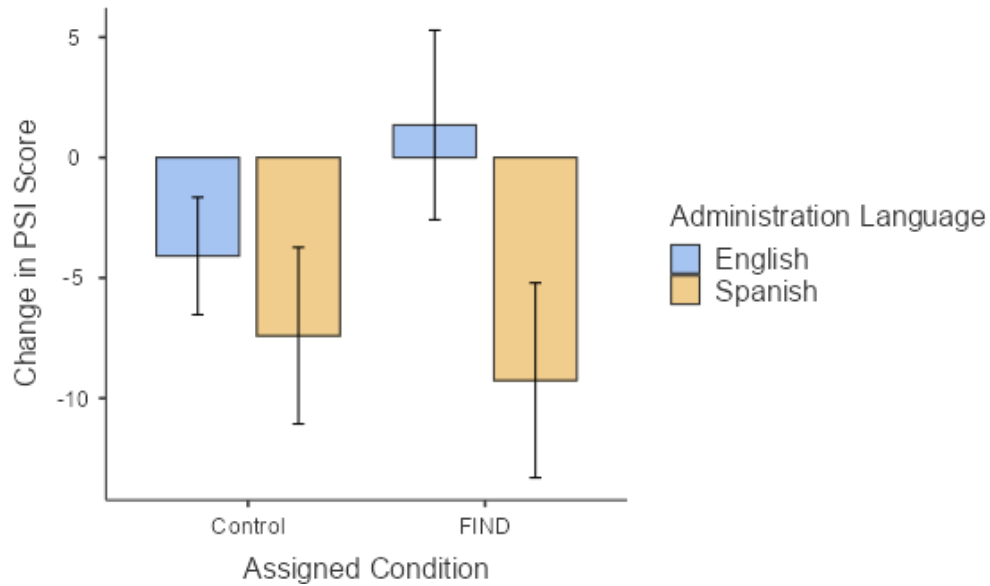
Note. Error bars indicate the 95% confidence interval. Among Spanish-speaking families (orange bars), analyses showed a significantly greater decrease in PSI scores compared to English-speaking families (blue bars). No significant interaction effect was found.

Figure 4

Estimated Marginal Means of Change in SEPTI-TS Scores
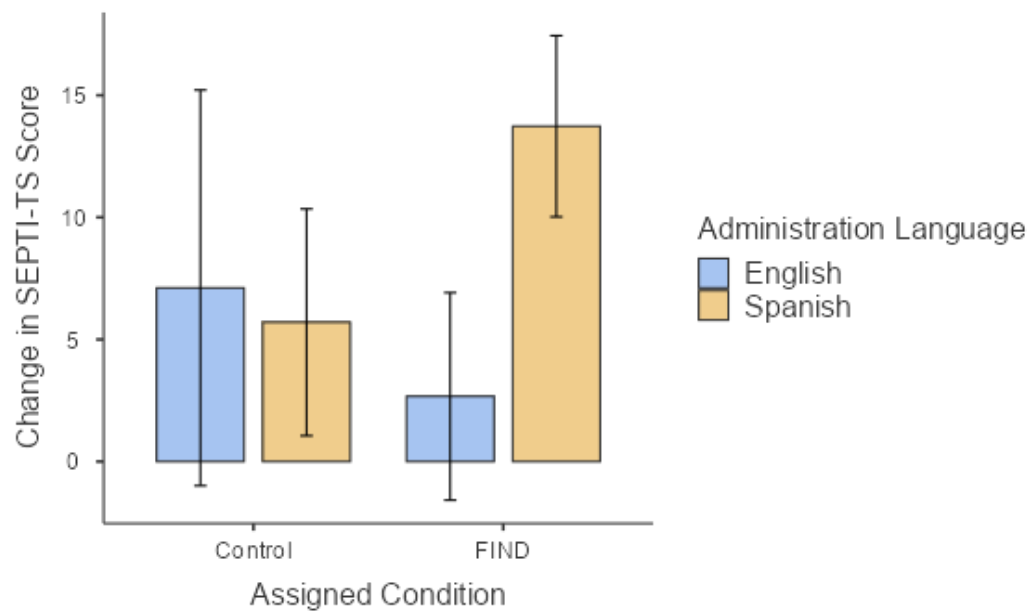
Note. Error bars indicate the 95% confidence interval. No significant main effects or interaction effects were found.

List of Tables

Table 1

ANOVA Summary Table for Change in PSOC, PSI, and SEPTI-TS Scores

		df_1, df_2	MS	F	p
PSOC	Assigned Condition*	1, 84	105.68	6.22	.015
	Administration Language	1, 84	5.72	5.72	.563
	Assigned Condition x Administration Language	1, 84	19.88	1.17	.282
PSI	Assigned Condition	1, 83	0.85	0.00	.953
	Administration Language*	1, 83	1330.24	5.36	.023
	Assigned Condition x Administration Language	1, 83	498.85	2.01	.160
SEPTI-TS	Assigned Condition	1, 63	45.46	0.12	.732
	Administration Language	1, 63	329.37	0.87	.358
	Assigned Condition x Administration Language	1, 63	550.12	1.43	.236

Note. MS = mean squares, df = degrees of freedom, *p < .05.

References

Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., & Juffer, F. (2003). Less is more: Meta-analyses of sensitivity and attachment interventions in early childhood. *Psychological Bulletin*, *129*(2), 195–215. https://doi.org/10.1037/0033-2909.129.2.195

Balldin, S., Fisher, P. A., & Wirtberg, I. (2016). Video feedback intervention with children. *Research on Social Work Practice*, *28*(6), 682–695. https://doi.org/10.1177/1049731516671809

Boruszak-Kiziukiewicz, J., & Kmita, G. (2020). Parenting self-efficacy in immigrant families—a systematic review. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.00985

Calzada, E. J. (2010). Bringing culture into parent training with Latinos. *Cognitive and Behavioral Practice*, *17*(2), 167–175. https://doi.org/10.1016/j.cbpra.2010.01.003

Dozier, M., & Bernard, K. (2017). Attachment and biobehavioral catch-up: Addressing the needs of infants and toddlers exposed to inadequate or problematic caregiving. *Current Opinion in Psychology*, *15*, 111–117. https://doi.org/10.1016/j.copsyc.2017.03.003

Fisher, P. A., Frenkel, T. I., Noll, L. K., Berry, M., & Yockelson, M. (2016). Promoting healthy child development via a two-generation translational neuroscience framework: The filming interactions to nurture development video coaching program. *Child Development Perspectives*, *10*(4), 251–256. https://doi.org/10.1111/cdep.12195

Fisher, P. A., Stoolmiller, M., Gunnar, M. R., & Burraston, B. O. (2007). Effects of a therapeutic intervention for foster preschoolers on diurnal cortisol activity. *Psychoneuroendocrinology*, *32*(8-10), 892–905. https://doi.org/10.1016/j.psyneuen.2007.06.008

Flannery, J. E., Beauchamp, K. G., & Fisher, P. A. (2016). The role of social buffering on chronic disruptions in quality of care: Evidence from caregiver-based interventions in foster children. *Social Neuroscience*, *12*(1), 86–91. https://doi.org/10.1080/17470919.2016.1170725

Fukkink, R. G. (2008). Video feedback in widescreen: A meta-analysis of family programs. *Clinical Psychology Review*, *28*(6), 904–916. https://doi.org/10.1016/j.cpr.2008.01.003

Gallo, L. C., Penedo, F. J., Espinosa de los Monteros, K., & Arguelles, W. (2009). Resiliency in the face of disadvantage: Do Hispanic cultural characteristics protect health outcomes? *Journal of Personality*, *77*(6), 1707–1746. https://doi.org/10.1111/j.1467-6494.2009.00598.x

Giuliani, N. R., Beauchamp, K. G., Noll, L. K., & Fisher, P. A. (2019). A preliminary study investigating neurocognitive mechanisms underlying a child-supportive parenting

intervention. *Frontiers in Behavioral Neuroscience*, *13*. https://doi.org/10.3389/fnbeh.2019.00016

Guerry, J. D., & Hastings, P. D. (2011). In search of HPA axis dysregulation in child and adolescent depression. *Clinical Child and Family Psychology Review*, *14*(2), 135–160. https://doi.org/10.1007/s10567-011-0084-5

Guilliams, T. G., & Edwards, L. (2010). Chronic stress and the HPA axis: Clinical assessment and therapeutic considerations. *The Standard: A Review of Natural & Nutraceutical Therapies for Clinical Practice*, *9*(2).

Jensen, E., Jones, N., Rabe, M., Pratt, B., Medina, L., Orozco, K., & Spell, L. (2021). *The Chance That Two People Chosen at Random Are of Different Race or Ethnicity Groups Has Increased Since 2010*. United States Census Bureau.

Joseph, G. E., Soderberg, J., Abbott, R., Garzon, R., & Scott, C. (2022). Improving language support for infants and toddlers: Results of FIND coaching in childcare. *Infants & Young Children*, *35*(2), 91–105. https://doi.org/10.1097/iyc.0000000000000214

Kang, S.-M. (2006). Measurement of acculturation, scale formats, and language competence. *Journal of Cross-Cultural Psychology*, *37*(6), 669–693. https://doi.org/10.1177/0022022106292077

Kim, J.-S. (2018). Social support, acculturation stress, and parenting stress among marriage-migrant women. *Archives of Psychiatric Nursing*, *32*(6), 809–814. https://doi.org/10.1016/j.apnu.2018.06.004

Lee, S. J., Gopalan, G., & Harrington, D. (2014). Validation of the Parenting Stress Index–Short Form with minority caregivers. *Research on Social Work Practice*, *26*(4), 429–440. https://doi.org/10.1177/1049731514554854

Lee, S., Nguyen, H. A., & Tsui, J. (2009). Interview language: A proxy measure for acculturation among Asian Americans in a population-based survey. *Journal of Immigrant and Minority Health*, *13*(2), 244–252. https://doi.org/10.1007/s10903-009-9278-z

Liu, S., Phu, T., Dominguez, A., Hurwich-Reiss, E., McGee, D., Watamura, S., & Fisher, P. (2021). Improving caregiver self-efficacy and children's behavioral outcomes via a brief strength-based video coaching intervention: Results from a randomized controlled trial. *Prevention Science*. https://doi.org/10.1007/s11121-021-01251-6

Lorenzo-Blanco, E. I., Meca, A., Unger, J. B., Romero, A., Gonzales-Backen, M., Piña-Watson, B., Cano, M. Á., Zamboanga, B. L., Des Rosiers, S. E., Soto, D. W., Villamar, J. A., Lizzi, K. M., Pattarroyo, M., & Schwartz, S. J. (2016). Latino parent acculturation stress: Longitudinal effects on family functioning and youth emotional and behavioral health. *Journal of Family Psychology*, *30*(8), 966–976. https://doi.org/10.1037/fam0000223

Marin, G., & Gamba, R. J. (1996). A new measurement of acculturation for Hispanics: The bidimensional acculturation scale for Hispanics (BAS). *Hispanic Journal of Behavioral Sciences*, *18*(3), 297–316. https://doi.org/10.1177/07399863960183002

Miao, S. W., Costigan, C. L., & MacDonald, S. W. S. (2018). Spillover of stress to Chinese Canadian immigrants' parenting: Impact of acculturation and parent–child stressors. *Asian American Journal of Psychology*, *9*(3), 190–199. https://doi.org/10.1037/aap0000105

Miller, E. B., Farkas, G., & Duncan, G. J. (2016). Does Head Start differentially benefit children with risks targeted by the program's service model? *Early Childhood Research Quarterly*, *34*, 1–12. https://doi.org/10.1016/j.ecresq.2015.08.001

Nese, R. N. T., Anderson, C. M., Ruppert, T., & Fisher, P. A. (2016). Effects of a video feedback parent training program during child welfare visitation. *Children and Youth Services Review*, *71*, 266–276. https://doi.org/10.1016/j.childyouth.2016.11.007

Ohan, J. L., Leung, D. W., & Johnston, C. (2000). The Parenting Sense of Competence scale: Evidence of a stable factor structure and validity. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, *32*(4), 251–261. https://doi.org/10.1037/h0087122

Pettit, G. S., Bates, J. E., & Dodge, K. A. (1997). Supportive parenting, ecological context, and children's adjustment: A seven-year longitudinal study. *Child Development*, *68*(5), 908–923. https://doi.org/10.1111/j.1467-8624.1997.tb01970.x

Quillian, L. (2012). Segregation and poverty concentration. *American Sociological Review*, *77*(3), 354–379. https://doi.org/10.1177/0003122412447793

Reitman, D., Currier, R. O., & Stickle, T. R. (2002). A critical evaluation of the Parenting Stress Index-Short Form (PSI-SF) in a Head Start population. *Journal of Clinical Child & Adolescent Psychology*, *31*(3), 384–392. https://doi.org/10.1207/s15374424jccp3103_10

Salisbury, M. R., Roos, L. E., Horn, S. R., Peake, S. J., & Fisher, P. A. (2022). The effectiveness of KEEP for families of children with developmental delays: Integrating FIND video coaching into parent management training—Oregon model: A randomized trial. *Prevention Science*. https://doi.org/10.1007/s11121-022-01344-w

Schindler, H. S., Fisher, P. A., & Shonkoff, J. P. (2017). From innovation to impact at scale: Lessons learned from a cluster of research-community partnerships. *Child Development*, *88*(5), 1435–1446. https://doi.org/10.1111/cdev.12904

Sheely-Moore, A. I., & Bratton, S. C. (2010). A strengths-based parenting intervention with low-income African American families. *Professional School Counseling*, *13*(3), 2156759X1001300. https://doi.org/10.1177/2156759x1001300305

Shonkoff, J. P., & Bales, S. N. (2011). Science does not speak for itself: Translating child development research for the public and its policymakers. *Child Development*, *82*(1), 17–32. https://doi.org/10.1111/j.1467-8624.2010.01538.x

Shrider, E. A., Kollar, M., Chen, F., & Semega, J. (2021). *Income and Poverty in the United States: 2020* (pp. 60–273). U.S. Census Bureau.

Thomson, M. D., & Hoffman-Goetz, L. (2009). Defining and measuring acculturation: A systematic review of public health studies with Hispanic populations in the United States. *Social Science & Medicine*, *69*(7), 983–991. https://doi.org/10.1016/j.socscimed.2009.05.011

van Rijen, E. H. M., Gasanova, N., Boonstra, A. M., & Huijding, J. (2013). Psychometric qualities of the Short Form of the Self-efficacy for Parenting Tasks Index-Toddler Scale. *Child Psychiatry & Human Development*. https://doi.org/10.1007/s10578-013-0414-6

SOCIAL CONNECTION AND FICTION:

THE POSSIBLE BENEFIT OF "INTERACTING" WITH FICTIONAL
CHARACTERS

by

BRINNA E. MAWHINNEY

A THESIS

Presented to the Department of Psychology
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

May 2022

An Abstract of the Thesis of

Brinna E. Mawhinney for the degree of Bachelor of Science
in the Department of Psychology to be taken June 2022

Title: Social Connection and Fiction: The Possible Benefit of "Interacting" with
Fictional Characters

Approved: _____*Dr. Sara D. Hodges*_____
Primary Thesis Advisor

This study addresses one role that fiction may play in people's lives – specifically,

providing social "interaction." Participants (265 University of Oregon students)

completed a writing task that involved writing about fictional characters and completed

measures of social fulfillment to see if that interaction may fulfill social needs and

alleviate loneliness. We hypothesized that higher transportation scores – the immersion

produced by the story as judged by an outside reader's perspective – would predict

lower participant loneliness scores for participants who are writing from the perspective

of a fictional character, to a fictional character, or their own journal entry. Furthermore,

we hypothesized that the media source of the fictional character chosen by the writer

will moderate this relationship, with written source media producing higher

transportation scores and lower loneliness scores than visual source media. Finally, we

hypothesized that participants who wrote more fiction or journaled outside of the

context of the study would write passages that earned higher transportation scores and

also report lower loneliness scores. Results indicated that coder-rated transportation

does not significantly predict a larger reduction in loneliness scores. Neither media type

nor participants' own writing outside of the study moderated the relationship between

transportation and change in loneliness. Results may have implications for developing a writing intervention to alleviate loneliness.

Keywords: Loneliness, transportation, fiction, fictional characters, writing

Acknowledgements

Table of Contents

List of Figures

List of Tables

Introduction

The first two years of the COVID-19 pandemic saw a large increase in self-identified loneliness paired with a massive increase in people interacting with media. With the COVID-19 pandemic's lockdowns and stay at home orders, researchers have found that during the first six months of the pandemic self- identified loneliness has significantly increased (Killgore, Cloonan, Taylor et al., 2020). Similarly, media interaction has also significantly increased. While Netflix added 28 million subscribers in all of 2019, they added 26 million new subscribers in just the first six months of the pandemic (Vlassis, 2021). This phenomenon has not just been limited to movies and television: Overdrive, a digital distributor of online content, including ebooks, reports that weekly library ebook lending saw an increase by 50% in the same first six months of the pandemic (Guren, McIlroy, Siek et al., 2021). The simultaneous increases in loneliness and increases in media consumption suggest that media may serve as a sort of social proxy in people's lives.

With pandemic lockdowns, people have also been interacting with fiction in a more immersive way than consuming television or books. People have also been writing and reading fan written narratives. Fanfiction has been a popular way for many people to interact with fictional characters, even before the Covid-19 pandemic. Many people spend their free time writing fictional narratives about their favorite fictional characters, while also interacting with an online community of fans that support and read their fanfiction. Archive of Our Own (Archive of Our Own, 2020), a popular fanfiction website, supported over 6 million unique works of fanfiction, with 2.5 million registered users, as of 2020 (Archive of Our Own, 2020). Furthermore, Archive of Our

Own reported increased comments on fanfiction works and overall number of written fanfictions when lockdowns began to happen in March 2020. As loneliness has increased during the pandemic, more people have been turning to entertainment, including fanfiction. This could indicate a connection between interacting with fictional characters, especially in written narratives like fanfiction, and loneliness.

Although the popularity of fanfiction websites indicates that many people are interacting with fictional characters through writing and could indicate a connection between fanfiction use and loneliness, there has been limited research on the connection between loneliness and actively interacting with fictional characters through writing. Writing about fictional characters could reduce loneliness. The current rise in loneliness increases the importance of studying the potential effects writing can have on well-being. The current study aims to understand how interacting with fictional characters through writing may alleviate loneliness.

Current research reveals that there is a gap in understanding how writing can influence people's emotions in ways that are different from the influence of reading or watching fiction in the form of films or plays. Although there is support for the idea that *consuming* fiction can change emotions, few have included *writing* in their analyses (Bal & Veltkamp, 2013; Djicik, Oatley, Zoeterman et al., 2009; Mar & Oatley, 2008).

Many studies have concluded that interacting with fiction can change emotional states and self-concept. Bal and Veltkamp (2013) found that reading fiction (in this case, an excerpt of a Sherlock Holmes short story) compared to informative reading (about riots in Libya and the nuclear disaster in Japan), could increase a participant's level of empathic concern. In addition, Djikic et al. (2009) found that when participants

2

read *The Lady with The Toy Dog* by Chekhov, they had a significantly greater change in self-identified personality traits on the Big Five Inventory than when they read a comparison text that had the same information in documentary form. Despite both studies providing support for fiction evoking changes, there needs to be caution in the assumption that fiction could create permanent change. A more cautious route would be to interpret the changes as a temporary shift (Dijikic et al., 2009). However, even temporary change might still support at least a momentary lapse in loneliness through interacting with fiction.

Gabriel and colleagues (2017) found that thinking about fictional characters filled social needs for participants who had experienced trauma. This finding, coupled with prior research that has established that interacting with fictional characters through reading or watching media can change emotional states, supports the idea that interacting with fictional characters through writing could also influence how writers feel.

Research has found that fiction can also change one's sense of self. Derrick et al. (2008) found that people with low self-esteem can use parasocial relationships – one-sided relationships with media persona that develop through a sense of shared experiences and a sense of knowing the person (Derrick et al., 2008) – to move towards developing attributes that participants used to describe an ideal self. Similarly, Sestir and Green (2010) found that exposure to media characters can temporarily change the self- concept. They used a "Me/Not-Me" reaction time task that tested how fast participants responded to a list of traits as either descriptive of the self or not. This task allowed the researchers to gain a better understanding of how identification with a

character affects an individual's sense of self. However, Sestir and Green (2010) found that an essential component to experiencing a change in self-concept was participants' immersion in the piece of fiction. This immersion in the fiction, also known as transportation, moderated the relationship between fiction and change in self-concept.

Transportation

Other studies along with that by Sestir and Green (2010) have also indicated that in order for fiction to produce a change of self, transportation needs to be present (Bal & Veltkamp 2013; Mar, Oatley, Hirsch et al., 2006). According to Green and Brock (2000), transportation refers to the experience of a reader being cognitively, emotionally, and imaginatively immersed in a narrative world. As readers become more immersed in the story, the narrative subsequently becomes more "real," both emotionally and cognitively (Sestir & Green, 2010). Sestir and Green (2010) also found that the higher level of transportation participants experienced, the more likely they were to switch from initially not identifying with character traits to later self-identifying with those traits.

For the current study, we adapted the concept of transportation in a way that has not been represented in prior research. Due to the written component of our study, and the lack of research regarding transportation and writing, we aimed to analyze the transportive qualities of participant written narratives. Our conception of transportation analyzes the overall immersion *produced* by a writer's story as judged by an outside reader's perspective. This contrasts with other studies that include transportation as a variable and generally are interested in how transportation affects readers, not how transportive qualities of writing affect the writer. We aimed to operationalize a

participant's ability to write a transportive story rather than what prior research has analyzed, which was the reader's experience of transportation. A high amount of writer immersion in the story would seem to be needed to communicate a highly transportive narrative. Therefore, our concept of transportation may potentially be conceptually related with how prior research has handled the reader's experience of transportation.

Loneliness and Perceived Realness

Researchers have deemed social relationships as vital to mental and emotional health. Life without social relationships is related to significant psychological distress and poor health outcomes (Heinrich & Gullone, 2006). These negative outcomes highlight the need for a comprehensive understanding of loneliness and ways to combat it. Heinrich and Gullone (2006) identify loneliness as not only an unpleasant emotional experience, but that it also includes a cognitive aspect where people perceive a discrepancy between social relationships that they wish to have versus what they perceive they do have.

Although researchers have extensively studied loneliness in the context of social relationships with other people, there has been little research on loneliness in the context of non-traditional means of social fulfillment, which could include interacting with fictional characters. However, Paravati and colleagues (2020) examined the role of traditional and nontraditional means of social fulfillment in regards to life satisfaction. They created a new measure, the "Social Fuel Tank," to analyze the different means of social fulfillment. They included both traditional means of social fulfillment, such as family, friends, and romantic partner, and non-traditional means of social fulfillment, including gaming, reading books, and watching TV. They found that both traditional

and nontraditional means of social fulfillment uniquely predicted feeling less lonely. This study illuminates how non-traditional means of social fulfillment, potentially including writing, may be essential to understanding loneliness.

As one example of a non-traditional means of social fulfillment, studies have examined how fictional characters could assume a role in people's social relationships. Fictional characters encompass a one-way connection, but people can have distinct emotional experiences when interacting with them (McGee, 2005). These real experiences with fictional characters can cover a variety of emotions, including distress at the loss of a fictional character that is similar to losing a real social relationship (Cohen, 2004). Another study by Gardner and Knowles (2008) found that undergraduate students perceived their favorite television character as more real than non-favorite characters. These studies provide evidence of the perceived realness of fictional characters show the blurry line between real and fictional relationships.

In the context of writing, Taylor and colleagues (2003) examined how fiction writers experienced the *illusion of independent agency* (IIA). IIA encompasses how many fiction writers experience their characters as having their own independent thoughts and actions. The researchers found that 92% of authors in the sample they interviewed had at least some experience of IIA and that writers scored higher on measures of empathy and dissociation than population norms. The experience of IIA, along with higher dissociative tendencies that include the ability to become engrossed in activities, could contribute to an author obtaining social fulfillment through a fictional character.

The current study

      Thus, the pairing of perceived realness and the ability of non-traditional means of social fulfillment to alleviate loneliness lend support to the idea that interacting with fictional characters could provide social fulfillment. Drawing from a larger study on this general topic (Doyle, 2021), the current study aimed to dive deeper into the relationship between writing and loneliness to examine the levels of transportation in writing and loneliness. Unlike previous studies that have established a connection between *consuming* fiction and social fulfillment, the current study focuses on analyzing the level of transportation produced by participants' written narratives and its relationship to their loneliness. Furthermore, we hope to establish how the type of media that someone's favorite fictional character comes from, and how the amount of time someone spends writing fiction outside of the context of the current study, affect the relationship between transportation and loneliness.

Research Questions

RQ1: Does amount of transportation produced by a piece of writing about a fictional character predict loneliness scores?

> **Hypothesis 1:** Participants who produce writing about a fictional character with higher transportation scores will report less loneliness. This will be the case whether the participants write **FROM** the perspective of the fictional character, write **TO** the fictional character, or write a journal entry reflecting their own perspective, without any engagement of a fictional character, but after being asked to choose a favorite fictional character. (In the larger study from which the data for the current study are being drawn, Doyle (2021) already established that these between-subjects conditions do not have a main effect on loneliness.)

RQ2: Does the type of media that a fictional character comes from moderate the relationship between transportation and loneliness?

> **Hypothesis 2:** There will be a stronger relationship between transportation and loneliness for participants who write about characters that come from written media than there is between transportation and loneliness for participants who write about characters that come from visual media.

> Due to the written aspect of the study, this hypothesis was based in how written media may be better connected than visual media to a writing task. Furthermore, written media may also be more cognitively taxing in comparison to visual media. This could potentially limit people's ability to become

8

distracted and allow someone to become more immersed in writing, impacting

how much social satisfaction someone obtains from interacting with the fictional

characters.

RQ3: Does the amount of time participants spend writing (fiction or personal

journaling) on their own time moderate the relationship between transportation

and loneliness?

> **Hypothesis 3:** There will be a stronger relationship between
>
> transportation and loneliness in participants who habitually write fiction
>
> on their own time.

Method

The data for the current study come from a larger study (Doyle, 2021) that encompassed additional measures that are not included in the analyses below.

Participants

The participants of the study were 265 undergraduate students who were recruited from the Psychology and Linguistics Subject Pool at the University of Oregon in exchange for partial fulfillment of a class research requirement. Participant data collection occurred Spring quarter 2021 through Fall quarter 2021. Thus, all data collection occurred during the Covid-19 pandemic, during which participants experienced variability in lock downs and surges in Covid-19 cases. Before beginning the study, participants had no prior knowledge of the study or its goals. Participants were 18 to 36 years of age ($M = 19.7$, $SD = 2.4$), with 66.8% identifying as female (n = 177), 31.32% identifying as male (n = 83), 1.13% identifying as non-binary (n = 3), and two participants who chose not to answer the question. Of the participants, a majority were white at 59.24% (n = 157) and 91.32% (n = 242) were fluent in English with English as their first language. (In some cases, participants were missing some data and thus degrees of freedom in the analyses that follow fluctuate slightly.)

Procedure

After consenting to the online study (conducted via Qualtrics and Pavlovia), participants first completed some pre-manipulation measures of loneliness and mood (relevant measures are described below). Next, participants selected their favorite fictional character and were asked to identify the media source (e.g., book or television

show) that the character came from and the approximate year that the media with this character first came out. Participants were also asked to think about a person that they were very close to in their own life (e.g., a best friend or romantic partner). They were then asked to identify the person, by using first name, initials, or a nickname, and to specify what their relationship with the person is. They were also asked how long they had known the close other.

After their selection of these two individuals, participants completed a survey of traits (e.g., respectful, impulsive, childish, etc.). They were asked how well they thought each trait described themselves on a 5-point scale from "extremely inaccurate" to "extremely accurate." Participants also completed the Interpersonal Reactivity Index, a measure of self-reported empathy (Davis, 1983). However, neither of these measures were part of the present study.

Participants were randomly assigned to one of three writing prompts. They were asked to write 200 words as if writing *from* the perspective of the fictional character they earlier selected ("From" Group); as if writing a letter from themselves *to* the fictional character ("To" Group); or as a journal entry about their own life (Control Group). In all cases, the writing prompt asked participants to describe an instance of bad news, so that, for example, in the From Group, participants wrote from the perspective of the character they chose as the character received bad news; in the To Group participants wrote a letter from themselves to the character about bad news the character was receiving or had received; and in the Control Group participants wrote about a time that they received bad news. The choice of the "bad news" prompt was made as way to increase the chances that participants would make-up a new instance of bad news that

their chosen character had not already experienced in previous works. We wanted a prompt that would be compelling and unexpected, so that participants would need to put original thought into their writing. We chose to ask participants to write about bad news to try to elicit an emotional response, and left the prompt open ended so that participants could take the prompt in any direction they desired. Participants in the From Group and the To Group were asked to indicate how much the situation that they wrote about was similar to something experienced in previously produced media about the character, using a 5-point scale ranging from 1 = "not at all similar" to 5 = "almost exactly the same" (i.e., was the situation similar to something the character had experienced in their original story). Participants in the Control group were asked to rate on a 5-point scale how carefully they remembered their personal instance of receiving bad news, with 1 = "I don't remember it at all", and 5 = "I remember it very carefully, with great detail".

After the writing task, participants then completed the same loneliness and mood surveys, along with multiple other surveys that were not analyzed for the current study (see Appendix A for unrelated tasks and measures). Before participants finished the study, they answered questions about their personal writing practices outside of the context of the study, answered demographic questions, and completed other measures that were also not analyzed for this study. Participants were debriefed upon completion of the study.

Measures

Loneliness

The revised UCLA Loneliness Scale (Russel et al., 1980) was used to assess loneliness. This scale is made up of 20 items that include statements such as "I feel

isolated from others" and "I am unhappy being so withdrawn." Participants rated these statements on a 4-point scale that ranged from 1 to 4 (points were labelled never, rarely, sometimes, and often; higher scores indicated greater loneliness). The participants completed the loneliness scale twice --- once before and once after the writing prompt.

To analyze loneliness, our study utilized the loneliness change score for each participant, using post-writing minus pre-writing task loneliness scores. Thus, positive change scores indicated that a participant's loneliness increased during the study; negative change scores indicated that a participant's loneliness decreased during the study.

Past Creative Writing Experience

To assess participants' writing experience, participants answered four questions about how frequently they had engaged in creative fiction writing in the last year and prior to the last year; and how frequently they had engaged in journal or diary writing in the last year and prior to the last year. Answer options for these questions were "1 = No, never"; "2 = Yes, sometimes"; and "3 = Yes, often." In addition, participants were asked whether or not they had an imaginary companion as a child.

To analyze participants' writing experience, we summed each participant's score across all four writing habit questions. A score of 4 indicated that a participant did not participate in any form of habitual writing outside of the study and a score of 12 indicated that a participant had completed journal/diary and creative fiction writing often in the last year and prior to the last year. The mean on the scale was 6.44 ($sd =$ 1.73). There were some participants who reported never having written (i.e., score of 4); the highest reported score was 11.

Coding of Transportation

The current study also utilized a coding scheme completed by seven trained undergraduate research assistants (including the author) to further understand trends in the participants' written narratives (see Appendix B for full coding manual). Research assistants were trained to code each written narrative for transportation and quality. Transportation, in this context, measured outside readers' (i.e., the trained research assistants') feelings of transportation while reading the participants' writing samples. Research assistants used a 5-point Likert scale that ranged from 1 to 5, with 1 indicating no transportation and 5 indicating that the coder experienced a lot of transportation, whether cognitively, emotionally, and/or imaginatively, into the written narrative. Coder reliability for transportation indicated an alpha of $a = .79$ with a confidence interval ranging from .75-.83. Mean transportation across the narratives was 2.97 ($sd =$ .741).

Results

We tested our first hypothesis, that participants who produced writing that had higher transportation scores would report lower loneliness scores, with a linear regression. As stated previously, we utilized the loneliness change scores (post minus pre writing task). These change scores ($M = .006$, $sd = .213$) had a mean closer to zero and a small standard deviation, indicating that participants changed very little from pre to post writing task and exhibited little variability in loneliness. Our first hypothesis was not supported and transportation did not predict change in loneliness, $b = -.011$, $t(262) = -0.65$, $p = .517$. (Note that data points are jittered in the figures that follow.)
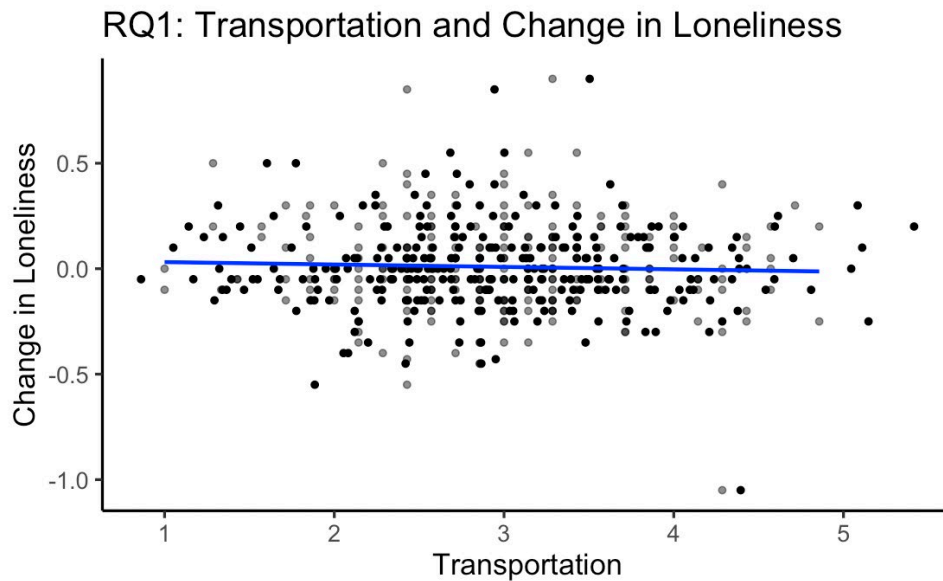
Figure 1. Transportation and Change in Loneliness

A nonsignificant relationship between transportation and change in loneliness

We tested our second hypothesis, that there would be a stronger relationship between transportation and change in loneliness if the participant chose their fictional

15

character from written media (as opposed to visual media). In order to test this, we ran a multiple regression predicting changes in loneliness from an interaction between transportation and type of media. The type of media variable used Books or Book Series as the reference category, so that all other media types (TV, Movies, Comic, Other) were in comparison to Books (see distribution of media types in Table 1). This hypothesis was not supported by our data. Transportation and media type did not significantly interact to predict change in loneliness in this model ($F(3, 260) = .2692$, $R^2 = .003$, $p = .848$). Transportation did not have a significant main effect on change in loneliness in this model ($b_1 = .005$, $t(260) = .084$, $p = .933$). Media type did not significantly predict change in loneliness in this model ($b_2 = .037$, $t(260) = .172$, $p = .864$). The interaction between transportation and media type also did not significantly predict change in loneliness ($b_3 = -.019$, $t(260) = -.290$, $p = .772$).

Books	TV	Movies	Comic	Other
26	146	68	6	19

Table 1: Number of Participants that Chose Each Media Type

RQ2: Change in Loneliness Predicted by Transportation and Media Type
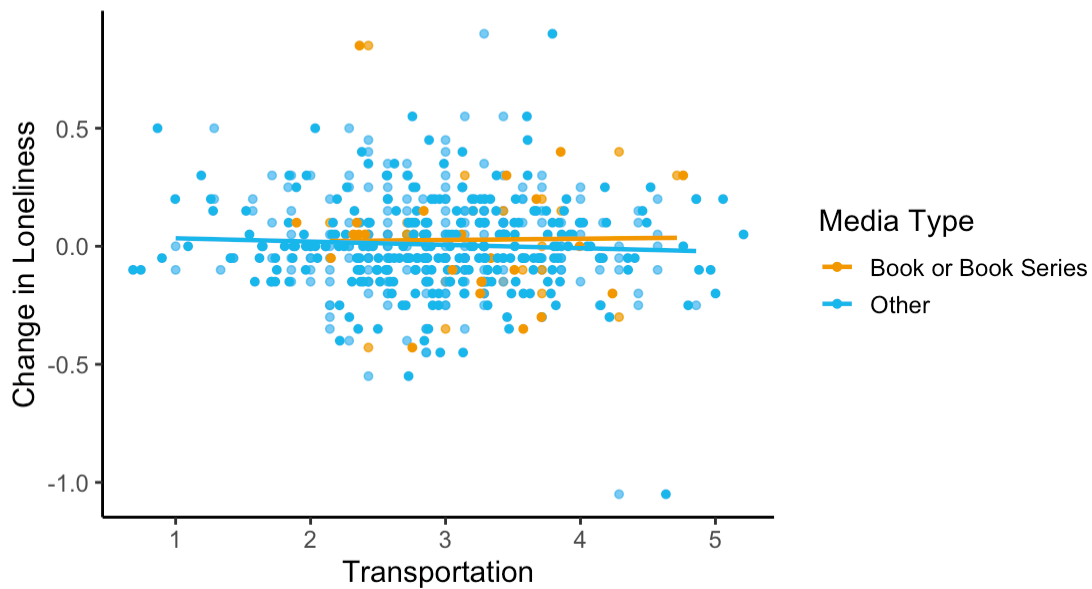
Figure 2. Change in Loneliness Predicted by Transportation and Media Type

No moderation by type of media on the relationship between transportation and change in loneliness

We tested our third hypothesis, that there would be a stronger relationship between transportation and loneliness if participants wrote outside of the context of the study, with a multiple regression. This hypothesis was not supported. Transportation and participants' own writing frequency did not significantly interact to predict change in loneliness ($F(3, 259) = .2691$, $R^2 = .003$, $p = .848$). Transportation had no significant main effect on loneliness in this model ($b_1 = -.041$, $t(259) = -.598$, $p = .55$). Participants' own writing frequency also did not significantly predict change in loneliness in this model ($b_2 = -.016$, $t(259) = -.488$, $p = .626$). The interaction between transportation and

17

participants' own writing also did not significantly predict change in loneliness (b_3 =

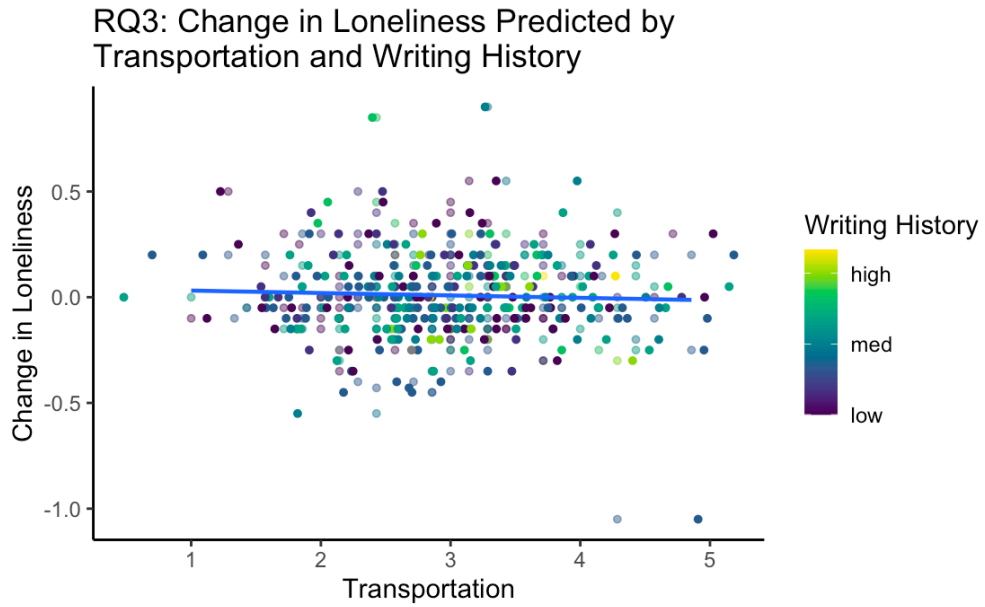.005, $t(259) = .435$, $p = .664$).

Figure 3. Change in Loneliness Predicted by Transportation and Writing History

 No moderation by participant's writing history on the relationship between

transportation and change in loneliness

Discussion

In this investigation of whether writing could help reduce loneliness, we found little support for our hypotheses. We did not find significant support for our first hypothesis that indicated participants who produced writing that had higher transportation scores would report lower loneliness scores (Hypothesis 1). We also did not find support for our other two hypotheses that indicated there would be a stronger relationship between transportation and loneliness if the participant chose their fictional character from written media (Hypothesis 2) or if the participant was someone who frequently wrote outside of the study (Hypothesis 3).

The results of the study did not support the idea that interacting with fictional characters through writing alleviates loneliness. This could indicate that there is no effect on loneliness from interacting with fictional characters. However, this study was limited in statistical power due to the sample size. Although the 265 participants may have been a large enough sample for our simple regression analysis that revealed no relationship between transportation and loneliness, it could be possible that our sample was not large enough to explore the interactions between variables tested here.

Furthermore, these hypotheses were exploratory. Previous research has been done on participant transportation (Bal & Veltkamp 2013; Mar, Oatley, Hirsch et al., 2006; Sestir & Green 2010), but not on the audience transportation that our study aimed to capture. We did not have participants rate their *own* transportation for their own pieces of writing; this was instead rated by trained undergraduate coders. Although our coders were reliable when coding our construct of transportation (a version of participant-rated transportation), a participant's own experience of transportation with

their own writing could have produced a different relationship between transportation

and loneliness. In the current study, all we can conclude is that there was no effect

between audience-perceived transportation and loneliness, but there is little to no other

previous research on this topic and further research would need to be conducted to make

firmer conclusions about this relationship.

Limitations and Implications for Future Research

Certain limitations of this study could be addressed in future research. For

example, from a world-wide perspective, this study utilized an overwhelmingly

Western, educated, industrialized, rich, and democratic (WEIRD) population that was

further constricted to university students who were majority white. Due to our study

participant limitations and the vast majority of research on engagement with fictional

characters being done in North America and Europe, it remains unclear whether the

study's findings would be generalizable to other populations.

Another possible influence on the results may have been the fact that

participants were writing in a context that was not purely for themselves. This could

have impacted how self-conscious participants may have felt while writing their

narratives, especially considering these were provided in the context of an hour-long

study and were accessible to researchers after participants finished. Participants may

have had the idea that someone was going to read their writing while completing the

written narrative which may have affected what they wrote, due to a desire for potential

audience approval or to entertain the audience. It may be that people find more benefit

from interacting with fictional characters through writing when they are only writing for

themselves (i.e., intrinsically motivate) and are afforded more privacy in their interactions.

Further analyses might also have yielded different results. Data analysis could have been broken down to multiple groups, such as comparing writing involving the fictional character to a participant's own journal/diary writing, to analyze the relationship between transportation and loneliness. This comparison may have proved fruitful; comparisons between writing groups is a central focus of the larger study from which these data were drawn.

This study was constricted to one hour, which limited the amount of time a participant could write for and did not allow the opportunity to test for more gradual or greater changes that may have been found if the study had been conducted over weeks or months. A future study on how loneliness could be impacted by writing may benefit from extending the duration of the study to weeks or months to have a fuller grasp on how habitual writing may influence loneliness.

The limited duration of the study also needs to be considered when looking at the loneliness measurement. The overall change in loneliness mean was close to zero ($M = .006$) and variability in change scores was low, indicating that most participants were experiencing very little change in loneliness. This could be due to the fact that participants completed the pre- and post-writing task loneliness measures within roughly 45 minutes of each other. This could have created pressure to provide similar answers for both surveys. Furthermore, loneliness may be too narrow of a construct to truly capture the benefits that people may experience from writing. We aimed to reduce the negative experience of loneliness within our study, but writing may create additional

positive experiences that were not captured through the narrow construct of loneliness. Previous research indicates that non-traditional means of social fulfillment, such as reading a book, can add depth and extra meaning to people's lives (Pavarti et al., 2020). Measures that capture additional positive benefits to people's lives, such as creative fulfillment, may be useful for future research to try to understand why many people write.

Despite our null results, this study has encouraged us to conduct future research that explores writing as a possible intervention for loneliness and how other factors may hinder or contribute to loneliness reduction. As previously mentioned, future research may consider lengthening the study itself to create habitual writing or may consider exploring workshops about writing activities that may help bolster loneliness reduction. Some of our null results may actually be seen as positive, in terms of developing writing as a means for reducing loneliness. Our second and third hypotheses indicate that character's media type and participant's own writing habits did not moderate the (lack of a) relationship between transportation and loneliness. This could indicate that it does not matter what media a fictional character is from for someone to receive benefits from interacting with that fictional character. In addition, a participant may not need to have their own writing habits outside of a writing intervention for a participant to receive the benefits of the writing intervention. These findings could have implications for future writing interventions and should be kept in mind when designing future research to try to reduce loneliness through interacting with fictional characters.

There is still work that needs to be done to fully unpack and understand the implications of immersive interactions with fictional characters, such as writing. Many

people associate writers with greater experiences of loneliness, but there are still questions around whether this is because writers write due to feelings of loneliness or if the (largely solitary act of) writing itself causes writers to be lonelier. Our research did not indicate that writing causes an increase in loneliness; however, more research is needed examining writers' emotional experiences. In addition, many people who write or interact with fictional characters do so in many more ways than simply writing alone. This leads to the conclusion that other variations on ways to "interact" with fictional characters may be explored. For example, fanfiction is a highly popular way to interact with fictional characters even after their franchises have ended. Many people form online communities around fanfiction and regularly interact through comment sections on websites, such as Archive of Our Own, but there is a currently limited understanding of how and why fanfiction is so popular. In addition, many people also participate in live action role play, cosplay, and other role play games where individuals actively take on the roles of fictional characters, whether through clothes, voices, or immersive experiences. In these contexts, not only do people interact with fictional characters by themselves, but many of these activities also include a group component that could provide fruitful insights into how and why people interact with fictional characters.

This research can be seen as a first step to understanding how interacting with fictional characters through writing affects loneliness, although the results suggest that there may not be a relationship between our primary predictor (audience-perceived transportation) and change in loneliness, and that this predictor is also not moderated by the character's media type and a person's own habitual writing. Despite these results, there still seems to be promise for the ability of writing to serve a purpose in people's

lives. This may not be realizable in loneliness reduction, but rather could be found in an added social/creative fulfillment or reduction in boredom, for example. Future research could lead towards creating a writing intervention to help ease loneliness by lengthening the study to form habitual writing, and by extension, potentially deeper relationships with characters. Overall, this study indicates that there needs to be more research on immersive fictional experiences and how fictional characters benefit people's lives.

Appendix A

Complete Larger Study Measures

Reaction Time – Self Traits

Participants responded either "Me" or "Not Me" to a series of 90 trait words and their reaction times were recorded using Pavlovia.

Trait Ratings

Participants rated the same 90 trait words for how well they describe themselves, the fictional character they chose at the beginning of the study, and for the real close other that they designated. Participants rated these traits for each of the three targets (self, character, and close other) on a 7-point scale that ranges from 1 = extremely inaccurate to 7 = extremely accurate. Before participants completed the writing intervention, they also rated another 45 of the trait words.

Measures:

Interpersonal Reactivity Index

To measure individual differences in self-reported empathy, participants completed the Interpersonal Reactivity Index (Davis, 1983) which consists of four subscales, each with seven items. Each subscale assesses a different facet of empathy, including perspective taking, empathic concern, personal distress, and fantasy (Davis, 1983). The index includes statements such as "I am often quite touched by things that I see happen" (empathic concern subscale) and "I really get involved with the feelings of characters in a novel" (fantasy subscale). Participants rated these statements using a 7-point scale that ranged from 1 ("does not describe me well") to 7 ("describes me very well").

Big Five Personality

To assess personality, embedded within the traits used for the reaction times were the traits from the Mini-Markers Scale (Saucier, 1994), which is a short measure of Big Five personality factors (Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism). Scores for the Mini-Markers were also collected during the trait ratings for the participant, the fictional character they chose, and their close other.

Mood

To assess participants' mood, participants completed the Positive and Negative Affect Schedule (PANAS; Watson et al., 1988). They indicated how they felt in the present moment by rating words, such as enthusiastic and ashamed, on a Likert scale ranging from 1 (very slightly or not at all) to 5 (extremely). Participants completed the PANAS both before and after the writing prompt.

Character Liking

To assess how the participants felt about the fictional character that they chose at the beginning of the study, participants rated statements about their character, including "I think the character is like an old friend" and "I find the character to be attractive/cute." These items were based on those used in Gardner and Knowles (2008). Participants rated these statements on a 5-point scale that ranged from strongly disagree to strongly agree.

Demographic Information

At the conclusion of the study, participants provided demographic information about their age, race and ethnicity, gender, major or intended major, and fluency in English.

Social Fuel Tank

Participants completed the Social Fuel Tank measure developed by Paravati et al. (2020). This measure assesses participants' strategies for social need fulfillment by having participants assign percentages to 17 possible strategies that they might use to fill their social needs that include items such as watching TV or eating favorite foods. For each strategy, they indicated how well they felt the strategy met their social needs, ranging from 0 to 100 percent. Percentages were intended to add up to no more than 100 percent; a percentage less than 100 percent would indicate that the participants felt as though their social needs were not entirely being met.

Appendix B

Fiction Narrative Coding Manual

Number of Real Words:

Under the "number of real words" column, we will find the number of words that are used in each narrative using a word count function. If the "Completed" column indicates "no," then we will enter the number of words that are actual engagement with the prompt.

Followed Directions:

Under the "Followed Directions" column, we indicate whether or not participants followed the directions of the prompt:

- **Yes:** The participant followed the prompt's directions. For a narrative to be coded as "Yes" the participant only needs to start off their narrative following directions and they do not have to name the bad news.

- **No:** The participant did not follow the prompt's directions in their narrative. A "No" could include that they did not write about bad news, that they did not follow whether they were directed to write To the character or From the character's perspective or their own narrative, and other indications of not following the prompt not specified here. Coders should still continue to code the other categories even if the narrative is coded as "No."

- **Other:** The "Other" code encompasses narratives that do not fit into either the "Yes" or "No" category. These narratives tend to follow certain aspects of the prompt, but not others. These narratives include writing that switches to another perspective, does not follow the same narrative throughout the entire piece of writing, provides context notes to the audience, writing that is in 3rd person, and other indications of a partly followed prompt. Coders should still continue to code the other categories even if the narrative is coded as "Other."

Examples:

- **Yes:** My mom walked into the room and told me terrible news. I couldn't believe what I was hearing. It was such a terrible feeling.

- **No:** All I heard was blah blah blah blah blahhhh blahhhh blooblahhhhh blah blah blah blah blah blah blah blahhhh blahhhh blooblahhhhh blah blah blah blah blah blah blah blahhhh blahhhh blooblahhhhh blah blah blah. blah blah blah blah blahhhh blahhhh blooblahhhhh blah blah blah blah blah blah blah blahhhh blahhhh blooblahhhhh blah blah blah blah blah blah blah blahhhh blahhhh.

- **Other:** Jake found out that Captain Holt would be putting him under witness protection. This is significant to Jake because he loves to work and hates not being able to work a case. He thinks about Holt's words, but ultimately decides to work a case behind Holt's back.

Completed:

Under the "Completed" column we indicate whether or not the participant fulfilled the prompt.

- **Yes:** A "Yes" indicates that the participant fulfilled the prompt. The narrative does not have to achieve the word length indicated in the prompt and as long as the participant stays with the same idea throughout the narrative it can be coded as a "Yes."

- **No:** A "No" encompasses narratives that do not stick with the same idea throughout or narratives that include non-words that do not fit the overall narrative. Coders should still continue to code the other categories even if the narrative is coded as "No."

Examples:

- **Yes:** My mom walked into the room and told me that my dog had died. I couldn't believe what I was hearing. He had just been here yesterday, wagging his tail, ready to go for a walk… I hope I never lose another pet suddenly again.

- **No:** My mom walked into the room and told me that my dog had died. I couldn't believe what I was hearing. All I heard was blah blah blah blah blahhhh blahhhh blooblahhhhh blah blah blah blah blah blah blah blahhhh blahhhh blooblahhhhh

blah blah blah blah blah blah blah blahhhh blahhhh blooblahhhhh blah blah blah. My mom walked into the room and told me that my dog had died. I couldn't believe what I was hearing. All I heard was blah blah blah blah blahhhh blahhhh blooblahhhhh blah blah blah blah blah blah blah blahhhh blahhhh blooblahhhhh blah blah blah blah blah blah blah blahhhh blahhhh blooblahhhhh.

Context Notes to the Audience:

This category indicates whether or not the participant's narrative provided notes to the audience that give background information about the characters or their situations.

- Yes: A "Yes" includes sentences that add context to the audience that does not seem to be in place with the narrative.
- No: A "No" indicates that the participant does not include context to that they think the audience needs to know for their narrative.

Examples:

- Yes: Dear Prince Charming, Cinderella lost her shoe at the ball. This is important because the Fairy Godmother made these shoes to fit Cinderella perfectly.
- No: Dear Prince Charming, I have to tell you about Cinderella. She lost her shoe at the ball.

1st/2nd/3rd Person Point of View:

This column will indicate what point of view the narrative is written in.

Dialogue:

This column indicates how much dialogue a participant uses in their narrative. This needs to be dialogue that is literally written using quotations in the narrative. Everything in Group 2 (Social, written **To** the character) does not count as dialogue.

- **None:** "None" indicates that there was no dialogue in the narrative.
- **Some:** "Some" indicates that there was dialogue in the narrative, but that the narrative did not have a heavy amount of dialogue.

- **A Lot:** "A Lot" indicates that the narrative had a large amount of dialogue or was mostly dialogue.

Examples:

- **None:** I found out that Nemo wasn't there. I shouted his name into the void, but heard no response. I started to panic.

- **Some:** I found out that Nemo wasn't there. I shouted "NEMO" into the void, but heard no response. I started to panic.

- **A lot:** I found out that Nemo wasn't there.

  Dory kept singing "Just keep swimming. Just keep swimming."

  I told her "Be quiet. I think I heard something over there. We need to swim to it."

  But Dory didn't listen to me. She kept singing "Just keep swimming just keep swimming.

Seriousness of Bad News:

Under the "Seriousness of Bad News" column, we indicate how serious the participant believes the news to be in the narrative. The coders judgement about the seriousness of the bad news should not factor in.

- **Not Serious:** This indicates that through the context of the narrative the writer did not perceive the bad news to be serious.

- **Somewhat Serious:** This indicates news that, to the writer, seems a little serious, but that it is not life changing or hugely impactful for the character in the narrative.

- **Very Serious:** Through the narrative's context, this indicates that the writer believes the bad news to be very impactful.

- **Unclear:** This code is used when the writer does not specify the bad news.

Examples:

- **Not Serious:** I had just found out from Nick that my fern died. I mean yeah it isn't great to have the best plant in the world die, but I'm not like heart broken about it. He maybe should have taken better care of it, but I love him anyways.

- **Somewhat Serious:** Wow Nick just told me my fern died. I am a little sad and I definitely cried (only for three days…), but it was already dying I suppose. I hope the fern can live a better life in plant heaven.
- **Very Serious:** I can't believe that Nick just told me that my fern died! He was supposed to take care of it, but I guess the best plant in the world means nothing to him. I will never forgive him.
- **Unclear:** Nick just told me some not so great news. I am definitely feeling down about it.

Transportation:

The category of "Transportation" encompasses how much a narrative can transport the reader into the story. This can be experienced cognitively, emotionally, or imaginatively. This category is coded using a 5-point Likert scale that measures how much the coders were transported when reading each narrative.

Some useful questions to consider when coding this category: How immersive is the text? Do you get swept up in what is happening in the text while you are reading? Is it exciting or emotionally compelling? Do you feel like you are there with the characters/writer while you are reading?

- **1= None**
- **3= Some**
- **5= A Lot**

Quality:

The "Quality" category captures the overall quality of the writing in the narrative. This category is coded using a 5-point Likert scale that rates the overall quality of the writing. This can include things such as, grammar, voice, style, and flow.

Some useful questions to consider when coding this category: How well-written is the text? Is it easy to understand what's going on? How effective is the text at communicating action/mood/ideas/etc? Does the writer seem to be making writing choices deliberately and adeptly?

- **1= Bad**

- **3= Okay**
- **5= Good**

Hope and Sympathy:

The "Hope and Sympathy" column captures how much the narrative continually describes and *"sits in"* the bad news.

- **None:** "None" indicates that the narrative only discusses the negative side of the bad news. There does not seem to be a bright side or a way out of the situation. The narrative is not exploring the space outside of the bad news. A useful metaphor for coding is if the narrative seems to *"sit in"* the bad news then it would be considered "none" for hope and sympathy.
- **A Little:** "A Little" indicates that the narrative included some hope and sympathy. It included lines that commented on the character's strengths or the possible good things to come. It does not just focus on the bad news.
- **A Lot:** "A Lot" indicates that the narrative hardly focuses on the negatives of the bad news. It instead looks to the positive and focuses on good things to come. It seems comforting or focuses on strengths.

Examples:

- **None:** I can't believe that Cece is with another man! I am just… GAH. This is the literal worst. It's my worst nightmare come to life. I am supposed to be with her! Who else is going to buy her mango chutney?? It sure won't be that heathen she's with now. MY life is over. Let me eat my butter in peace.
- **A Little:** I can't believe that Cece is with another man! I am just… GAH. This is the literal worst. It's my worst nightmare come to life, but I mean she does deserve happiness. Maybe he will buy her mango chutney, but for now leave me to be with the pigeons.
- **A Lot:** Dear Schmidt, I know that it must be really hard to see Cece with someone else, but I know you'll find someone amazing. You are witty and so in shape I'm sure another woman will want to date you. If you need anything please let me know. I'll always be here as your friend.

Details:

The "Details" column indicates whether or not the narrative provided details about the situation, the bad news, characters, or other aspects important to the narrative.

- **Yes:** A "Yes" is when a narrative includes any details about the characters, the bad news, or other important aspects. The coder gets a strong sense of what is happening in the narrative. The coder should see an attempt to engage with the source material.
- **No:** A "No" is when there are no details used in the narrative. There is no explanation of characters, setting, the bad news, and other important aspects. It is a generic narrative that could be applied to any source material.

Examples:

- **Yes:** Gideon couldn't believe what Harrowhark was telling her. She never thought that she would get to leave their planet, but with her? How dare she think that she can manipulate her like that. She would refuse to be her cavalier. She would never give up her long sword, especially not for Harrowhark, but she wouldn't stay on this dark planet to whither either.
- **No:** The bad news was terrible. She felt like she could cry. Bad news is never good for anyone.

Action or Drama:

The "Action or Drama" column indicates whether the narrative focuses more on an action sequence or more on emotions and drama.

- **Action:** "Action" indicates that the characters are actively influencing the world around them. The narrative focuses more on actions than feelings. A narrative may be coded as "action" if the narrative sounds similar to a plot synopsis.
- **Drama:** "Drama" indicates that the narrative focuses on the emotions of the characters. It does not focus on actions that the character/s takes, but rather how the character/s is feeling.

Examples:

- **Action:** I slammed my fist into the table after hearing about Pekka. He had already screwed me over once he wasn't going to get away with it again. I grabbed my cloak and strode onto the streets of Ketterdam. I wasn't going to go down without a fight.

- **Drama:** I couldn't believe what I was hearing. Pekka was back? In Ketterdam? He wasn't going to escape this time. I wouldn't give him the chance. I was ready for revenge.

Familiarity with the Character:

The "Familiarity with the Character" column indicates whether or not the coder was familiar with the character referenced in the narrative.

- **Yes:** A "Yes" indicates that the coder is almost definitely familiar with the character referenced in the narrative in any way.

- **No:** A "No" indicates that the coder is almost definitely not familiar with the character referenced in the narrative.

- **Unsure:** "Unsure" indicates that it was hard for the coder to tell from the text whether or not they were familiar with the character in the narrative.

Further Notes:

- All categories should be coded to the best of the coders' ability regardless of the answer to each column.

- A notes column may be included for coders to mark narratives that seem to not fit into the coding sequence or for coders to note their reasoning behind coding difficult narratives a certain way. Coders do not need to write notes for all, most, or even any.

References

Archive of Our Own. (2020, May 9). *A03 Statistics 2020: A Look Behind the Scenes*. https://archiveofourown.org/admin_posts/15931

Bal, P. M., & Veltkamp, M. (2013). How does fiction reading influence empathy? An experimental investigation on the role of emotional transportation. *PloS One*, *8*(1), e5534.

Cohen, J. (2004). Parasocial break-up from favorite television characters: The role of attachment styles and relationship intensity. *Journal of Social and Personal Relationships*, *21*(2), 187-202.

Davis, M. H. (1983). Measuring individual differences in empathy: Evidence for a multidimensional approach. *Journal of Personality and Social Psychology*, *44*(1), 113-126.

Derrick, J. L., Gabriel, S., & Tippin, B. (2008). Parasocial relationships and self-discrepancies: Faux relationships have benefits for low self-esteem individuals. *Personal Relationships*, *15*(2), 261-280.

Djikic, M., Oatley, K., Zoeterman, S., & Peterson, J. B. (2009). On being moved by art: How reading fiction transforms the self. *Creativity Research Journal*, *21*(1), 24-29.

Doyle, E. K. (2021). *Favored, Familiar, and Friendly: Identification and Relationships with Fictional Characters.* Unpublished master's thesis, University of Oregon.

Gabriel, S., Read, J. P., Young, A. F., Bachrach, R. L., & Troisi, J. D. (2017). Social surrogate use in those exposed to trauma: I get by with a little help from my (fictional) friends. *Journal of Social and Clinical Psychology*, *36*(1), 41-63.

Gardner, W. L., & Knowles, M. L. (2008). Love makes you real: Favorite television characters are perceived as "real" in a social facilitation paradigm. *Social Cognition, 26*(2), 156-168.

Gerrig, R. J. (1993). *Experiencing narrative worlds.* New Haven, CT: Yale University Press.

Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasive- ness of public narratives. *Journal of Personality and Social Psychology, 79,* 701–721.

Guren, C., McIlroy, T., & Sieck, S. (2021). COVID-19 and book publishing: Impacts and insights for 2021. *Publishing Research Quarterly*, *37*(1), 1-14.

Heinrich, L. M., & Gullone, E. (2006). The clinical significance of loneliness: A literature review. *Clinical Psychology Review*, *26*(6), 695-718.

Killgore, W. D., Cloonan, S. A., Taylor, E. C., Lucas, D. A., & Dailey, N. S. (2020). Loneliness during the first half-year of COVID-19 Lockdowns. *Psychiatry Research*, *294*, 113551.

Mar, R. A., & Oatley, K. (2008). The function of fiction is the abstraction and simulation of social experience. *Perspectives on Psychological Science*, *3*(3), 173-192.

Mar, R. A., Oatley, K., Hirsh, J., Dela Paz, J., & Peterson, J. B. (2006). Bookworms versus nerds: Exposure to fiction versus non-fiction, divergent associations with social ability, and the simulation of fictional social worlds. *Journal of Research in Personality, 40*(5), 694-712.

McGee, J. (2005). The ethics of fanfiction and real person fiction. *Communication Ethics, Media & Popular Culture*, *9*, 161-180.

Paravati, E., Naidu, E., & Gabriel, S. (2020). From "love actually" to love, actually: The sociometer takes every kind of fuel. *Self and Identity*, *20*(1), 1-19.

Russell, D., Peplau, L. A., & Cutrona, C. E. (1980). The revised UCLA Loneliness Scale: Concurrent and discriminant validity evidence. *Journal of Personality and Social Psychology*, *39*(3), 472-480.

Saucier, G. (1994). Mini-Markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment*, *63*(3), 506-516.

Sestir, M., & Green, M. C. (2010). You are who you watch: Identification and transportation effects on temporary self-concept. *Social Influence, 5*(4), 272-288.

Taylor, M., Hodges, S. D., & Kohányi, A. (2003). The illusion of independent agency: Do adult fiction writers experience their characters as having minds of their own?. *Imagination, Cognition and Personality*, *22*(4), 361-380.

Vlassis, A. (2021). Global online platforms, COVID-19, and culture: The global pandemic, an accelerator towards which direction?. *Media, Culture & Society*, *43*(5), 957-969.

Watson, D., Clark, L. A., & Tellegen, A. (1988). Development and validation of brief measures of positive and negative affect: the PANAS scales. *Journal of Personality and Social Psychology*, *54*(6), 1063.

UNDERSTANDING HOW MENTAL HEALTH CARE DELIVERY

AFFECTS CLIENT OUTCOMES AND SATISFACTION

by

BRIANA MOFHITZ-FAIETA

A THESIS

Presented to the Department of Psychology
and the Robert D. Clark Honors College
in partial fulfillment of the requirements for the degree of
Bachelor of Science

March 2022

An Abstract of the Thesis of

Briana Mofhitz-Faieta for the degree of Bachelor of Science
in the Department of Psychology to be taken June 2022

Title: Understanding How Mental Health Care Delivery Affects Client Outcomes and
Satisfaction

Approved: _____*Dr. Ruth Ellingsen*_____
Primary Thesis Advisor

The COVID-19 pandemic forced society to make many abrupt shifts in order to ensure the safety of the general public. In the mental health care industry, this required the field to take an unforeseen leap into remote delivery of clinical services. This study looks at the effects of the shift to telehealth through the lens of the University of Oregon Psychology Clinic. Clients receiving treatment at the clinic completed pre- and post-treatment outcome measures and had the opportunity to complete a clinic evaluation survey. Based on these measures, we found that telehealth services were able to produce clinically and statistically significant reductions in symptoms. Participants also indicated a strong satisfaction for treatment that they received, suggesting that, with proper framework and training opportunities, telehealth may be a viable resource in the mental health care industry.

Acknowledgements

I would like to thank Dr. Ruth Ellingsen, Dr. Sara Hodges and Dr. Carol Paty for their dedication to helping me conduct the present research. Each of these brilliant women inspired me through their individual accomplishments, alongside their influence in mine. Their help and support made it all possible, from the impromptu meetings to the countless email streams. I would also like to thank the University of Oregon, Clark Honors College and the Psychology Clinic for giving me the privilege of completing the captivating, yet strenuous, process of conducting undergraduate research. To my family, I would like to thank you for your unconditional support throughout every triumph and tribulation. And, finally, to my friends, thank you for patiently listening to me as I have talked incessantly about the rigorous process to complete this thesis.

Table of Contents

List of Tables

Introduction

As society's awareness of the impacts of mental health and treatment for mental illness has led to a significant increase in accessibility and resources available to those who need it, the field still has a long way to go in order to ensure that its recipients are receiving care that best suits their individual needs. According to the National Institute of Mental Health, in 2019 there were an estimated 51.5 million adults (approximately 20.6% of the population) in the United States living with mental illness. Of those, approximately 40.1% received treatment in the form of in-patient therapy, out-patient therapy, or prescription medication (NIMH, 2021). Considering that one in five people in the United States deals with mental illness, and that the majority do not receive treatment, it is important that treatment options are as efficient and effective as possible, while also taking into account individual needs.

Given recent circumstances of the COVID-19 pandemic, there has been a substantial influx of telehealth options for therapy, providing a new perspective on how feasible and effective online mental health treatment can be. Prior to the onset of the pandemic, approximately 7% of mental health services provided by psychologists were taking place through telehealth. Of the various types of practice settings, the Veteran's Affairs Medical Center had the highest usage of remote mental health services prior to the onset of the pandemic, due to the geographically remote locations in which clients resided (Godleski, Darkins, & Peters, 2012; Pierce, Perrin, Tyler, Mckee, & Watson, 2021). Based on a report from the RAND Corporation, there were more than 300,000 military service members and one million family dependents that lived in geographically remote locations, requiring that accessible options be made available

6

(Luxton, Pruitt, Wagner, Smolenski, Jenkins-Guarnieri, & Gahm, 2016). In an effort to understand the feasibility, safety and clinical efficacy of Home-Based Telebehavioral Health, Luxton and colleagues (2016) compared the effects of the online setting with those of a traditional in-office setting on military personnel and veterans dealing with depression. In their study, they randomly assigned participants to either condition, then assessed at baseline, mid-treatment (four weeks), post-treatment (eight weeks), and three months after treatment. They found that there were significant reductions in depression symptoms and hopelessness in both groups, along with improvement observed on measures of PTSD symptoms and anxiety. In terms of safety, they also found that there was no evidence of clinical worsening for the telehealth condition, making online treatment no less safe than traditional in-office treatment.

Despite telehealth being an effective and accessible means of delivering therapy, many therapists did not choose to adopt it into their practice prior to the COVID-19 pandemic; only about 21% of psychologists had used telehealth in their practice (Pierce, Perrin, & McDonald, 2020). Furthermore, approximately 75% of psychologists and other mental health providers claimed that they were unwilling to refer clients to telehealth therapy options (Perle, Burt, & Higgins, 2014). Along with the rapid changes in society surrounding the COVID-19 pandemic, the mental health field had to adopt a new structure, with a 12-fold increase in the amount of mental health treatment delivered through telehealth (Pierce, et al., 2021). A national study conducted during the pandemic assessed how psychologists thought mental health care delivery might change in the post-pandemic future. They found that participants projected that 35% of their work would be conducted via telehealth even after the pandemic restrictions are lifted,

7

as compared to the 7% that was being conducted prior to the pandemic (Pierce et al., 2021). These results indicate an important shift in attitudes surrounding the use of telehealth among mental health providers.

Risks and Benefits of Telehealth

The dawn of the Covid-19 pandemic served as the perfect storm to allow the mental health care field to reassess the feasibility of remote therapy. Prior to the pandemic, there were hesitations and negative attitudes among therapists surrounding the usage of telehealth, alongside concerns regarding the efficacy of providing online therapy. Given that telehealth relies on the functionality of an online resource, there are technical interferences and experiential differences that can be expected with the use of this modality, such as unstable connection, frozen screen, bad audio, and poor lighting (Fernández-Álvarez, & Fernández-Álvarez, 2021; Payne, et.al., 2020). Beyond these concerns, there is apprehension regarding the therapeutic experience in an online therapy environment. Given that there is no physicality in a remote model, social facets such as eye contact, tone of voice, open posture, body movements, and synchrony may be impacted. Furthermore, clients and therapists may be prone to distraction or may pay too much attention to their own appearance on the screen, making it difficult to fully engage in the session (Markowitz, Milrod, Heckman, Bergman, & Amsalem, 2020).

Additionally, when it comes to telehealth there are operational concerns regarding informed consent, confidentiality, crisis management, and client identities (Rochlen & Speyer, 2004); these are important safety measures taken by the mental health care industry to ensure the protection of clients and the information they disclose. Without being face-to-face, there is an inhibited ability to guarantee these facets are

8

sufficient in serving their intended purpose. Beyond this, there is uneasiness surrounding the perceived decrease in formality that comes with an online therapy environment; with the lack of physical proximity, there is concern that this may lead to boundary crossing and deleterious effects to professionalism associated with the therapeutic relationship and treatment outcomes (Simpson, et al., 2020).

Alongside a general lack of training in providing remote therapy services, many therapists are not eager to implement telehealth in their practice due to the above concerns confounding its implementation. In multiple studies, it has been established that working alliance (i.e., the ability to create a meaningful relationship between therapist and client), is interpreted as significantly worse through the eyes of the practicing therapist when conducting therapy online compared to traditional, in-person services (Norwood, Malins, & Sabina-Farrell, 2018). In contrast, research has indicated that clients tend to have positive attitudes toward online therapy and have less concern about potential problems than their therapists (Fernández-Álvarez & Fernández-Álvarez, 2021). Furthermore, research has found that the telehealth modality is capable of providing similar outcomes, and that clients tend to be open to utilizing it (Fernández-Álvarez & Fernández-Álvarez, 2021; Norwood, Malins, & Sabina-Farrell, 2018).

Research has also indicated that remote therapy can provide benefits that are not available in traditional settings. As opposed to the in-person therapy setting, an online environment has the opportunity to promote a sense of safety, alongside a neutral power balance, that has potential to facilitate greater disclosure (Fernández-Álvarez, & Fernández-Álvarez, 2021; Simpson, et al., 2020). For example, those facing mood

9

disorders, interpersonal avoidance, social anxiety, or individuals who may find in-person therapy to be overwhelming, might benefit from the decreased arousal levels associated with conducting therapy in an online setting (Reynolds, Stiles, Bailer, Hughes, 2013; Simpson, et al., 2020). Additionally, a remote modality has the ability to promote attendance to therapy sessions without concern of being observed or stigmatized, which may be beneficial for demographics who may avoid receiving treatment in traditional settings. Evidence also suggests that children and teenagers may be more open to online psychotherapy given their familiarity and comfort with the modality (Simpson, et al., 2020).

Importantly, telehealth has the opportunity to mitigate issues associated with the mental health treatment gap that exists in today's society. Beyond the scope of the United States, it is estimated that there is a significant treatment gap in other developed countries (between 44-70%). And in developing countries, over 80% of those facing mental health problems are unable to access treatment (Mohsen, 2020; Ghebreyesus, 2019). Providing remote access to mental health treatment may provide a viable solution to confounds associated with this treatment gap, overcoming some of the barriers that make it difficult to receive treatment for mental illness.

Along with the unforeseen circumstances of the Covid-19 pandemic came a unique opportunity to understand the practical implications of telehealth options for mental health treatment. The risks and rewards of the modality have been more widely discussed and addressed. Potential barriers to the success of online therapy can be found in technical issues, operational concerns and negative attitudes held by therapists regarding therapeutic alliance. However, despite these confounds, telehealth has the

10

potential to increase the accessibility of mental health care substantially, while also providing benefits such as greater disclosure due to increased comfortability and a heightened sense of safety.

Objectives

The goal of the current study is to investigate therapeutic outcomes and service satisfaction after the abrupt shift from in-person to online therapy among clients receiving treatment at the University of Oregon Psychology Clinic. This research will add to our understanding of the accessibility and efficacy of therapy conducted via telehealth. Based on previous findings that telehealth is an effective way to deliver mental health services, we expect to find that adult clients will see significant improvements in their treatment outcomes (i.e., reduction in anxiety symptoms, depression symptoms, and emotion regulation difficulties). Among minor clients (those under 18 years of age), we hypothesize that overall child behavior difficulties will significantly decrease with the use of telehealth services. For exploratory purposes, we will also investigate possible impact of client age, gender, student status, and number of total sessions on treatment outcomes. Further, we investigate service satisfaction among telehealth clients, including their ratings of their telehealth experience and their sense of progress toward their therapy goals.

Methodology

The present research used archival data collected by the University of Oregon Psychology Clinic, a training clinic for clinical psychology doctoral students located on the University of Oregon campus. Data used in the current study were collected from May 2020 through September 2021 and includes pre- and post- treatment questionnaires from 29 clients who received telehealth services during the COVID-19 pandemic. All participants/participants' legal guardians signed a consent form prior to services that allows for the use of their archival data for research. The sample consists of 10 minor clients and 19 adult clients; data from these groups were analyzed separately. Participants who had received any in-person services, or who did not have complete outcome measure data were not part of this study.

Participants

Adult Group

The adult group consisted of 19 clients and included six males, twelve females, and one non-binary client, with ages ranging from 18 to 45 years old ($M = 25.74$, $SD = 7.50$). Participants in this group received treatment for a wide range of diagnoses (e.g., anxiety, depression, and trauma), and on average attended 17.11 sessions ($SD = 7.55$). In the adult group, 79% of the participants identified as white, 14% identified as Hispanic, 5% identified as Asian, and 2% identified as something else or did not respond. As this clinic is located on the University of Oregon campus, 53% of the sample consisted of university students.

13

Minor Group

The minor group consisted of three male and seven female clients, ranging in age from 5 to 17 years old (M = 11.5, SD = 4.43). The average number of sessions attended was 34.2 (SD = 9.83). In this group, 80% identified as white, 10% identified as Hispanic and 10% identified as mixed. Clients presented with a range of presenting issues/diagnoses, including anxiety, obsessive-compulsive disorder (OCD), and trauma.

Measures

Adult Group

Adult client outcomes were assessed with three measures: the Patient-Reported Outcomes Measurement Information System (PROMIS) Depression Short-Form, PROMIS Anxiety Short-Form, and Difficulty in Emotion Regulation Scale (DERS), all of which were collected pre- and post-treatment. The PROMIS Depression Short-Form (Pilkonis et al., 2011) is an eight item self-report questionnaire assessing depressive symptoms over the course of the last seven days, scored on a 5-point Likert scale (1 = Never, 5 = Always); statements include "I felt worthless" and "I felt that I had nothing to look forward to."

The PROMIS Anxiety Short-Form (Pilkonis et al., 2011) is a seven item self-report questionnaire assessing symptoms associated with anxiety and emotional distress over the course of the past seven days, scored on a 5-point Likert scale (1 = Never, 5 = Always); statements include "I felt worried" and "I found it hard to focus on anything other than my anxiety." PROMIS measures are scored on a T-score metric with a mean of 50 and standard deviation of 10 in the U.S. general population. Scores are

categorized as *within normal limits* (T < 55), *mild* (T = 55-60), *moderate* (T = 60-70), and *severe* (T > 70).

The DERS (Gratz & Roemer, 2004) is a 36 item self-report questionnaire, scored on a 5-point Likert scale (1 = Almost never, 5 = Almost always) designed to assess multiple aspects of emotional dysregulation, including nonacceptance of emotional responses, difficulty engaging in goal-directed behavior, impulse control difficulties, lack of emotional awareness, limited access to emotion regulation strategies, and lack of emotional clarity. Statements include "I am confused about the way I feel" and "When I am upset, I have difficulty getting work done." Total scores range from 36 to 180 and higher scores suggest greater problems with emotion regulation. There are no standardized clinical cutoffs for the DERS; however, the mean total score for participants in a group treatment for Borderline Personality Disorder (BPD) is 127.92.

Client treatment satisfaction was assessed with items from a Clinic Evaluation provided at the end of services (i.e., post-treatment). All items on this form are rated on a 5-point scale ranging from "Strongly Disagree" to "Strongly Agree." Responses to the item "Services have been effective in helping me reach my goals" were evaluated in the current study. Another measure of client satisfaction was a telehealth rating scale; clients rated telehealth services on a scale of 0-100, with a score of 100 indicating ultimate satisfaction with services. Clients were also given the option of leaving comments regarding their treatment experience.

Minor Group

Minor client outcomes were assessed with pre- and post-treatment caregiver-report on the Strength and Difficulties Questionnaire (SDQ; Goodman, 1997). The SDQ is a 25-question screening questionnaire that assesses child/adolescent emotional symptoms, conduct problems, hyperactivity/inattention, peer relationship problems and prosocial behavior. Items are rated as 0 = Never, 1 = Somewhat True, and 2 = Certainly True. Total scores of 17 and above are considered to be clinically elevated.

Treatment satisfaction was assessed with items from a Clinic Evaluation provided to caregivers of minor clients at the end of services (i.e., post-treatment). The Clinic Evaluation is identical to the one used in the adult sample, although questions are written for caregivers of minor clients. Responses to the item "Services have been effective in helping my child reach their goals" were used in the current study. Caregivers of minor clients also provided a rating of telehealth satisfaction on a scale of 0-100 and were given the option of leaving comments.

Results

In the following analysis of outcome measures, it is important to note that a single regression analysis was conducted on change scores from post- to pre-treatment to investigate the significance of changes in measures over the course of treatment. Descriptive statistics, such as the mean (M) as a measure of central tendency, standard deviation (SD) as a measure of variability, and number of participants (n) have been reported, along with the t-statistic (t), degrees of freedom ($n - 1$), 95% confidence interval (95%CI) and the p-value (p) where a $p < .05$ is considered to be statistically

significant. Negative t-statistics indicate that post-treatment scores were lower (i.e., less pathology) than pre-treatment scores. In the regression analyses that are reported below, F-statistics and p-values are reported for statistical significance, and η_p^2 is reported as a measure of effect size.

Adult Group

Results from a general linear model analysis indicated significant clinical improvements in the adult group. Change scores on the PROMIS depression measure from pre-treatment ($M = 57.49$, $SD = 6.60$) to post-treatment ($M = 50.26$ $SD = 6.40$) significantly decreased, indicating improvement in depressive symptoms over the course of treatment, $t(18) = -4.99$, 95%CI [-10.27, -4.19], $p < .001$. On the PROMIS anxiety measure, change scores between pre-treatment ($M = 61.64$, $SD = 6.02$) and post-treatment ($M = 54.04$, $SD = 7.49$) significantly decreased, suggesting improvement in symptoms of anxiety over the course of treatment, $t(18) = -5.89$, 95%CI [-10.31, -4.89], $p < .001$. In regard to the DERS measure, changes scores from pre-treatment ($M = 96.22$, $SD = 22.59$) to post-treatment ($M = 73.94$, $SD = 20.64$) significantly decreased, indicating reduced difficulties in emotional regulation over the course of treatment, $t(16) = -5.98$, 95%CI [-28.13, -13.40], $p < .001$. Note that two clients in the sample were missing data for the DERS measure; therefore, they were not included in the outcome analysis.

After establishing that outcome measures for the adult group were significant, the decision was made to run an exploratory regression analysis to analyze the predictive ability of four variables on the three clinical outcome measures: age, gender, student status and total number of sessions. However, it is important to note that given

the small sample size and low statistical power, results from this analysis should be taken with a heed of caution. This analysis was done for exploratory purposes and not for hypothesis testing. Across the three outcome measures, a general liner model analysis indicated that none of these predictors were significantly related to improvement on these measures. For regression analysis of predictor variables, see Table 1.

Data from the clinic evaluation question "Services have been effective in helping me reach my goals" were evaluated to assess treatment satisfaction. Responses to this measure ranged from "Strongly Disagree" to "Strongly Agree". In the adult group (n = 14), 71.4% (n = 10) of clients responded with "Strongly Agree", 21.4% (n = 3) responded with "Agree", and 7.1% (n = 1) of clients responded with "Strongly Disagree". On the telehealth rating scale, clients (n = 14) ranged in responses from 70-100 (M = 90, SD = 9.20).

Minor Group

A general linear model analysis indicated significant improvements in mental health outcomes of minors with the use of telehealth services. The change scores of the SDQ measure from pre-treatment (M = 16.60, SD = 7.00) to post-treatment (M = 10.40, SD = 3.44) significantly decreased, indicating reduced negative mental health outcomes, $t(9)$ = -3.66, 95%CI [-10.03, -2.37], p = 0.005. After establishing that the measure had significant effects, another exploratory regression model was conducted to analyze the contributions of three predictor variables (age, gender and number of total sessions) to the clinical outcome variables. The model analysis suggested that none of the predictors significantly predicted change scores on the SDQ. However, once again, due to the

18

small sample and low power, results from this analysis should be taken with caution and considered only to be exploratory. See Table 1 for regression analysis of the predictor variables.

Minor clients' caregivers ($n = 10$) responded to the Clinic Evaluation item "Services have been effective in helping my child reach their goals," with 90% ($n = 9$) of parents responding with "Strongly Agree" and 10% ($n = 1$) of participants responding with "Agree". Telehealth satisfaction rated by minor client caregivers ranged from 80-100 ($M = 95.00$, $SD = 7.07$).

Table 1

Regression Analysis of Predictor Variables for Change Scores of Client Outcome
Measures

	PROMIS-DEP			PROMIS-ANX			DERS			SDQ		
	p-value	*F*	η_p^2	*p-value*	*F*	η_p^2	*p-value*	*F*	η_p^2	*p-value*	*F*	η_p^2
AGE	0.110	2.34	0.11	0.086	3.46	0.24	0.962	0.002	0.14	0.326	1.15	0.16
GENDER	0.243	1.50	0.10	0.778	0.08	0.004	0.129	2.69	0.18	0.903	0.02	0.003
TOTAL SESSIONS	0.851	0.04	0.003	0.527	0.42	0.03	0.487	0.52	0.04	0.960	0.003	0.08
STUDENT STATUS	0.247	1.47	0.33	0.903	0.02	0.15	0.415	0.72	0.08	N/A		

Statistics Key:

p-value: probability of obtaining these results by chance if there is no effect; $p < .05$
considered to be statistically significant

F-statistic: comparison of predictor against a model without predictors; higher values
indicate greater proportion of explained variance relative to unexplained variance

Partial eta squared (η_p^2): Measure of effect size (proportion reduction of error
associated with the inclusion of the predictor in the regression model); Values range
from 0 to 1, with a score close to one indicating a high proportion of variance explained

Discussion

The primary goal of this study was to investigate therapeutic outcomes and service satisfaction after an abrupt shift to telehealth for those receiving services at the University of Oregon Psychology Clinic during the COVID-19 pandemic. In our sample, we found statistically significant improvements on mental health outcomes, indicating that services provided via telehealth were effective. Beyond the scope of statistical significance, results were also clinically meaningful.

The PROMIS depression measure had a pre-treatment average in the mild range of depressive symptoms. From pre- to post-treatment there was a significant reduction in symptoms, with a post-treatment average in the normal range of depressive symptoms. The PROMIS anxiety measure had a pre-treatment average in the moderate range of anxiety symptoms and significantly dropped to a post-treatment average in the normal range of anxiety symptoms. The reductions in symptoms of depression and anxiety were both statistically and clinically significant. Depression and anxiety symptoms started in the clinical range (mild and moderate levels, respectively) and at post-treatment the average symptoms were within normal limits. Furthermore, we saw a 22-point decrease – a significant reduction – in emotion regulation difficulties. As a whole, these results indicate statistically significant and clinically meaningful change in client symptoms as a result of telehealth mental health services. Findings were similar in the minor group. Pre-treatment, the minor group average of overall mental health difficulties on the SDQ was right at the clinical cut-off; the post-treatment average was significantly lower and no longer clinically elevated. Overall, results on all four

21

outcome measures suggest that telehealth services were, on average, effective in reducing clinical-level symptoms to post-treatment scores within normal limits.

To gather more information about the implications of telehealth, we ran a regression analysis to see if different variables, such as age, gender, total number of sessions or student status (adult group only) had an effect on client outcomes across the four measures. Our results were unable to confirm whether these variables were predictors of the clinical outcomes. Because of the small sample size, we likely did not have the power to determine whether these null results are because clinical outcomes with the use of telehealth do not significantly differ among clients who vary in age, gender identity, whether or not they were a student at the university, or how many sessions they had. It is possible that these variables could show significant effects on outcomes in a larger sample, but our inclusive results are at least not inconsistent with the idea that telehealth services may be able improve mental health outcomes across a wide range of people and circumstances.

In addition to examining treatment outcomes, we investigated telehealth satisfaction to understand how clients felt about the treatment they received. Based on responses to one of the questions on the clinic evaluation, we found that 97% of our entire sample agreed or strongly agreed that receiving telehealth services was effective in helping them reach their therapy goals. Furthermore, telehealth ratings on a scale from 1-100 had means of 90 and 95 in the adult and child groups, respectively. Both of these measures indicate high satisfaction among clients who received services delivered via telehealth, and that clients felt like they were able to make progress in their mental health.

22

The above outcome and satisfaction results were corroborated by additional comments provided by clients on the clinic evaluation. In the adult sample, these comments included "She was great and I felt really listened to and comfortable," "I think [Therapist] did an amazing job as my therapist and although I have a lot to work on, I now feel positive about the direction I am moving in," and "[Therapist] was the best… she did everything I needed and more. I am grateful for this experience and her help will stay with me forever." In the minor sample, caregivers of clients also provided positive remarks, including "I was surprised how much rapport [Therapist] was able to build with [Client], considering they only ever met via Zoom. She's a very talented clinician and we are very thankful for our 6 months with her," "[Therapist] is amazing and I couldn't be more pleased with our progress and relationship! The therapy our son received has been life changing," and "We are so thankful for [Therapist]'s help. She was wonderful with [Client] and always met her wherever she was each week, and made her feel very comfortable. [Client]'s improvement has gone beyond our expectations!" This feedback illustrates that even though clients never met with their therapists in person, they were still able to create meaningful connections that helped them reach their therapeutic goals.

Prior to the COVID-19 pandemic, therapists were hesitant to adopt telehealth into general practice due to the fear that the functionality of the modality could have detrimental impacts to therapeutic alliance and mental health outcomes (Norwood et al., 2018; Fernández-Álvarez & Fernández-Álvarez, 2021; Simpson, et al., 2020). Our findings add additional evidence to the relatively small, but quickly growing, evidence base for telehealth services, with the implication that clients are satisfied with the

experience of telehealth and that mental health outcomes can significantly improve over the course of telehealth therapy.

Limitations

As can be expected, there are limitations present in the current research study, including sample size, population, and confounds surrounding the COVID-19 pandemic. The sample consisted of clients who received services from the University of Oregon Psychology Clinic after the shift to telehealth during the COVID-19 pandemic. This made it difficult to collect a large sample, especially considering that minors and adults were given different outcome measures. In particular, this had an impact on the regression analysis, increasing the likelihood of a type II error, wherein an effect was present, but the study did not have the statistical power to detect it. In the realm of research, there is controversy over the correct ratio of predictor variables to number of participants. Some argue that there should be over 100 participants to run a model with a single predictor (Green, 1991), whereas others suggest that two subjects per variable is sufficient (Austin & Steyerberg, 2015). Therefore, future studies would benefit from running a regression analysis with more statistical power. For example, it could be possible that age is a significant predictor of clinical outcomes with treatment via telehealth, considering that young people tend to be more familiar and open-minded with technology. However, these effects would likely have been undetectable due to the limited number of participants and limited variability found in the present study.

It is also worthwhile to note that our research did not utilize random assignment to compare results across in-person and remote delivery of mental health services. By utilizing random assignment, a study would be better able to conclude that any outcome

differences were due to differences in how telehealth compares to the traditional, in-person setting. Furthermore, this type of design would allow the field to take a deeper look into how these effects play out across various populations and diagnoses.

It is important to recognize that given the circumstances of the pandemic, clients were not given the option to choose between in-person or telehealth services. This could have influenced the satisfaction ratings of telehealth, considering pandemic conditions (e.g., isolation, heightened anxiety, unpredictability, etc.) and that the only option for treatment was via telehealth. Attention should also be given to the external influence that the pandemic could have had on client's improvements. Mental health outcomes may have improved with the adjustment to the pandemic over time, suggesting that as the pandemic became more familiar, clients may have naturally seen improvement. In the future, it would be important to conduct a study about the effects of telehealth during a time where mandates, restrictions, and unpredictability do not confound daily life.

Future Directions

The COVID-19 pandemic has provided a unique situation to investigate new opportunities within the field of mental health care. As more studies are published, it is important that research focuses on which circumstances are suitable for telehealth services, as well as providing adequate structure and training to facilitate its adoption into general practice. In our study, we were able to cover a wide range of diagnoses; however, a more individualized approach to the uses of telehealth would be helpful toward understanding who would benefit most from having the option to utilize telehealth services. Individuals receiving treatment for mood disorders, such as

25

generalized anxiety disorder, depression or social anxiety may benefit from the accessibility and decreased arousal levels associated with telehealth. In contrast, an individual receiving treatment for attention deficit hyperactive disorder may be under-stimulated by telehealth and prone to distraction. Continuing research in this area could allow for a more personalized approach to mental health care, while also making services more accessible to those who may have circumstances that make a traditional in-person approach less viable.

With the rapid shift to remote services, the University of Oregon Psychology Clinic was required to quickly pivot to a new format of service delivery in order to meet the needs of their clients. It is promising that a small training clinic was able to make these adjustments in an efficient manner, suggesting that the implementation of telehealth could be one that is more feasible than previously expected. However, the field as a whole could benefit from additional infrastructure and therapist training to better facilitate the adoption of telehealth into general practice.

With the unforeseen circumstances that stemmed from COVID-19, the mental health care industry was forced to look at telehealth in a new light. Prior to the pandemic, there were few options for practicing and receiving remote mental health services; however, once it was the only safe option for delivering services, the field shifted their attitudes toward the modality. As more therapists have considered permanently adopting telehealth into their practice, it is imperative that research focuses on establishing credibility and framework to ensure that the field is able to provide services that best suit the needs of the client.

Bibliography

Austin, P. C., & Steyerberg, E. W. (2015). The number of subjects per variable required in linear regression analyses. *Journal of Clinical Epidemiology*, *68*, 627–636. https://doi.org/10.1016/j.jclinepi.2014.12.014.

Fernández-Álvarez, J., & Fernández-Álvarez, H. (2021). Videoconferencing psychotherapy during the pandemic: Exceptional times with enduring effects? *Frontiers in Psychology*, *12,* 1-8. https://doi.org/10.3389/fpsyg.2021.589536

Ghebreyesus, T. A. (2019). *The WHO special initiative for mental health (2019-2023): Universal Health Coverage for Mental Health*. World Health Organization. Retrieved February 5, 2022, from https://apps.who.int/iris/handle/10665/310981

Godleski, L., Darkins, A., & Peters, J. (2012). Outcomes of 98,609 U.S. Department of Veterans Affairs patients enrolled in telemental health services, 2006 –2010. *Psychiatric Services, 63*, 383–38. http://dx.doi.org/10.1176/appi.ps.201100206

Goodman, R. (1997). Strengths and Difficulties Questionnaire. *PsycTESTS Dataset*. https://doi.org/10.1037/t00540-000.

Gratz, K. L. & Roemer, L. (2004). Multidimensional assessment of emotion regulation and dysregulation: Development, factor structure, and initial validation of the Difficulties in Emotion Regulation Scale. *Journal of Psychopathology and Behavioral Assessment, 26,* 41-54.

Green, S. B. (1991). How many subjects does it take to do a regression analysis. *Multivariate Behavioral Research, 26,* 499-510. DOI: 10.1207/s15327906mbr2603_7

Luxton, D. D., Pruitt, L. D., Wagner, A., Smolenski, D. J., Jenkins-Guarnieri, M. A., & Gahm, G. (2016). Home-based telebehavioral health for U.S. military personnel and veterans with depression: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 84*(11), 923–934. https://doi.org/10.1037/ccp0000135

Markowitz, J. C., Milrod, B., Heckman, T. G., Ph, D., Bergman, M., & Amsalem, D. (2020). Psychotherapy at a distance. *American Journal of Psychiatry*, *178,* 1‐7. doi:10.1176/appi.ajp.2020.20050557

Mohsen, L. (2020, October 26). *Covid-19 and the need for action on Mental Health*. United Nations. Retrieved February 05, 2022, from https://www.un.org/en/un-chronicle/covid-19-and-need-action-mental-health

National Institute of Mental Health. (2021, January). *Mental illness*. Retrieved February 05, 2021, from https://www.nimh.nih.gov/health/statistics/mental-illness.shtml

Norwood, C., Moghaddam, N. G., Malins, S., & Sabina-Farrell, R. (2018). Working alliance and outcome effectiveness in videoconferencing psychotherapy: A systematic review and Noninferiority Meta-analysis. *Clinical Psychology & Psychotherapy*, *25*(6), 797- 808. https://doi.org/10.1002/cpp.2315

Payne, L., Flannery, H., Gedara, C. K., Daniilidi, X., Hitchcock, M., Lambert, D., et al. (2020). Business as usual? Psychological support at a distance. *Clinical Child Psychology and Psychiatry, 25*, 672‑686. doi: 10.1177/1359104520937378

Perle, J. G., Burt, J., & Higgins, W. J. (2014). Psychologist and physician interest in telehealth training and referral for mental health services: An exploratory study. *Journal of Technology in Human Services*, 32, 158 – 185.

Pierce, B. S., Perrin, P. B., & McDonald, S. D. (2020). Path analytic modeling of psychologists' openness to performing clinical work with telepsychology: A national study. *Journal of Clinical Psychology*, *76,* 1135–1150. http://dx.doi.org/10.1002/jclp.22851

Pierce, B. S., Perrin, P. B., Tyler, C. M., McKee, G. B., & Watson, J. D. (2021). The COVID-19 telepsychology revolution: A national study of pandemic-based changes in U.S. mental health care delivery. *American Psychologist, 76*(1), 14–25. https://doi.org/10.1037/amp0000722

Pilkonis, P.A., Choi, S.W., Reise, S.P., Stover, A.M., Riley, W.T., & Cella, D. (2011). Item banks for measuring emotional distress from the client-reported outcomes measurement information system (PROMIS): Depression, anxiety, and anger. *Assessment*, *18*(3), 263-283.

Reynolds, D., Stiles, W., Bailer, A., & Hughes, M. (2013). Impact of exchanges and client therapist alliance in online-text psychotherapy. *Cyberpsychology, Behavior and Social Networking, 16*(5), 370‑377. https://doi.org/10.1089/cyber.2012.0195

Rochlen, A., Zack, J., & Speyer, C. (2004). Online therapy: Review of relevant definitions, debates, and current empirical support. *Journal of Clinical Psychology, 60*(3), 269‑283. https://doi.org/10.1002/jclp.10263

Simpson, S., Richardson, L., Pietrabissa, G., Castelnuovo, G., & Reid, C. (2020). Videotherapy and therapeutic alliance in the age of Covid-19. *Clinical Psychology & Psychotherapy*, *28*(2), 409-421. https://doi.org/10.1002/cpp.2521

28

Following the Child's Lead: Evaluating Responsive Caregiving Behaviors in a Strength-Based

Video Coaching Intervention

Heather Ralph

Department of Psychology, University of Oregon

Author Note

Correspondence concerning this article should be addressed to Heather Ralph, Department of Psychology, University of Oregon, Eugene, OR 97403.

E-mail: hralph@uoregon.edu

Abstract

Interventions that emphasize responsive caregiving can reverse the negative effects of early life stress exposure on development in early childhood. Despite this knowledge, there is a lack of consensus in the field about which behaviors define "responsive caregiving". The Filming Interactions to Nurture Development (FIND) Intervention is a responsive caregiving intervention that guides caregivers towards serve and return interactions that follow the child's lead. Preliminary evidence from pilot trials suggests that the FIND intervention may significantly impact both caregiver and child outcomes, but it is not yet known how FIND changes the way caregivers and children interact. This pilot trial evaluated whether FIND increases the frequency of caregiver "following" behaviors during dyadic freeplay tasks. Results from a pilot trial using 18 mother-infant dyads (11 FIND families and 7 control families) showed an increase in following behaviors in FIND families and no change in control families, although this difference was nonsignificant. ANCOVA analyses showed that the level of following behaviors before the intervention significantly predicts the level of following behaviors after the intervention. These results suggest that the FIND intervention may increase following behaviors in caregivers, but a larger sample size is needed to determine whether these effects are the result of regression to the mean or intervention findings.

Introduction

Responsive caregiving provides a supportive foundation for healthy development in children (Ainsworth et al., 1974). A lack of responsive caregiving can lead to negative neurobiological effects and physiological changes during development (Bourne et al., 2022), which can affect future physical and mental wellbeing (Levine, 2005; Shonkoff et al., 2009). Research has shown that interventions that emphasize responsive caregiving can potentially reverse the negative neurobiological effects of stress exposure (Bruce et al., 2013; Miller et al., 2014). However, there is a lack of consensus in the field over the behaviors that comprise "responsive caregiving". This study utilizes a framework for responsive caregiving laid out by Fisher et al. 2016 in the Filming Interactions to Nurture Development (FIND) Intervention, a strength-based video coaching intervention program that encourages following the child's lead. We evaluated how the FIND intervention impacts the way that parents follow their child's lead, a critical mechanism driving positive downstream caregiver and child outcomes.

The Importance of Responsive Caregiving

The first few years of life are a critical period for infant development. Environmental experiences can significantly affect socio-emotional development as well as cognitive functioning (Evans, 2006; Gilmore et al., 2018). Infants that experience early life stress due to environmental factors are at higher risk of developmental difficulties and poor stress regulation (Levine, 2005). Since many environmental experiences may come from parent-child interactions, a lack of supportive parenting behaviors can contribute to early life stress. Because of this, there has been an increasing focus on developing interventions that encourage responsive caregiving to ensure positive parenting skills and prevent future developmental difficulties in infants (Blair & Raver, 2016; Bruce et al., 2013). Research has shown that caregiver responsiveness and secure

attachment can buffer the reactivity of the stress regulation system in infants (Gunnar, 1998; Laundry et al., 2006; Mihelic et al., 2017). Scherer and colleagues (2019) found that when parents display responsive caregiving behaviors more often, children are more likely to experience positive socioemotional outcomes. These studies suggest that increasing reciprocal and supportive interactions encourages healthy neurobiological development in infants as well as reduced maternal depression (Lucas et al., 2018).

What is Responsive Caregiving?

Although "responsive caregiving" is known to produce positive outcomes for both caregivers and infants, there is a lack of consensus in the field about exactly which behaviors define this term. Two common conceptual definitions of responsive caregiving are contingent responsiveness and parental sensitivity. Both concepts allow for healthy development of caregiver-infant relationships but are expressed through different behaviors.

Many researchers consider contingent responsiveness to be an essential component of responsive caregiving. Caregivers engage in "contingent responsiveness" when they engage in back-and-forth interactions with their child, follow their child's cues, and respond with proper feedback. Serve and return interactions are an example of this: a child "serves" by starting an interaction through gaze, vocalization, or action, and the caregiver "returns" that serve by following the child's cues and responding to the child in a developmentally supportive way that caters to the child's needs (Shonkoff & Bales, 2011). These interactions are inherently rewarding and promote positive interactions that further support the child's cognitive growth (Fisher et al., 2016). Contingent interactions can be found in various communities and support the development of turn taking in early childhood socialization (Bornstein et al., 2015). Improving these interactions by increasing the amount of meaningful responses the parent makes to the

child is associated with shaping children's language and communication skills (Bornstein et al., 2008).

Sensitive parenting focuses on improving secure attachments between caregivers and infants by promoting parental warmth, noticing the child's actions, and responding in an affectionate and supportive manner (MacDonald, 1992). Parents that are sensitive to infant cues are more likely to foster secure attachments with their child which facilitate healthy development of self-reliance (Ainsworth, 1979). Secure attachment styles can promote positive internal working models in children which can lead to beneficial downstream effects. For example, Wolke and colleagues (2013) found that increased sensitive parenting by harmonious and supportive play can protect against the adverse effects of preterm and low birthweight on school success. They state since sensitivity is bidirectional and is adapted to the child's skills, they allow for self-regulation and scaffolding of children's attentional focus. Barnett et al. (2012) found that parent sensitivity by increased warmth and child-centered play was associated with increased social competence and language skills in children. These studies show that increased caregiver sensitivity can lead to stronger secure attachment and future positive outcomes such as language skills, school success, and socio-emotional development.

The FIND Intervention

The Filming Interactions to Nurture Development (FIND) Intervention is an intervention that aims to enhance responsive caregiving by providing strength-based video feedback to caregivers. This intervention incorporates both contingent responsiveness and sensitive parenting behaviors through individualized parent coaching sessions. By combining these two components of responsive caregiving, the FIND intervention teaches caregivers to pay attention to their child's "serves", follow their lead, and respond by "returning" the child's serve in a supportive

manner (Fisher et al., 2016). Coaches highlight five main types of serve and return interactions in coaching sessions, which are taught as the five core FIND elements: (1) sharing the focus; (2) supporting and encouraging; (3) naming; (4) back-and-forth; (5) endings and beginnings. This process allows caregivers to learn what behaviors are supportive to their child in a way that is encouraging.

The FIND theory of change predicts that by increasing the frequency of parent following behaviors, caregivers will form stronger attachment relationships with their child and reduce the impact of environmental stressors on the child's wellbeing (Fisher et al., 2016). Following the child's lead creates parent-child interactions that are optimal, sustained, and reciprocal, which provides a strong foundation towards healthy development and positive downstream effects for both the caregiver and the child (Shonkoff and Fisher, 2013).

What is Next: The Current Study

Preliminary evidence suggests that the FIND intervention improves caregiver self-efficacy, parental executive function, and child outcomes (Liu et al., 2021; Giuliani et al., 2019). However, we have yet to determine how the FIND intervention changes the way caregivers and children interact with each other. The FIND theory of change predicts that by strengthening "serve and return" interactions, caregivers will increase the amount of time they follow their child's lead, increasing the likelihood that both the caregiver and the child experience positive future outcomes (Fisher et al., 2016). This study aims to evaluate whether the FIND intervention increases the frequency of caregiver following behaviors. Since responsive caregiving is an important predictor of child development and positive outcomes, it is critical to understand how the FIND intervention changes the way caregivers interact with their children. This study analyzed the frequency of following behaviors during dyadic freeplay interaction filmed pre- and

post- the intervention period using a novel observational coding method to quantify caregiver following and leading behaviors. We predict that the FIND intervention will increase the percent of caregiver following behaviors; next steps will examine how changing dyadic interactions are correlated with other caregiver and child outcomes.

Methods

Participants

Participants were recruited from Early Head Start (EHS) programs located in the Denver, Colorado metropolitan area. To be eligible for screening, participants needed to be eligible to receive EHS services, have children between the age of 4 to 36 months old, and be fluent in either English or Spanish. Data for this study were drawn from a larger sample size ($n = 138$) collected from a large-scale RCT of the FIND intervention (for full recruitment procedures and sample size, see Liu et al. 2021); this study included only families from the larger RCT who spoke English during dyadic freeplay videos and who completed both pre- and post- intervention visits. 18 parent-infant dyads met all inclusion criteria and were included in this pilot trial: 7 families participated in an active control group, and 11 participated in the FIND intervention group. Families in the active control group received services from EHS without additional support, while families in the FIND intervention group completed 10 intervention sessions in addition to EHS services. All but one of the caregivers were mothers ($n = 17$, 94.4%) and the mean age of the caregivers was 32.50 years ($SD = 8.91$). There were 44.4% ($n = 8$) female children in the sample and the mean age of the children was 20.83 months ($SD = 8.75$). In the English-speaking videos, 22.2% of caregivers identified as Hispanic/Latino ($n = 4$).

The FIND Intervention

In order to improve parent-child interactions, the FIND intervention aims to increase "serve and return" interactions as well as following behaviors (Fisher et al., 2016). Participants within the FIND intervention received ten weekly home visits that alternate between filming dyadic freeplay and coaching sessions with the caregiver. In between recording and coaching sections, clips of freeplay were edited to highlight times when caregivers displayed the five FIND elements. The first element is *sharing the child's focus*, which focuses on supporting the caregiver in noticing what the child is actively focused on. The second element is *supporting and encouraging*, which occurs when a caregiver provides positive or encouraging feedback, including responding to the child's needs. The third element is *naming,* which occurs when the caregiver is labeling objects that are within the child's focus. The fourth element is *back and forth interaction*, which is when a caregiver notices the child's serve, returns that serve, and waits for the child's next serve. Finally, the fifth element is *endings and beginnings*, which encourages parents to notice when a child changes their focus from one activity ("ending") to start a new one ("beginnings"). The FIND elements were predicted to promote reciprocal and contingent interactions between caregivers and their children. Clips containing these elements were presented to caregivers during coaching sessions where FIND coaches highlighted positive "serve and return" behaviors the caregiver had displayed.

FLO Coding

To determine if caregiver following behaviors changed throughout the intervention period, a team of coders from the University of Oregon developed and followed a coding protocol to quantify caregiver following and leading behaviors. Videos used for the FIND intervention were coded into segments that differentiate when the caregiver was "following", "leading", or "other". A caregiver was coded as "following" when they were sharing the focus

with the child and were supplementing the child's play. A caregiver was coded as "leading" when they were taking charge of the play or changing the direction of the play. A caregiver was coded as "other" when they were not engaging with the child or were distracted with other toys, objects, or people. When the face and hands of both the caregiver and the child were out of the frame and undistinguishable, the video segment was coded as "no code" and was removed from the overall percentage of codable film (for the FLO Coding FLO Chart and Glossary, see Appendices A and B). Coders used Noldus Observer XT observational coding software to code films; analyses performed using Noldus Observer XT documented the percentage of the time the caregiver participated in following/leading/other behaviors for each freeplay film. Pre- and post-intervention videos were compared to see if there was a change in the percentage of following behaviors during freeplay across the intervention period.

Out of the 36 videos, 36% were double coded to determine inter-rater reliability (n = 13). Reliability was calculated using Noldus Observer XT in two ways: a) by duration and sequence, or how well the codes match at a particular timestamp, and b) by duration, or the similarity in the amount of time a particular code is used is between coders. Average inter-rater reliability for duration and sequence was 79.37%, κ = 0.54, and average reliability for duration was 82.86%, κ = 0.67, which met the threshold of acceptable inter-rater reliability set for this study.

Analytical Plan

Analyses were conducted using Jamovi version 2.0. First, independent samples t-tests were conducted to compare group differences in the percentage of following behaviors between the FIND group and the control group before the intervention and after the intervention. An independent samples t-test was also conducted to compare changes in following behaviors between the FIND and control groups. Next, a two-way ANOVA was conducted to examine pre-

and post- intervention effects on caregivers' following behaviors between groups. Finally, an ANCOVA analysis was done to control for the difference between groups pre-intervention. Residual analyses were performed to test for the assumptions of the two-way ANOVA. Outliers were assessed by inspection of a boxplot, normality was assessed using Shapiro-Wilk's normality test, and homogeneity of variances was assessed by Levene's test.

Results

Independent samples t-tests were first conducted to compare group differences in following behaviors before and after the intervention. There was a difference in the percentage of following behaviors between the FIND ($M = 0.45$, $SD = 0.23$) and control ($M = 0.56$, $SD = 0.25$) conditions pre- intervention, although an independent samples t-test revealed this difference was not significant ($t(16) = 0.11$, $p = 0.333$). An independent samples t-test also showed that there was no significant difference in following behaviors between control ($M = 0.57$, $SD = 0.26$) and FIND groups ($M = 0.63$, $SD = 0.27$) after the intervention period ($t(16) = -0.06$, $p = 0.670$). See Table 1 for a summary of means and standard deviations.

A two-way ANOVA was conducted to examine the effects of condition (FIND or control), time (PRE or POST intervention), and the interaction between the two on following behaviors. The main effect of condition was not significant, $F(1, 32) = 0.11$, $p = 0.738$, $\eta^2 = 0.00$, as well as the main effect of time, $F(1, 32) = 1.25$, $p = 0.272$, $\eta^2 = 0.04$. Families who received the FIND intervention increased in following behaviors across the intervention period ($M_{diff} = 0.18$, $SD = 0.27$) while the control group did not show change across the intervention period ($M_{diff} = 0.01$, $SD = 0.20$); notably, the two-way ANOVA revealed no significant interaction between condition and time, $F(1, 32) = 0.97$, $p = 0.331$, $\eta^2 = 0.03$ (see Figure 1).

Follow-up analysis using an ANCOVA to control for the difference between groups at pre-intervention revealed that pre-intervention scores significantly predicted following behaviors post-intervention, $F(1, 15) = 5.36$, $p = 0.035$, $\eta^2 = 0.25$. On the other hand, the condition the families were in did not significantly predict their post- percent following behaviors $F(1, 15) = 1.07$, $p = 0.317$, $\eta^2 = 0.05$). See Table 2 for ANCOVA results.

Discussion

Results from t-tests indicated no significant difference between groups before the intervention and after the intervention. T-tests also showed no significant difference in the change in following behaviors between groups. While there was an increase in following behaviors in the FIND group and no change in the control group, results from ANOVA analyses revealed that the main effect of condition, time, and the interaction were nonsignificant. Lastly, an ANCOVA analysis showed that pre-intervention percent following behaviors significantly predicted post-intervention following behaviors; this means that lower percent following behavior before the intervention period were likely to lead to higher percent following behavior after the intervention period.

The results of this study suggest two potential possibilities. Our ANCOVA analyses revealed that pre-intervention following behaviors significantly predict the percentage of following behaviors at post-; specifically, families with lower percentages of following behaviors at baseline were more likely to see an increase in following behaviors across the intervention period. In other words, a lower percentage of following pre-intervention predicts a higher percentage of following post-intervention: a regression to the mean effect. Because results of the FIND group had overall lower percentages of following behaviors pre-intervention compared to

the control group, their increase in percent following behaviors across the intervention period

may simply be a regression to the mean and not an intervention effect.

A second possibility follows our hypothesized mechanism of change for the FIND

intervention: that this intervention increases caregiver following behaviors over time. This study

predicted that percent following behaviors in the FIND intervention group would increase across

the intervention period compared to the control group. The results were consistent with this

prediction and showed increased following behaviors in the FIND group (M_{diff} = 0.18, SD =

0.27) while the control group showed little to no change in following behaviors (M_{diff} = 0.01, SD

= 0.20). Although the difference between groups was not significant (p = 0.331), the results from

the ANCOVA (pre- scores significantly predicting post- scores) may highlight an important

intervention effect that is difficult to differentiate with a small sample size and unequal baseline

scores. Since these results may be affected by this study's limitations, a larger sample size is

required to differentiate which of these possibilities is likely true for this specific intervention.

This study had several limitations that impact the implications of these findings. First,

because this was a pilot trial of a novel behavioral coding scheme, we utilized only a small

sample of English-speaking families from a larger intervention trial. Our small sample of 18

English-speaking families allowed us to examine trends in behavioral change but does not have

enough power to draw conclusions about intervention effects. Completing this study – using the

same methodology with the full sample of English and Spanish-speaking families who

completed the intervention – will allow us to better differentiate whether following behaviors are

truly impacted by the FIND intervention. Secondly, our research team experienced some

challenges with inter-rater reliability as a part of the early-stage development of the FLO coding

scheme. Part of this study included pilot testing of the FLO Coding Tool, which is still in the

early stages of development. While reliability among coders was strong for this specific study, we had not yet finalized the development of the coding scheme for use with Spanish-speaking families. This contributed to why this study was only able to use a small pilot sample of English-speaking families from the larger study. The next steps for the FLO coding measure include translating and adapting this coding scheme for use with Spanish-speaking families. Future analyses should include videos from both English and Spanish-speaking families, and video coders fluent in Spanish would contribute to coding the Spanish films.

Overall, this study suggests that the FIND intervention may increase caregiver following behaviors over the course of the intervention period. Increases in following behaviors were observed for the FIND intervention group, but not for the control group, although the interaction effect of condition by time was not significant. We also found that parents who displayed lower following behaviors before the intervention were more likely to show higher following behaviors after the intervention. The small sample size included in this pilot trial makes it difficult to determine whether these effects stem from regression to the mean or if the FIND intervention may be most effective for caregivers who tend to dominate or lead interactions with their children (i.e. those with fewer following behaviors at the start of the intervention). This study offers the first evidence of quantifying and evaluating responsive caregiving within the FIND intervention. Although this was a pilot trial, the results indicate a possible way to evaluate changes in caregiver behavior and parent-child interactions. Continuing this study using a larger sample size should clarify the extent to which the FIND intervention directly impacts these changes in caregiver behaviors.

References

Ainsworth, M. D. S. (1979). Infant–mother attachment. *American Psychologist, 34*(10), 932–
        937. https://doi-org.libproxy.uoregon.edu/10.1037/0003-066X.34.10.932

Ainsworth, M.D.S., Bell, S.M., & Stayton, D.J. (1974) Infant-mother attachment and social
        development: 'Socialization' as a product of reciprocal responsiveness to signals. In: *The
        integration of a child into a social world* (ed. M.P.M. Richards), pp. 99–135. Cambridge
        University Press, Cambridge, UK.

Barnett, M. A., Gustafsson, H., Deng, M., Mills-Koonce, W. R., & Cox, M. (2012). Bidirectional
        Associations Among Sensitive Parenting, Language Development, and Social
        Competence. *Infant and child development*, *21*(4), 374–393. https://doi-
        org.libproxy.uoregon.edu/10.1002/icd.1750

Blair, C., & Raver, C. C. (2016). Poverty, Stress, and Brain Development: New Directions for
        Prevention and Intervention. *Academic pediatrics*, *16*(3 Suppl), S30–S36.
        https://doi.org/10.1016/j.acap.2016.01.010

Bornstein, M. H., Putnick, D. L., Cote, L. R., Haynes, M., & Suwalsky, J. T. D. (2015). Mother-
        infant contingent vocalizations in 11 countries. *Psychological Science, 26*(8), 1272-1284.

Bornstein, M. H., Tamis-LeMonda, C. S., Hahn, C. S., & Haynes, O. M. (2008). Maternal
        responsiveness to young children at three ages: Longitudinal analysis of a multidimensional,
        modular, and specific parenting construct. *Developmental Psychology, 44*(3), 867-74.

Bourne, S. V., Korom, M., & Dozier, M. (2022). Consequences of inadequate caregiving for
        children's attachment, neurobiological development, and adaptive functioning. *Clinical Child
        and Family Psychology Review.* Advance online publication. https://doi-
        org.libproxy.uoregon.edu/10.1007/s10567-022-00386-4

Bruce, J., Gunnar, M. R., Pears, K. C., & Fisher, P. A. (2013). Early adverse care, stress
     neurobiology, and prevention science: Lessons learned. *Prevention Science*, 14, 247–256.
     doi:10.1007/s11121-012-0354-6

Evans G. W. (2006). Child development and the physical environment. *Annual review of
     psychology*, *57*, 423–451. https://doi.org/10.1146/annurev.psych.57.102904.190057

Gilmore, J. H., Knickmeyer, R. C., & Gao, W. (2018). Imaging structural and functional brain
     development in early childhood. *Nature reviews. Neuroscience*, *19*(3), 123–137.
     https://doi.org/10.1038/nrn.2018.1

Giuliani NR, Beauchamp KG, Noll LK and Fisher PA (2019) A Preliminary Study Investigating
     Maternal Neurocognitive Mechanisms Underlying a Child-Supportive Parenting
     Intervention. *Front. Behav. Neurosci.* 13:16. doi: 10.3389/fnbeh.2019.00016

Gunnar M. R. (1998). Quality of early care and buffering of neuroendocrine stress reactions:
     potential effects on the developing human brain. *Preventive medicine*, *27*(2), 208–211.
     https://doi.org/10.1006/pmed.1998.0276

Landry, S. H., Smith, K. E., & Swank, P. R. (2006). Responsive parenting: establishing early
     foundations for social, communication, and independent problem-solving
     skills. *Developmental psychology*, *42*(4), 627–642. https://doi.org/10.1037/0012-
     1649.42.4.627

Levine, S. (2005). 1-Stress: An historical perspective. *Techniques in the Behavioral and Neural
     Sciences*, 15, 3–23. doi:10.1016/S0921- 0709(05)80004-5

Liu, S., Phu, T., Dominguez, A., Hurwich-Reiss, E., McGee, D., Watamura, S., & Fisher, P.
     (2021). Improving Caregiver Self-Efficacy and Children's Behavioral Outcomes via a Brief
     Strength-Based Video Coaching Intervention: Results from a Randomized Controlled

Trial. *Prevention science : the official journal of the Society for Prevention Research*, 10.1007/s11121-021-01251-6. Advance online publication. https://doi.org/10.1007/s11121-021-01251-6

Lucas, J. E., Richter, L. M., & Daelmans, B. (2018). Care for Child Development: An intervention in support of responsive caregiving and early child development. *Child: Care, Health and Development, 44*(1), 41–49. https://doiorg.libproxy.uoregon.edu/10.1111/-cch.12544

MacDonald, K. (1992). Warmth as a Developmental Construct: An Evolutionary Analysis. Child Development, 63(4), 753–773. https://doi.org/10.2307/1131231

Mihelic, M., Morawska, A., & Filus, A. (2017). Effects of early parenting interventions on parents and infants: A meta-analytic review. *Journal of Child and Family Studies, 26*(6), 1507–1526. https://doi-org.libproxy.uoregon.edu/10.1007/s10826-017-0675-y

Miller, G. E., Brody, G. H., Yu, T., & Chen, E. (2014). A family-oriented psychosocial intervention reduces inflammation in low-SES African American youth. Proceedings of the National Academy of Sciences of the United States of America, 111, 11287–11292. doi:10.1073/pnas.1406578111

Scherer, E., Hagaman, A., Chung, E. et al. (2019). The relationship between responsive caregiving and child outcomes: evidence from direct observations of mother-child dyads in Pakistan. *BMC Public Health, 19*, 252. https://doi.org/10.1186/s12889-019-6571-1

Shonkoff, J. P., & Bales, S. N. (2011). Science does not speak for itself: translating child development research for the public and its policymakers. *Child development, 82*(1), 17–32. https://doi-org.libproxy.uoregon.edu/10.1111/j.1467-8624.2010.01538.x

Shonkoff, J. P., Boyce, W. T., & McEwen, B. S. (2009). Neuroscience, molecular biology, and

    the childhood roots of health disparities: Building a new framework for health promotion and

    disease prevention. *Journal of the American Medical Association, 301*, 2252–2259.

    doi:10.1001/jama.2009.754

Shonkoff, J. P., & Fisher, P. A. (2013). Rethinking evidence-based practice and two-generation

    programs to create the future of early childhood policy. *Development and Psychopathology,*

    25, 1635–1653. doi:10.1017/S0954579413000813

Wolke, D., Jaekel, J., Hall, J., & Baumann, N. (2013). Effects of sensitive parenting on the

    academic resilience of very preterm and very low birth weight adolescents. *Journal of*

    Adolescent Health, 53(5), 642–647. https://doiorg.libproxy.uoregon.edu/10.1016/-

    j.jadohealth.2013.06.014

Table 1

Descriptive Statistics of FIND and Control Groups Pre and Post Intervention

Condition	n	Mean % Following	Standard Deviation
Pre/Control	7	0.56	0.25
Post/Control	7	0.57	0.26
Pre/FIND	11	0.45	0.23
Post/FIND	11	0.63	0.27

Note. This table includes the relevant means and standard deviations of following behaviors for both the control and FIND conditions before and after the intervention period.

Figure 1

Change in Following Behaviors for Control and FIND Groups Before and After the Intervention Period
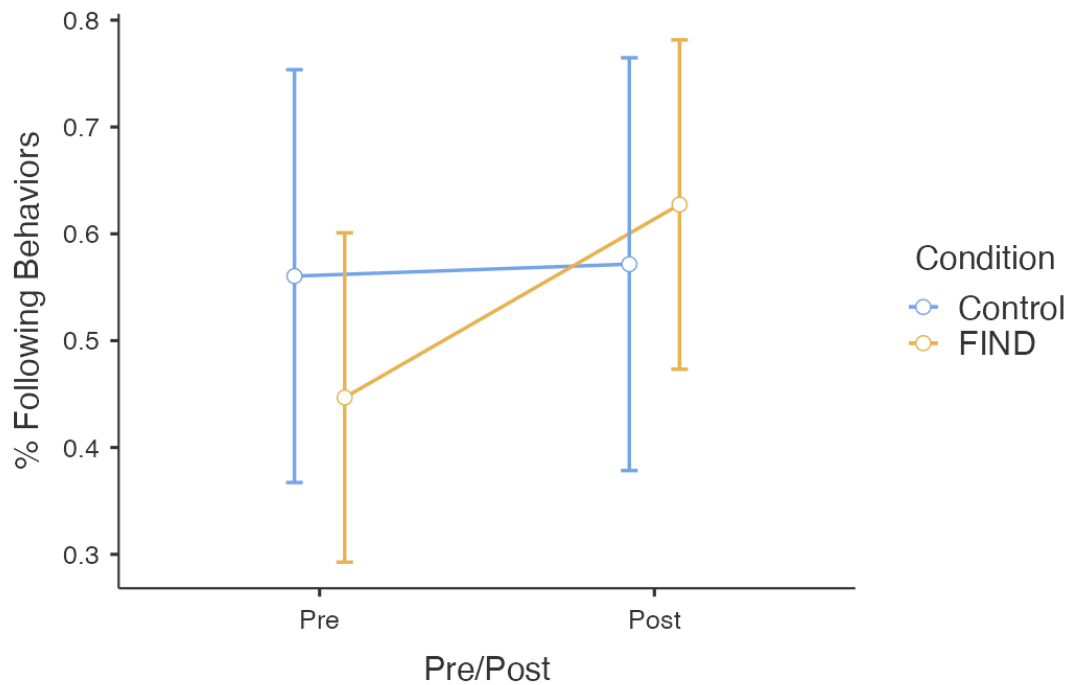
Note. Results showed a promising trend: The FIND intervention group increased in percent following behaviors (M_{diff} = 0.18, *SD* = 0.27), while the control group showed no change in following behaviors across the intervention period (M_{diff} = 0.01, *SD* = 0.20). ANOVA results were nonsignificant.
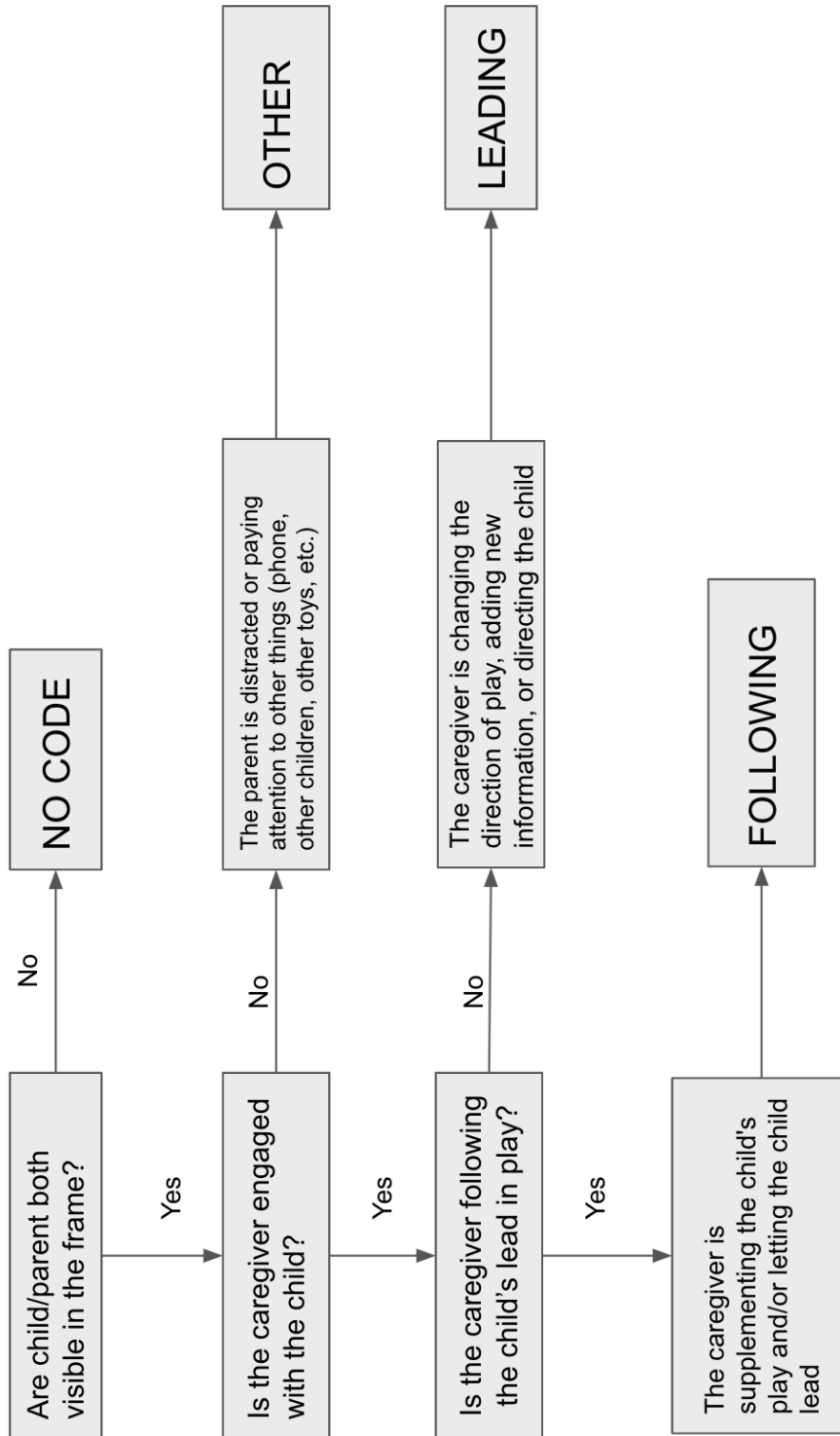
Table 2

ANCOVA Analysis Results

	Sum of Squares	df	Mean Square	F	p	η^2
Overall model	0.36	2	0.18	2.80	0.093	
Condition	0.06	1	0.06	1.07	0.317	0.05
% Following Pre	0.30	1	0.30	5.36	0.035	0.25
Residuals	0.83	15	0.06			

Note. This table demonstrates the results of ANCOVA analyses. Results revealed that pre-intervention scores significantly predicted following behaviors post-intervention. The condition the families participated in did not significantly predict their post- percent following behaviors

Appendix A

FLO Coding Flowchart

```
Are child/parent both          No
visible in the frame?    ────────────────→   NO CODE

        │ Yes
        ▼

Is the caregiver engaged       No        The parent is distracted or paying
with the child?          ────────────────→   attention to other things (phone,    ──→   OTHER
                                             other children, other toys, etc.)

        │ Yes
        ▼

Is the caregiver following     No        The caregiver is changing the
the child's lead in play? ────────────────→   direction of play, adding new        ──→   LEADING
                                             information, or directing the child

        │ Yes
        ▼

The caregiver is
supplementing the child's  ────────────────→   FOLLOWING
play and/or letting the child
lead
```

Appendix B

FLO Coding Glossary

Overall, this coding scheme aims to quantify *caregiver behaviors* within the context of dyadic freeplay interactions. During play, we consider the child's "spotlight" of attention. The determination of "following" and "leading" refers to whether the caregiver is in control of the child's spotlight.

This glossary is designed to complement the FLO Coding flowchart. Coders should follow coding decisions outlined by the flowchart. Definitions of each behavior are outlined in more detail below.

Following (F):

- Caregiver is attentively engaged with the child and paying attention to what the child is doing (i.e., the child is in control of the "spotlight")
  - Caregiver is sharing the child's focus
  - Caregiver is not physically controlling the play (e.g., they may be sitting back to let the child lead)
- Caregiver is supplementing the child's activity without changing the direction of play
  - Examples:
    - Caregiver naming something the child is doing/playing with
    - Responding to a child's serve without changing the direction of play
    - Caregiver is curious about something the child is doing (e.g., "wondering")

Leading (L):

- Caregiver is actively engaged in directing the trajectory of the child's play (i.e., the caregiver is in control of the "spotlight")
  - Caregiver may be physically controlling/manipulating the toy or the play
  - Caregiver may be "teaching" or guiding the child's actions
  - Caregiver may be directing the child through commands
    - E.g., limit setting ("No") would fall into this category
  - Caregiver may change the direction of play from one thing to another
  - Caregiver may add things to the play that the child is currently not focused on
  - Caregiver may be diverting the child's attention away from their current interest

Other (O):

- Caregiver is disengaged or not interacting with child; caregiver is not sharing the child's focus
  - Examples:
    - Caregiver using their phone or distracted by something else
    - Caregiver is talking to someone else (another child, researcher)
    - Caregiver and child are interacting with different toys (parallel play)

Not Codable (N)

- If any part of the interaction is not visible/audible and the quality of the film makes it too difficult to make a coding decision, the coder should choose "No Code"
  - Both caregiver and child need to be in the frame
    - If either child or caregiver is off-screen it should be coded "No Code"
    - Caregiver's face needs to be visible in order for a video to be considered "codeable"

  - Interaction between caregiver and child needs to be audible

OXYTOCIN: A POTENTIAL PATHWAY FOR THE

INTERGENERATIONAL IMPACTS OF EARLY TRUAMA

by

GIOVANNI RICCI

A THESIS

Presented to the University of Oregon Department of Psychology
in partial fulfillment of the requirements for the degree of
Bachelor of Arts with departmental honors

July 2022

Approved: _____*Title and Full Name (or Title at end of Full Name)*_____
Primary Thesis Advisor

Abstract

This study examined the correlations between early childhood adversity, baseline salivary oxytocin, and self-reported maternal-child bonding and attachment in a population of mothers of infants. The aim of this study was to use a new combination of measures and a novel salivary oxytocin immunoassay to provide support for previous findings on (1) the association between experiences of early childhood adversity and salivary oxytocin in mothers of infants, (2) the association between maternal salivary oxytocin and mothers' perceived bonding and attachment with their child, (3) examine oxytocin as a mediator between the effects of early childhood trauma on maternal-child bonding and attachment, and to (4) propose a conceptual model where oxytocin acts as a pathway for the intergenerational transmission of trauma, particularly through its role in promoting healthy maternal-child bonding and attachment. Results of a multiple regression analysis demonstrated mothers' experiences of early childhood adversity were not correlated with baseline salivary oxytocin, mothers' adverse childhood experiences did not significantly predict their self-reported bonding and attachment with their infants, and maternal baseline salivary oxytocin levels did not significantly predict their self-reported bonding and attachment with their infants. This calls into question the strength of the associations suggested in prior research regarding early adversity, oxytocin, and maternal-child bonding and attachment. Future research examining oxytocin as a pathway for the intergenerational impacts of early trauma through the mother-child relationship would benefit from examining the role of factors such as the level of severity of early adversity, the type of adversity experienced, and the effects of such variables on specific components of bonding and attachment behavior rather than bonding and attachment as a whole.

Acknowledgements

Table of Contents

Trauma, especially early life trauma, significantly impacts the health and well-being of those who experience it. And yet, the scars trauma can leave on individuals are often experienced by future generations in various ways. Emerging research elucidates one potential pathway for the intergenerational impacts of trauma: oxytocin and the maternal-child relationship (Donadon et al., 2018; Scatliffe et al., 2019; Szymanska et al., 2017). Oxytocin is an evolutionarily ancient peptide hormone and neuropeptide produced by the hypothalamus and released by the pituitary gland. As such, oxytocin plays an important role in many aspects of human life and can have lasting impacts on human development. Along with several physiological effects within the body, research has shown oxytocin is involved in regulating various social behaviors particularly within the maternal-child relationship. Most notably, oxytocin is involved with regulating maternal bonding and the facilitation of healthy maternal-child attachment (Feldman et al., 2010a; Scatliffe et al., 2019; Szymanska et al., 2017). However, research also suggests early or chronic trauma may impair oxytocin production later in life, and decreased oxytocin has been associated with experiences of childhood abuse, maltreatment, and dysfunctional parent-infant relationships (Donadon et al., 2018; Szymanska et al., 2017). As such, research suggests oxytocin may be one pathway in which the effects of trauma can be transmitted to future generations through the impact early childhood trauma can have on oxytocin production later in life and the role oxytocin may play as a mediator for the effects of early childhood trauma on maternal-child bonding and attachment in a way that promotes the continuation of this cycle across generations. However, the majority of research on the association among oxytocin, trauma, and maternal-child bonding and attachment has not used salivary oxytocin as a measure—partly due to past challenges in salivary oxytocin measures and debates over their efficacy (McCullough et al., 2013)—and salivary oxytocin has very rarely been used in combination with the widely supported Adverse Childhood Experiences Scale (ACES).

The present study aims to provide further evidence for oxytocin as a pathway for the intergenerational impacts of trauma by examining the association between early life adversity and impaired oxytocin production in moms of infants and examining the role oxytocin plays as a mediator of the effects of early childhood adversity on maternal-child bonding and attachment. Specifically, this study aims to bridge this gap in research through the use of a novel salivary oxytocin immunoassay in combination with the ACE questionnaire to propose a model that recognizes the role oxytocin may play in the intergenerational transmission of early life trauma and adversity with a particular emphasis on the mediating effects of oxytocin on early adversity and maternal-child bonding and attachment. See Figure 1.

Impact of Early Childhood Adversity on Oxytocin Production Later in Life

Trauma, particularly chronic or childhood trauma, may impair oxytocin production later in life. A systematic review by Donadon et al. (2018) reviewed 35 studies that examined the relation between emotional trauma and endogenous oxytocin production and found that emotional trauma and the presence of PTSD symptoms was moderately associated with decreased oxytocin levels. These associations were particularly strong in cases where emotional trauma occurred during early childhood. Similarly, studies have shown childhood abuse, maltreatment, and adverse life experiences during childhood are correlated with decreased endogenous oxytocin levels later in life regardless of gender. Specifically, Heim et al. (2008) demonstrated that women exposed to emotional abuse, physical abuse, and emotional neglect during childhood had decreased oxytocin concentrations in cerebrospinal fluid—with a particularly strong effect of emotional abuse—and greater experiences of overall maltreatment associated with decreased oxytocin levels. Similar results have been demonstrated in men and women who experienced significant early life stress

during childhood using plasma oxytocin measures (Bertsch et al., 2013; Opacka-Juffry & Mohiyeddini, 2011), with Bertsch et al. (2013) demonstrating a particularly strong effect due to emotional abuse and neglect during childhood. Interestingly, results from Mizuki and Fujiwara (2015) showed basal oxytocin levels may depend on the intensity of adverse childhood events, as adult participants with less severe forms of childhood physical abuse showed higher basal oxytocin levels in urine than participants reporting no physical abuse during childhood, contrasting previous findings. The researchers theorized that this may be due to elevated oxytocin levels acting as a defense mechanism that helps manage maltreatment and social stress by promoting 'tend-and-befriend' behaviors to mitigate development of psychopathology in response to less severe forms of trauma (Mizuki & Fujiwara, 2015), a mechanism which may become overwhelmed in the presence of more severe trauma and adversity. Overall, research using peripheral measures of oxytocin in urine and plasma—as well as a smaller subset of studies using cerebrospinal fluid and salivary oxytocin measures—suggests experiences of trauma and adversity during childhood likely reduce oxytocin production later in life.

Why might we see this association between early trauma and adversity and oxytocin levels? Existing research suggests the impact trauma has on decreased oxytocin production later in life is likely due to long-term impairments in hypothalamic structures and negative feedback mechanisms within the hypothalamic-pituitary-adrenal (HPA) axis. Specifically, oxytocin plays a crucial role in the negative feedback mechanisms of cortisol and helps ensure a return to baseline cortisol levels after exposure to psychologically stressful stimuli (Amico et al., 2004; Gulpinar & Yegen, 2004; Heinrichs et al., 2004). As summarized by Donadon et al. (2018), under situations of chronic stress especially during early childhood, the functioning of the suprachiasmatic nucleus may become impaired over time through the prolonged presence of excess cortisol, as seen in conditions

such as hypercortisolemia (Gonzalez et al., 2009), which likely decreases the synthesis and release of oxytocin well into adulthood (Brown et al., 2016). Behaviorally, oxytocin has several prosocial functions that can reduce the neurophysiological and neurochemical impacts of stress on the brain through the facilitation of resiliency and coping strategies in response to adversity. As mentioned previously, the dysregulation of the HPA axis due to a prolonged cortisol response likely reduces endogenous oxytocin synthesis and release, reducing resilience and coping mechanisms associated with oxytocin, ultimately contributing to the continued buildup of cortisol and impaired HPA axis functioning. Overall, research suggests early life adversity may reduce oxytocin production later in life through long-term impairments in HPA axis functioning and behavioral pathways involving coping and resiliency that aggravate these effects. However, the impact of trauma does not stop at the individual level, particularly when considering the role oxytocin plays in maternal-child bonding and attachment.

Oxytocin and Maternal-Child Bonding and Attachment

In addition to several physiological effects within the body, research has established the role oxytocin plays in interpersonal relationships, particularly between a mother and child. Specifically, studies have shown oxytocin is involved with parental sensitivity—defined as the ability to recognize, interpret, and respond appropriately to a child's cues during distress—which is an important determinant of maternal-child attachment (Szymanska et al., 2017). Research using self-report (Feldman et al., 2007), neuroimaging (Atzil et al., 2011), and behavioral measures (Feldman et al., 2010a) have reliably demonstrated positive correlations between parental sensitivity and oxytocin levels in mothers, with higher oxytocin levels associated with greater parental sensitivity. Studies using exogenous oxytocin administration suggest a causal association

between the two (Bakermans-Kranenburg et al., 2011). Interestingly, recent studies have demonstrated conflicting results, with one study showing that mothers who were characterized as low in parental sensitivity demonstrated higher baseline plasma oxytocin compared to mothers characterized as high in parental sensitivity (Elmadih et al., 2014), and a similar study found that while plasma oxytocin overall was not associated with parental sensitivity, for mothers who reported high levels of psychosocial stress, higher levels of plasma oxytocin were associated with higher parental sensitivity (Zelkowitz et al., 2014). The authors argued that oxytocin may act as a buffer against the deleterious effects of stress, promoting more sensitive maternal interactive behavior for high risk women specifically. Overall, while more research is needed given the conflicting results in recent literature, the role oxytocin plays in parental sensitivity likely influences caregiving quality due to the importance of appropriate detection and recognition of children's signals in times of distress.

Additionally, oxytocin has been shown to play an important role in synchrony between a mother and child. Synchrony can be defined as the dynamic process during social contact between parent and child in which hormonal, physiological, and behavioral cues are exchanged and coordinated (Atzil et al., 2013), and likely influences the quality of maternal-child bonding and attachment (Szymanska et al., 2017). Several studies have demonstrated maternal oxytocin levels predict the quality of maternal postpartum behavior and its synchronization with infant state as well as increased affectionate contact involved in healthy bonding and attachment. Specifically, higher oxytocin levels have been shown to predict increased maternal-child synchrony as measured by plasma oxytocin and parent-infant interactions examining frequency and duration of synchronous gaze and vocalizations, positive affect, and affectionate contact (Apter-Levi et al., 2014; Feldman et al., 2007; Gordon et al., 2010), which has also been supported by research from

Feldman et al. (2010b) using an enzyme-linked immunosorbent assay (ELISA) method with salivary oxytocin. Experimental studies using intranasal oxytocin administration have shown an increase in synchronous behaviors after administration, such as social reciprocity, reciprocal parent-infant gaze, and parent-infant touch synchrony, suggesting a causal association between oxytocin and synchrony (Weisman et al., 2012).

Not only does oxytocin appear to be involved with various processes that promote bonding and attachment, it has been shown to influence bonding and attachment more directly. Bonding can be defined as the time in which mothers bond with their offspring, particularly during the first days after birth, and is characterized by a mother's perception of a unique bond with their child and feelings of responsibility and a desire for protection (Szymanska et al., 2017). Maternal-child attachment, by comparison, can be more broadly defined as an emotional bond between a mother and infant. However, both processes are intimately related and often used interchangeably in research. The months preceding childbirth are thought to be a particularly sensitive period for maternal-fetal bonding, and studies have shown an increase in oxytocin levels during this period is associated with greater maternal-infant bonding (Levine et al., 2007). Similarly, oxytocin levels have been shown to increase immediately following childbirth and are involved in neuroendocrine processes that stimulate maternal behavior (Nissen et al., 1995), particularly maternal behavior that promotes healthy maternal-infant bonding and attachment behaviors such as gaze, vocalizations, positive affect, affectionate touch, attachment-related thoughts, and frequent checking of the infant (Feldman et al., 2007). As summarized in Szymanska et al. (2017), during this period oxytocin appears to promote maternal-infant bonding and attachment processes by decreasing the stress response in mothers, enhancing a state of calmness immediately following birth, inhibiting amygdala function related to the processing of fear and threat stimuli, and possibly

decreasing negative attributions and anger towards the infant (Dębiec, 2005; Feldman et al., 2007; Kirsch, 2005). While studies have shown increases in oxytocin production following affectionate behavior between mother and infant, experimental studies using intranasal oxytocin administration have shown increases in brain activity involved with emotional regulation as well as decreased amygdala activation during exposure to a child's cues during states of distress (Riem et al., 2011a). This suggests attachment bonds help overcome social distance with offspring by decreasing critical social assessment and negative emotions, while simultaneously bonding individuals through reward circuitry in the brain (Bartels & Zeki, 2004). Experimental designs using oxytocin administration have also been shown to reduce negative emotional arousal in response to a child's demands and to promote responsiveness and increased proximity with a child (Riem et al., 2011b). Such experimental manipulation of participants' oxytocin levels through intranasal administration and subsequent behavioral observations suggests an increase in maternal-child bonding and attachment behavior in response to heightened oxytocin levels rather than the other way around.

Research specifically examining attachment has shown oxytocin in mothers and fathers was associated with parent and child's social engagement, affect synchrony, and positive communicative sequences between parent and child, and oxytocin levels were shown to be associated with mothers' and fathers' attachment relationships throughout life—to individuals' own parents, partners, and infants (Feldman et al., 2010a). However, some research suggests individual attachment styles might modulate the effects of oxytocin as well as endogenous oxytocin levels. Specifically, individuals with adverse attachment experiences during early childhood have shown decreased activation in brain reward regions after being exposed to images of their own infant's faces and lower oxytocin levels in response to infant touch (Strathearn et al., 2009), while individuals characterized by secure attachment who reported less suffering from

adverse life events have demonstrated heightened oxytocin levels (Tops et al., 2007). This suggests that in times of stress, healthy attachment may increase oxytocin levels and thus mediates oxytocin's stress-buffering effects in ways that are not seen in poor attachment relationships (Fang et al., 2014). In other words, while studies have shown heightened oxytocin levels may promote healthy maternal-child bonding and attachment—a likely causal association given results of experimental studies using intranasal oxytocin administration—there likely also exists a reciprocal association between oxytocin and attachment, where individual attachment styles may also influence oxytocin. More specifically, individual attachment styles may influence oxytocin's role in the activation of brain reward regions in response to images of one's own offspring as well as the oxytocin response to infant touch (Fang et al., 2014; Strathearn et al., 2009; Tops et al., 2007). Overall, the larger body of research suggests oxytocin plays a crucial role in facilitating processes that promote healthy maternal-child bonding and attachment, and while individual attachment styles may modulate specific functions of oxytocin, impaired oxytocin production in particular likely harms healthy maternal-child bonding and attachment.

Impaired Bonding and Attachment as a Risk Factor for Adverse Childhood Experiences

As highlighted above, oxytocin likely plays an important role in healthy maternal-child bonding and attachment, and experiences of early adversity likely reduce oxytocin levels later in life. However, evidence suggests impaired maternal-child bonding and attachment likely acts as a risk factor for adverse childhood experiences, implicating oxytocin as a potential pathway for the intergenerational impacts of early trauma. Specifically, healthy attachment to a caregiver acts as a crucial protective factor from the development of psychopathology and contributes to a child's resilience to adverse experiences (Kennison & Spooner, 2020; Miranda et al., 2012; Oldfield et

al., 2018; Stacy, 2006). This may be due to the role healthy attachment plays in providing emotional support to offspring in the presence of adversity and the potential damage insecure attachment can have on mental health outcomes of children. However, poor attachment itself can be a source of early adversity. Not only does a lack of healthy attachment act as a risk factor for psychopathology, but poor attachment relationships between a mother and child are more likely to be characterized by adverse experiences such as abuse, neglect, and household dysfunction (Erozkan, 2016). In other words, poor maternal-child bonding and attachment relationships characterized by a lack of sensitivity to a child's needs, less maternal warmth and affirmation, and increased negativity toward one's infant acts as a risk factor for multiple types of adverse childhood experiences, often with life-long implications (Murray et al., 1996; Pickreign Stronach et al., 2011)

One primary mechanism this may occur through likely involves resiliency. Resilience can be defined as the ability to thrive in the presence of adversity through adaptive psychological and physiological stress responses that allow individuals to maintain internal feelings of control and a sense of purpose and social connectivity (Sharma et al., 2020). The role oxytocin plays in pair bonding and attachment has been shown to reduce the neurophysiological and neurochemical effects trauma has on the brain by facilitating healthy psychological and social attachment, which likely increases resilience to traumatic events. For example, it has been well documented oxytocin plays an important role in resilience to adverse events due to its role in maintaining homeostasis within the HPA axis by counteracting increases in cortisol after exposure to stressful stimuli (Amico et al., 2004; Gulpinar & Yegen, 2004; Heinrichs et al., 2004). Additionally, oxytocin promotes resilience through its role in facilitating social attachment by enhancing the reward value of social stimuli and reducing potential fear responses that contribute to isolative and socially

anxious behavior that can occur in the aftermath of stress and trauma (Sharma et al., 2020). In other words, in the presence of stress and trauma oxytocin can promote individuals to seek emotional support from peers and loved ones, helping reduce the impacts of stress and increase resilience through social support and connection. Conversely, reduced oxytocin levels may act as a risk factor for adverse childhood experiences by impairing the ability for oxytocin to enable a return to baseline cortisol levels within the HPA axis—thus increasing the physiological impacts trauma can have on the brain—as well as reduce the seeking of social support and connection that would otherwise buffer against the effects of early trauma and adversity.

Oxytocin as a Pathway for the Intergenerational Impacts of Early Childhood Trauma

The prior sections have attempted to outline a model that can help account for the intergenerational transmission of early trauma. When considering the overall body of research, the literature suggests that (1) experiences of early life adversity can impair oxytocin production later in life, (2) reduced oxytocin levels likely impair healthy maternal-child bonding and attachment, and (3) impaired maternal-child bonding attachment can decrease child resilience to traumatic events as well as (4) act as a direct risk factor for adverse childhood experiences, ultimately promoting impaired oxytocin production in offspring of mothers with experiences of early life trauma and adversity. When considered holistically, this suggests oxytocin may function as one pathway for the intergenerational impacts of trauma in a way that promotes recurring impairments to oxytocin production across multiple generations in the absence of protective factors or intervention. See Figure 1.

However, despite this growing body of literature the vast majority of research on the association between oxytocin and early childhood trauma has not made use of salivary oxytocin

measures. For example, of the 35 studies on the relation between emotional trauma and endogenous oxytocin production reviewed by Donadon et al. (2018), only three included salivary oxytocin measures through the use of enzyme-linked immunosorbent assay (ELISA). This is partly due to the relative novelty and past challenges in salivary oxytocin measures and debates over their efficacy. Most notably, weak correlations with unextracted plasma oxytocin, the presence of biologically similar molecules that may be potentially tagged as oxytocin in unextracted samples, sensitivity challenges due to the small levels of oxytocin found in the periphery, and the questionable ability to estimate centrally released oxytocin from peripheral measures in general as documented by Martins et al. (2020) and McCullough et al. (2013). And yet, the relatively few studies that have used salivary oxytocin measures in relation to trauma and parenting behavior have generally demonstrated similar findings as other peripheral oxytocin measures (primarily through the use of ELISA with unextracted samples; Scatliffe et al., 2019), although there have been conflicting results (Julian et al., 2017; Martins et al., 2020; McCullough et al., 2013). This suggests more research is needed and would greatly benefit from a more accurate measure of salivary oxytocin than what was previously available. Additionally, salivary oxytocin measures have very rarely been used in combination with the Adverse Childhood Experiences Scale (ACES), a widely supported and well validated questionnaire for measuring childhood maltreatment and household dysfunction (Osofsky et al., 2021), which would add to this body of literature through its well documented associations with numerous negative health outcomes.

The aim of this study is to (1) examine the association between experiences of early childhood adversity and salivary oxytocin in mothers of infants, (2) examine the association between maternal salivary oxytocin and subsequent perceived bonding and attachment with their child, and (3) examine oxytocin as a mediator between the effects of early childhood trauma on

maternal-child bonding and attachment. Specifically, it is hypothesized that greater experiences of early childhood adversity will be associated with decreased baseline salivary oxytocin in mothers of infants, decreased baseline salivary oxytocin will be associated with impaired maternal-child bonding and attachment as perceived by the mothers, and baseline salivary oxytocin will act as a partial mediator for the effects of early childhood adversity on maternal-child bonding and attachment later in life. This study aims to support previous findings regarding childhood trauma, oxytocin, and maternal-child bonding and attachment by examining the association between salivary oxytocin and ACE scores specifically, two measures that have rarely been used in combination with one another. In addition, salivary oxytocin will be assayed using a novel electrochemiluminescence immunoassay method that has recently been developed and validated through internal testing by Salimetrics LLC., to contribute to this body of research through the use of a more precise measurement of salivary oxytocin that allows for increased sensitivity and controlled detection of oxytocin compared to previously used assay methods in salivary oxytocin research (Rhyne et al., 2009; Zhang et al., 2019). This novel salivary oxytocin assay will also be used in combination with the Postpartum Bonding Questionnaire (PBQ) and the Maternal Postnatal Attachment Scale (MPAS) to examine the association between baseline salivary oxytocin and self-reported maternal-child bonding and attachment respectively.

If supported, these results would suggest oxytocin may be one pathway in which the impacts of early childhood adversity can be transmitted to future generations through impaired maternal-child bonding and attachment. Given the role impaired maternal child-bonding and attachment can play as both a direct risk factor for adverse childhood experiences as well as contributing to an increased susceptibility to negative impacts of trauma and decreased resilience to traumatic events during childhood, it would be argued that reduced oxytocin production due to

trauma may have intergenerational impacts with potentially long-lasting consequences across generations. As such, the present study attempted to propose a model of oxytocin as a pathway for the intergenerational transmission of early trauma that places a prime importance on early adversity, maternal-child bonding and attachment, and the role of oxytocin as a mediator between the two.

Method

Participants

Participants included 33 mothers of infants who had previously participated in a prenatal study on stress and nutrition during their pregnancy. Mothers' age ranged from 19 to 39 years ($M = 31.45$, $SD = 5.62$), were 76% non-Hispanic White, 9% non-Hispanic other, 6% Hispanic White, 3% non-Hispanic Black, 3% Hispanic Black, and 3% Hispanic other. Mothers completed between 11 and 20+ years of education ($M = 16.06$, $SD = 2.87$), and reported an average subjective social status of 5.52 out of 10 on the MacArthur Scale of Subjective Social Status ($SD = 2.11$), indicating that on average participants reported feeling slightly above average in socioeconomic status relative to the rest of the country. All infants were the biological children of the mothers, were between 5 and 20 months ($M = 10.95$, $SD = 4.12$), and were 66.7% male. Participants in the present study were recruited from a pilot study in which a community sample of pregnant women agreed to participate in a prenatal investigation examining the role of stress and nutrition on health outcomes during pregnancy. Participant involvement in the original study concluded upon delivery. Taking advantage of the existing sample, participants in the present study were a subset of those participants who agreed to participate in a postnatal follow up visit with their child. Of the original 55 participants who were involved in the pilot study, 33 agreed to participate in the present study (reasons for declining included moving out of state, being too busy, not interested,

were not able to be reached, etc.). Inclusion criteria for the original prenatal study included mothers who were between the ages of 19 and 40, were in general good health, were not taking any anti-inflammatory or antipsychotic medications, were not a high risk or medically complicated pregnancy, did not conceive their child via egg or embryo donation, and were not pregnant with more than one child.

Procedure

Previous research has shown increases in salivary oxytocin up to 30 minutes before and after breastfeeding (Niwayama et al., 2017; White-Traut et al., 2009). There is also evidence of diurnal fluctuations in salivary oxytocin (Amico et al., 1989; Forsling et al., 1998), although recent evidence has challenged this assumption (Kagerbauer et al., 2019). To control for these potential confounds, participants were invited to the lab between 1 p.m. and 3 p.m. and were asked to avoid food or drink as well as breastfeeding for at least 30 minutes before and after their visit. This helped ensure individual variations would not be confounded with potential diurnal fluctuations in oxytocin or natural increases due to breastfeeding behavior, similar to prior research by Feldman et al. (2010). After arriving for their visit participants and their infants were brought into a carpeted observation room where they reviewed the consent with an examiner. They were then asked to complete a questionnaire on their recent activities and medications, which allowed participants and their infants to become acquainted with the testing environment and reduced mother-child interaction prior to the baseline oxytocin sample. Upon completing the initial questionnaire, responses were examined to confirm their latest time of breastfeeding was within the accepted range. The mothers were then asked to provide a baseline saliva sample before proceeding with the remainder of their visit, the protocols of which are not relevant to the present study. Prior to

their visit participants completed the MPAS and PBQ questionnaires online as part of a larger set of questionnaires used in the broader study, and ACE questionnaires were completed during their prenatal involvement in the original study. All procedures were conducted with approval from the Institutional Review Board and explained to the participants prior to their involvement. All participants signed an informed consent.

Measures

Maternal Adverse Childhood Experiences

The Adverse Childhood Experiences Scale (ACES; Felitti et al., 1998) was used to assess mothers' early life stress, trauma, and adversity. Participants responded yes or no to a series of 10 questions about their experiences prior to age 18 that covered physical, emotional, and sexual abuse, as well as neglect, domestic violence, household substance use, household mental illness, parental separation, and household incarceration. The ACES has shown moderate internal consistency (Cronbach's α = .67) and good test-retest reliability (Cohen's κ = .64; Dube et al., 2004; Folayan et al., 2020). Responses were summed to provide a total score, with higher values indicating greater experiences of early life trauma and adversity. ACE scores have a theoretical minimum of 0 and a maximum of 10, and research has shown a graded relation between ACEs and numerous negative health consequences, with scores of four or greater being associated with a particular risk for adverse health outcomes (Felitti et al., 1998). Mothers in the current sample reported ACE scores between zero and nine with an average ACE score of 2.91 ($SD = 2.67$).

Maternal-Child Bonding

The Postpartum Bonding Questionnaire (PBQ; Brockington et al., 2001) was used to assess impaired maternal-child bonding and was completed online by participants prior to their visit. The PBQ consists of 25 statements with a six-point Likert scale ranging from "Always" (0) to "Never" (5). Questions reflecting a negative emotion or attitude were reverse scored, and all scores were summed to provide a total PBQ score for each participant. The theoretical minimum and maximum scores are 0 and 125 respectively, with higher scores indicating problematic maternal-child bonding. The PBQ can also be divided into four subscales, but these were not used in the analysis. The PBQ has shown good internal consistency (Cronbach's α = .76) and excellent test-retest reliability across the four scales (r = .95, r = .95, r = .93 and r = .77; Wittkowski et al., 2007). Mothers within the current sample scored between 0 and 23, with an average score of 9.15 (*SD* = 5.82).

Maternal-Child Attachment

The Maternal Postpartum Attachment Scale (MPAS; Condon & Corkindale, 1998) was used to assess impaired maternal-child attachment and was completed online by participants prior to their visit. The MPAS consists of 19 statements, with each statement having a two-, three, four, or five-point scale response option. The MPAS has shown good internal consistency (Cronbach's α = .78) and test-retest reliability (r = .86; Condon & Corkindale, 1998). To ensure equal weighting of the questions, all response options were recoded to represent a score of 1 (low attachment) to 5 (high attachment; van Bussel et al., 2010). After completion, the MPAS scores were summed to provide a total MPAS score for each participant, with low scores initially indicating problematic maternal-child attachment. However, to maintain similar directionality with the PBQ for ease of

interpretation, the total MPAS scores were reversed in the present study, such that higher MPAS scores indicate impaired maternal-child attachment. The theoretical minimum and maximum MPAS scores are 19 and 95 respectively. Mothers within the current sample scored between 20.4 and 39.9 with an average total MPAS score of 29.26 ($SD = 5.32$). Additionally, the MPAS can be divided into three subscales, but these were not used in the analysis.

Salivary Oxytocin

Oxytocin from saliva was measured by Salimetrics, LLC (Carlsbad, California, United States). Mothers were instructed to complete a baseline saliva sample using the SalivaBio passive drool method after reviewing consent and becoming acquainted with the testing environment. Mothers were asked to avoid food and drink at least 30 minutes prior to the start of the visit as well as breastfeeding at least 30 minutes before and after their visit. Once collected, samples were stored at -20°C until being shipped on dry ice to be assayed. Salivary oxytocin was determined using a highly sensitive electrochemiluminescence immunoassay method recently developed and validated through internal tests by Salimetrics, LLC, and allowed for sensitive detection in whole, non-extracted saliva using low test volumes without requiring additional sample manipulation and adheres to the applicable NIH guidelines for Enhancing Reproducibility through Rigor and Transparency (Salimetrics, 2021). Sample test volume was 25 μL of saliva per determination and the lower limit of sensitivity of this assay is 8 pg/mL, with an upper limit of 1000 pg/mL. Internal tests by Salimetrics have reported intra-assay and inter-assay coefficients of variation of 16.03% and 3.90% respectively. To increase confidence in salivary oxytocin levels samples were assayed in triplicate and averages were used to provide baseline oxytocin levels for each participant. The average amount of unadjusted baseline salivary oxytocin for mothers in the current sample was

19.42 pg/mL (*SD* = 12.91), and the intra-assay coefficient of variation in the current sample was 27.76% and 20.62% when values under the lower limit of sensitivity were excluded. See Figure 2 for standard curve.

Statistical analysis

All statistical analyses were performed with SPSS version 28 for Macintosh. Prior to data analysis, maternal baseline salivary oxytocin levels were winsorized to avoid outlier effects, with the lowest seven outliers (below lower limit of sensitivity of 8pg/mL) adjusted by +3.25 and the highest four outliers adjusted by -6 to ensure normal distribution. There was no meaningful impact on results between adjusted oxytocin scores and unadjusted oxytocin scores in the following analyses, thus only results of adjusted oxytocin scores will be discussed. One participant was excluded from baseline oxytocin related analyses due to the quantity of saliva in their sample not being sufficient for an accurate assay.

Prior to running primary analyses, correlations were computed to test potential relationships between main variables of interest and various demographic variables, with the resulting correlation matrix used to identify potential covariates.

After identifying potential covariates, a series of multiple regression models were run to test the primary predictions. Models 1a and 1b were used to examine the extent to which ACE scores predict PBQ and MPAS scores respectively. Models 2a and 2b were used to examine the extent to which maternal baseline salivary oxytocin levels predict PBQ and MPAS scores respectively. The final regression models, Models 3a and 3b, were planned to test for the potential mediation effect of maternal baseline salivary oxytocin on ACE score's prediction of PBQ and MPAS scores respectively, as originally hypothesized.

Results

Preliminary Analyses

Of the main variables of interest, descriptives showed on average mothers within the current sample scored low in problematic bonding, with PBQ scores between zero and 23 ($M =$ 9.15, $SD = 5.82$), and low in impaired attachment, with MPAS scores between 20.4 and 39.9 ($M = 29.26$, $SD = 5.32$). ACE scores were between zero and nine with an average score of 2.91 ($SD = 2.67$). Unadjusted baseline oxytocin levels were between 2.88pg/mL and 46.70pg/mL, with an average of 19.42pg/mL ($SD = 12.91$). Participants' winsorized baseline oxytocin levels were between 6.13pg/mL and 40.70pg/mL, with an average of 19.38pg/mL ($SD = 10.62$). Twenty-three of the total 96 assays (24%) used to determine average baseline salivary oxytocin levels were below the lower limit of sensitivity. See Table 1 for sample characteristics and Table 2 for descriptive statistics.

Prior to data analysis, depending on data type Pearson and Spearman correlations were computed to test potential relationships between main variables of interest and demographic variables to identify potential covariates. These included maternal age, child age, sex of child, maternal race/ethnicity, maternal years of education, and maternal subjective social status. The resulting correlation matrix indicated there were no significant correlations between the above demographic variables and the main variables of interest, apart from child age and PBQ scores, which were moderately correlated ($r_s = -.498$, p $= .003$). The correlation matrix is available upon request. Given child age was only moderately correlated with PBQ scores, to avoid oversaturating the analyses and reducing statistical power given the sample size of the present study, child age was not included as a covariate in the subsequent multiple regression analysis. Of note, baseline

salivary oxytocin in the subsequent sections refers to winsorized oxytocin values. Results of the preliminary correlation matrix demonstrated that ACE scores and baseline salivary oxytocin were not significantly correlated (r_s = .045, p = .805), in contrast to what was originally hypothesized. Additionally, PBQ scores and MPAS scores were strongly correlated (r_s = .667, p < .001), indicating that as problematic maternal-child bonding increased, impaired maternal-child attachment increased, as self-reported by the mothers.

ACE Scores' Prediction of PBQ and MPAS Scores

Results from the multiple regression analysis demonstrated that participants' ACE scores did not significantly predict participants' total PBQ scores, β_1 = 0.382, $F(1, 31)$ = .988, p = .328. The effect of participants' ACE scores was small and only accounted for 3% of the variance in participants' total PBQ scores (f^2 = .03). Additionally, participants' ACE scores did not significantly predict participants' total MPAS scores, β_1 = .236, $F(1, 31)$ = .442, p = .511. The effect of participants' ACE scores was small and only accounted for 1% of the variance in participants' total MPAS scores (f^2 = .01). Overall, as participants' ACE scores increased by 1 their total PBQ and MPAS scores increased by .382 and .236 respectively, indicating that greater experiences of early childhood adversity predicted increased impaired bonding and attachment, but this effect was small and not significant. See Table 3, Figure 3, and Figure 4.

Baseline Oxytocin's Prediction of PBQ and MPAS Scores

Results from the multiple regression analysis demonstrated that participants' baseline salivary oxytocin levels did not significantly predict participants' total PBQ scores, β_1 = .038, $F(1, 30)$ = .146, p = .705. The effect of participants' baseline salivary oxytocin levels was small and

only accounted for 1% of the variance in participants' total PBQ scores ($f^2 = .01$). Additionally, participants' baseline salivary oxytocin levels did not significantly predict participants' total MPAS scores, $\beta_1 = .060$, $F(1, 30) = .436$, p = .514. The effect of participants' baseline salivary oxytocin levels was small and only accounted for 1% of the variance in participants' total MPAS scores ($f^2 = .01$). Overall, the model suggests that as participants' baseline oxytocin levels increased by 1pg/mL their total PBQ and MPAS scores increased by .038 and .060 respectively, indicating that higher baseline oxytocin levels predicted increased impaired bonding and attachment, but this effect was small and not significant. See Table 3, Figure 5, and Figure 6.

Mediation Effect of Oxytocin on ACE's Prediction of PBQ and MPAS Scores

Based on these results, further tests of mediation were not warranted to determine if baseline salivary oxytocin had a mediating effect on ACE score's prediction of total PBQ and MPAS scores. Specifically, tests of mediation were determined to be unnecessary because the criteria for a mediation analysis was not supported by the preliminary models (Kraemer et al., 2001).

Discussion

This study examined the correlations among early childhood adversity, baseline salivary oxytocin, and self-reported maternal-child bonding and attachment in a population of mothers of infants. The aim of this study was to use a new combination of measures and a novel salivary oxytocin immunoassay to provide support for previous findings on (1) the association between experiences of early childhood adversity and salivary oxytocin in mothers of infants, (2) the association between maternal salivary oxytocin and mothers' perceived bonding and attachment with their child, (3) examine oxytocin as a mediator between the effects of early childhood trauma

on maternal-child bonding and attachment, and to (4) propose a conceptual model where oxytocin acts as a pathway for the intergenerational transmission of trauma, particularly through its role in promoting healthy maternal-child bonding and attachment.

In contrast to what was originally hypothesized, results showed that when measured with the ACE Scale (Felitti et al., 1998) and salivary oxytocin, experiences of early childhood adversity were not significantly correlated with maternal baseline oxytocin levels. This contradicts several previous findings on the correlation between early adversity, trauma, and oxytocin (e.g., Bertsch et al., 2013; Donadon et al., 2018; Heim et al., 2008); Opacka-Juffry & Mohiyeddini, 2011), including a select few studies using salivary oxytocin measures (Frijling et al., 2015; Mizushima et al., 2015). Additionally, results of a multiple regression analysis demonstrated that while the directionality of the effect was as expected, when measured through the ACE Scale, PBQ (Brockington et al., 2001), and MPAS (Condon & Corkindale, 1998), mothers' experiences of early life adversity did not significantly predict either their self-reported bonding or attachment with their child. This finding contradicts previous research on the association between early adversity and maternal-child attachment and bonding (Erickson et al., 2019; Muzik et al., 2012), though the results regarding maternal-child bonding are not completely unfounded (Lara-Cinisomo et al., 2018), as discussed below. Finally, when measured through salivary oxytocin, the PBQ, and the MPAS, results of the multiple regression analysis demonstrated that mothers' baseline oxytocin levels did not significantly predict their self-reported bonding or attachment with their child. Interestingly, results showed that oxytocin levels were positively correlated with problematic bonding and attachment, though this effect was minimal and non-significant. This finding contradicts prior research which has shown positive correlations between oxytocin levels, parent-child attachment, and bonding (e.g., Feldman et al., 2010a; Galbally et al., 2011; Levine et

al., 2007), as well as behaviors that promote healthy maternal-child bonding and attachment (e.g., Atzil et al., 2011; Feldman et al., 2007; Feldman et al., 2010a). While a mediation analysis was originally planned to examine a potential mediating role of oxytocin on the effect of early adversity on maternal-child bonding and attachment as suggested by some researchers (Heim et al., 2008; Mizuki & Fujiwara, 2015), criteria for conducting a mediation model was not met due to a lack of significance of the main variables of interest (Kraemer et al., 2001).

Outside of methodological limitations, there are several considerations that may have contributed to the findings of the present study. First, while several studies have demonstrated experiences of early adversity and trauma are associated with decreased oxytocin levels later in life, relatively recent research suggests this relationship may depend on the degree of severity of that trauma. For example, Mizuki and Fujiwara (2015) found that less severe forms of childhood maltreatment history might enhance oxytocin concentrations as a response to coping with social stress within the family. Specifically, researchers found that less severe physical abuse was associated with higher oxytocin concentrations, as well as a positive dose-response association between the number of less severe childhood maltreatment types and oxytocin levels. The authors argued that less severe forms of childhood adversity might enhance oxytocin levels in order to cope with social stress (Mizuki & Fujiwara, 2015). Essentially, less severe experiences of early adversity that do not create long-term impairments in HPA axis functioning may promote resilience and coping strategies involving social support through increased oxytocin production compared to decreased oxytocin production seen with severe, chronic experiences of trauma during childhood. In the present study, the non-significant correlation between early adversity and baseline oxytocin may have been influenced by a moderating role of trauma severity, an effect that would likely not have been observed due to the measures used. Specifically, because the ACE

Scale does not assess severity of individual trauma types, and only the number of adverse experiences during childhood, this effect likely would have gone unnoticed in the present study. Similarly, the present study demonstrated that when assessed through the ACE Scale, PBQ, and MPAS, mother's experiences of early adversity did not predict the quality of their self-reported bonding and attachment with their infants. While surprising given the previously highlighted literature, similar results have been observed. Specifically, Lara-Cinisomo et al. (2018) found that the number of childhood traumatic events reported by mothers was not significantly correlated with PBQ scores. However, researchers found significant correlations with other types of trauma histories and specific PBQ subscales. While this suggests the relationship between early adversity and maternal-child bonding in general may not be as strong as currently assumed, the present study would have benefitted from the use of the specific PBQ subscales in order to better support that claim. Overall, future research would benefit from examining specific types of trauma as well as trauma severity rather than number of traumatic events, as seen in Lara-Cinisomo et al. (2018) and Mizuki and Fujiwara (2015).

Secondly, when measured using a novel salivary oxytocin immunoassay and the PBQ and MPAS, the present study did not find a significant correlation between mothers' baseline oxytocin levels and their self-reported bonding and attachment with their infants. While the non-significance of baseline oxytocin levels and maternal-child attachment may have been due to the use of a self-reported measure of attachment rather than a less subjective, observational measure, there may have been important moderation effects hiding these correlations. For example, while overall research suggests a positive correlation between oxytocin and maternal-child bonding and attachment, Elmadih et al. (2014) found that oxytocin may be functioning differently in mothers depending on their level of maternal sensitivity, an important aspect of bonding, attachment, and

caregiving. Specifically, mothers who were characterized as low in maternal sensitivity towards their infant showed higher baseline and post-interaction oxytocin levels than mothers characterized as high in maternal sensitivity. The researchers argued that oxytocin may be acting to reduce stress and anxiety in low-sensitivity mothers in response to the demands of caregiving, while in contrast, mothers characterized as high in maternal sensitivity may find interacting with their infants less stressful and thus observe a smaller increase in oxytocin needed to help return to baseline levels of cortisol. In contrast, Zelkowitz et al. (2014) found that plasma oxytocin was not associated with maternal sensitivity, however, for women who reported high levels of psychosocial stress, higher levels of oxytocin were associated with more sensitive maternal behavior. As the authors suggest, this may be due to oxytocin acting as a buffer against the negative impacts of stress, promoting more sensitive maternal interactive behavior in high-risk women. While these findings paint a confusing—and perhaps unsatisfying—picture of the influence of oxytocin in parenting behavior, they do offer an alternative perspective of the role maternal-sensitivity and psychosocial stress may have on the relationship between oxytocin and maternal-child bonding and attachment given the importance of sensitive parenting in caregiving relationships. In regard to the present study, one possibility is that the influence of factors such as maternal-sensitivity and the role of psychosocial stress that may moderate the differing effects of oxytocin depending on the sensitivity profile of mothers may have hidden the correlation between salivary oxytocin and maternal-child bonding and attachment, which were not examined in the methodology of the present study.

Finally, a possible explanation for the non-significant correlations between baseline oxytocin and maternal-child bonding and attachment the present study is that oxytocin may not act in expected ways in higher-risk populations, i.e., mothers with particularly high ACE scores. While overall studies have demonstrated oxytocin tends to be correlated with positive parenting

behaviors and healthy maternal-child bonding and attachment (e.g., Atzil et al., 2011; Feldman et al., 2007; Feldman et al., 2010a; Galbally et al., 2011; Levine et al., 2007), inconsistent results have been reported (Graustella & MacLeod, 2012), and some research has shown oxytocin may function differently in high-risk populations (Bartz et al., 2010a; Bartz et al., 2010b). In order to address these inconsistencies, Julian et al. (2017)—which to our knowledge is the only other study that has used the ACE Scale in combination with a salivary oxytocin measure—examined oxytocin and parenting behavior among impoverished mothers with low vs. high early life stress. Researchers used the ACE Scale to separate mothers of infants into groups of either high or low early life stress depending on their ACE score, and found that for mothers with low ACE scores, higher oxytocin secretion was associated with more positive parenting as expected. However, for mothers with high ACE scores, higher oxytocin secretion was associated with lower levels of positive parenting (Julian et al., 2017). The researchers theorized that oxytocin may be functioning differently in mothers who experienced harsh early social environments, supporting more defensive behaviors and harsh parenting rather than reducing stress and promoting positive parenting. This finding would corroborate previously discussed results by Mizuki and Fujiwara (2015), which suggests that experiences of less severe forms of early adversity may have beneficial effects in relation to oxytocin and parenting compared to more severe, chronic adversity during childhood. In the present study, while on average mothers had an ACE score of less than three, 36% of mothers had ACE scores of four or more, which has been associated with a particular risk for adverse health outcomes (Felitti et al., 1998), and 15% had an ACE score of seven or greater. A possible explanation for the observed phenomena in the present study is that for mothers with low ACE scores oxytocin may be functioning to promote healthy bonding and attachment through oxytocin's role in facilitating social attachment with their infants, but for mothers with high ACE

scores, oxytocin may be acting to promote more defensive and harsh parenting, which may lead to lower self-reported bonding and attachment with their infants. Due to the study design, such an effect would not have been observed, and future research would benefit from examining how oxytocin functions to promote healthy maternal-child bonding and attachment depending on the degree of trauma and adversity mothers experienced during childhood.

Overall, the results of the present study suggest the relationship between early adversity, oxytocin, and maternal-child bonding and attachment is less direct than originally hypothesized, and may depend on other important factors such as the level of severity of early adversity, the type of adversity experienced, maternal psychosocial stress profiles during child interactions, and their effects on specific components of bonding and attachment behavior rather than bonding and attachment as a whole.

Strengths and limitations

There are a number of strengths in the present study worth mentioning. First, when examining the role of oxytocin as a pathway for the intergenerational impacts of early trauma, the use of a population of recent mothers of infants in combination with measures of early childhood adversity and subsequent maternal-child bonding and attachment is ideal given the crucial vulnerability of early childhood and the impact trauma and poor attachment and bonding relationships can have on long-term health (Buss et al., 2015; Cross et al., 2017; Gunnar & Quevedo, 2007; Mueller & Tronick, 2019). Additionally, this study adds to extant literature through the use of the ACE Scale and a salivary oxytocin measure, two measures that have very rarely been used in combination with one another, adding to our understanding of observed phenomena through the use of more diverse measures. Similarly, the present study is the first to

use a validated electrochemiluminescence immunoassay method to measure salivary oxytocin, an assay method previously unavailable in salivary oxytocin research that allows for sensitive detection in whole, non-extracted saliva using low test volumes without requiring additional sample manipulation. This provides increased sensitivity and controlled detection of oxytocin compared to previously used assay methods in salivary oxytocin research (Rhyne et al., 2009; Zhang et al., 2019). In addition, potentially important confounding variables such as diurnal fluctuations of oxytocin and natural increases of oxytocin levels due to breastfeeding behavior were controlled for in the study procedures, increasing the confidence of oxytocin results in the present study.

However, there are a number of limitations that may limit the scope of the current findings. For one, with only 33 participants the small sample size reduces the external and internal generalizability of the findings, as well as reduces the statistical power of the analyses, which may have contributed to the overall lack of significant results. In addition, there was a relatively large age range of mothers' infants in the present study, between 5.5 and 20.5 months, which may have influenced the main variables of interest. For example, the only statistically significant covariate identified was child age and total PBQ scores, which was negatively correlated. This suggests that mothers with older children self-reported healthier bonding with their child compared to mothers of younger children, and it brings into question whether the age of participants' children may have also influenced their self-reported attachment in a way that was not detected in the correlation due to the small sample size. In addition, the results of the present study, which seem to contradict notable findings in previous literature, may be due to the use of self-reported measures of bonding and attachment compared to more objective measures. For example, while the present study assessed problematic bonding and attachment through the PBQ and MPAS respectively, two

measures based on self-report, other research has found differing results using behavioral coding of attachment and bonding behaviors (Atzil et al., 2011; Feldman et al., 2007; Feldman et al., 2010a). This poses the question of whether we measured mothers' actual bonding and attachment relationships with their children, or simply their perceptions of their bonding and attachment. While the PBQ and MPAS are easier to apply in research than behavioral coding and are commonly used as an approximation of actual maternal-child bonding and attachment, there may be a meaningful difference between the two in the present sample.

In addition, while the salivary oxytocin immunoassay used in the present study has been validated through internal testing by Salimetrics, LLC., and adheres to the applicable NIH guidelines for Enhancing Reproducibility through Rigor and Transparency, assay results for salivary oxytocin in the present study had a particularly high average intra-assay coefficient of variation of 27.76%, and 20.62% when values under the lower limit of sensitivity were excluded. While samples were assayed in triplicate to increase confidence by using individual's average oxytocin levels, such inter-assay variation cannot be ignored, and reflects a common challenge in research using a single baseline saliva sample to determine oxytocin levels (Martins et al., 2020). Similarly, out of the total 96 determinations used to provide average baseline oxytocin levels for each participant, 24% were below the lower limit of sensitivity of 8pg/mL. While the oxytocin values were winsorized to avoid outlier effects from these lower values, specific comparisons between participant's oxytocin levels below this lower limit lack the same degree of confidence compared to oxytocin levels above this limit. This is a limitation in the current study that reflects another common challenge with peripheral oxytocin measures in general (McCullough et al., 2013).

A final set of limitations involve procedural challenges and confounding variables. Specifically, efforts were made to reduce affectionate interaction between mothers and their infants prior to baseline salivary oxytocin collection by having participants spend several minutes reviewing the consent and completing a questionnaire on their recent activities and medications while an RA interacted with the infants. However, mothers and their infants were not completely separated prior to baseline sample collection, and there is a possibility that any interactive behavior between the mothers and their infants influenced baseline oxytocin levels to some degree. However, this was a necessary limitation, as separating the infants from their mothers prior to sample collection would likely have distressed the infants, and in turn, elevated maternal stress levels which may have influenced baseline oxytocin samples. Finally, while a correlation matrix was conducted to identify potential covariates between various demographic variables and the main variables of interest, there are a number of factors that protect against the deleterious effects of early trauma that were not statistically controlled for within the scope of the study. For example, social support has been shown to play a crucial role in reducing the negative impacts of trauma, both on a physiological and psychosocial level (e.g., Muzik et al., 2016; Ozbay et al., 2007). As such, interventions targeting social support both during childhood after exposure to trauma and in adulthood as a means of coping with traumatic childhood experiences are particularly beneficial for protecting against the negative impacts of early adversity (Hogan et al., 2002; Wagner et al., 2016). In the present study, factors such as the extent of social support mothers experienced during early life and in adulthood was not assessed, and may have hidden the correlation between participants' ACE scores and the main variables of interest due to the protective role social support can have on protecting against the negative effects of trauma.

Conclusion

This study examined the correlations between early childhood adversity, baseline salivary oxytocin, and self-reported maternal-child bonding and attachment in a population of mothers of infants. The aim of this study was to use a new combination of measures and a novel salivary oxytocin immunoassay to provide support for previous findings on (1) the association between experiences of early childhood adversity and salivary oxytocin in mothers of infants, (2) the association between maternal salivary oxytocin and mothers' perceived bonding and attachment with their child, (3) examine oxytocin as a mediator between the effects of early childhood trauma on maternal-child bonding and attachment, and to (4) propose a conceptual model where oxytocin acts as a pathway for the intergenerational transmission of trauma, particularly through its role in promoting healthy maternal-child bonding and attachment. Results of a multiple regression analysis demonstrated that mothers' experiences of early childhood adversity were not correlated with baseline salivary oxytocin, mothers' adverse childhood experiences did not significantly predict their self-reported bonding and attachment with their infants, and maternal baseline salivary oxytocin levels did not significantly predict their self-reported bonding and attachment with their infants. This calls into question the strength of the associations suggested in prior research regarding early adversity, oxytocin, and maternal-child bonding and attachment. Future research examining oxytocin as a pathway for the intergenerational impacts of early trauma through the mother-child relationship would benefit from examining the role of factors such as the level of severity of early adversity, the type of adversity experienced, maternal psychosocial stress profiles during child interactions, and the effects of such variables on specific components of bonding and attachment behavior rather than bonding and attachment as a whole.

Works Cited

Amico, J. A., Levin, S. C., & Cameron, J. L. (1989). Circadian rhythm of oxytocin in the cerebrospinal fluid of Rhesus and cynomolgus monkeys: Effects of castration and adrenalectomy and presence of a caudal-rostral gradient. *Neuroendocrinology*, *50*(6), 624–632. https://doi.org/10.1159/000125291

Amico, J. A., Mantella, R. C., Vollmer, R. R., & Li, X. (2004). Anxiety and stress responses in female oxytocin deficient mice. *Journal of Neuroendocrinology*, *16*(4), 319–324. https://doi.org/10.1111/j.0953-8194.2004.01161.x

Apter-Levi, Y., Zagoory-Sharon, O., & Feldman, R. (2014). Oxytocin and vasopressin support distinct configurations of social synchrony. *Brain Research*, *1580*, 124–132. https://doi.org/10.1016/j.brainres.2013.10.052

Atzil, S., Hendler, T., & Feldman, R. (2011). Specifying the neurobiological basis of human attachment: Brain, hormones, and behavior in synchronous and intrusive mothers. *Neuropsychopharmacology*, *36*(13), 2603–2615. https://doi.org/10.1038/npp.2011.172

Atzil, S., Hendler, T., & Feldman, R. (2013). The brain basis of social synchrony. *Social Cognitive and Affective Neuroscience*, *9*(8), 1193–1202. https://doi.org/10.1093/scan/nst105

Bakermans-Kranenburg, M. J., van IJzendoorn, M. H., Riem, M. M., Tops, M., & Alink, L. R. (2011). Oxytocin decreases handgrip force in reaction to infant crying in females without

harsh parenting experiences. *Social Cognitive and Affective Neuroscience*, *7*(8), 951–957. https://doi.org/10.1093/scan/nsr067

Bartels, A., & Zeki, S. (2004). The neural correlates of maternal and romantic love. *NeuroImage*, *21*(3), 1155–1166. https://doi.org/10.1016/j.neuroimage.2003.11.003

Bartz, J. A., Zaki, J., Ochsner, K. N., Bolger, N., Kolevzon, A., Ludwig, N., & Lydon, J. E. (2010a). Effects of oxytocin on recollections of maternal care and closeness. *Proceedings of the National Academy of Sciences*, *107*(50), 21371–21375. https://doi.org/10.1073/pnas.1012669107

Bartz, J., Simeon, D., Hamilton, H., Kim, S., Crystal, S., Braun, A., Vicens, V., & Hollander, E. (2010b). Oxytocin can hinder trust and cooperation in borderline personality disorder. *Social Cognitive and Affective Neuroscience*, *6*(5), 556–563. https://doi.org/10.1093/scan/nsq085

Bertsch, K., Schmidinger, I., Neumann, I. D., & Herpertz, S. C. (2013). Reduced plasma oxytocin levels in female patients with borderline personality disorder. *Hormones and Behavior*, *63*(3), 424–429. https://doi.org/10.1016/j.yhbeh.2012.11.013

Brockington, I. F., Oates, J., George, S., Turner, D., Vostanis, P., Sullivan, M., Loh, C., & Murdoch, C. (2001). A screening questionnaire for mother-infant bonding disorders. *Archives of Women's Mental Health*, *3*(4), 133–140. https://doi.org/10.1007/s007370170010

Brown, C. A., Cardoso, C., & Ellenbogen, M. A. (2016). A meta-analytic review of the

correlation between peripheral oxytocin and cortisol concentrations. *Frontiers in

Neuroendocrinology*, *43*, 19–27. https://doi.org/10.1016/j.yfrne.2016.11.001

Buss, K. E., Warren, J. M., & Horton, E. (2015). Trauma and treatment in early childhood: A

review of the historical and emerging literature for counselors. *The Professional

Counselor*, *5*(2), 225–237. https://doi.org/10.15241/keb.5.2.225

Condon, J. T., & Corkindale, C. J. (1998). The assessment of parent-to-infant attachment:

Development of a self-report questionnaire instrument. *Journal of Reproductive and Infant

Psychology*, *16*(1), 57–76. https://doi.org/10.1080/02646839808404558

Cross, D., Fani, N., Powers, A., & Bradley, B. (2017). Neurobiological development in the

context of childhood trauma. *Clinical Psychology: Science and Practice*, *24*(2), 111–124.

https://doi.org/10.1111/cpsp.12198

Donadon, M. F., Martin-Santos, R., & Osório, F. de. (2018). The associations between oxytocin

and trauma in humans: A systematic review. *Frontiers in Pharmacology*, *9*.

https://doi.org/10.3389/fphar.2018.00154

Dube, S. R., Williamson, D. F., Thompson, T., Felitti, V. J., & Anda, R. F. (2004). Assessing the

reliability of retrospective reports of adverse childhood experiences among adult HMO

members attending a primary care clinic. *Child Abuse & Neglect*, *28*(7), 729–737.

https://doi.org/10.1016/j.chiabu.2003.08.009

Dębiec, J. (2005). Peptides of love and fear: Vasopressin and oxytocin modulate the integration of information in the amygdala. *BioEssays*, *27*(9), 869–873. https://doi.org/10.1002/bies.20301

Elmadih, A., Wai Wan, M., Numan, M., Elliott, R., Downey, D., & Abel, K. M. (2014). Does oxytocin modulate variation in maternal caregiving in Healthy New Mothers? *Brain Research*, *1580*, 143–150. https://doi.org/10.1016/j.brainres.2014.01.020

Erickson, N., Julian, M., & Muzik, M. (2019). Perinatal Depression, PTSD, and trauma: Impact on mother–infant attachment and interventions to mitigate the transmission of risk. *International Review of Psychiatry*, *31*(3), 245–263. https://doi.org/10.1080/09540261.2018.1563529

Erozkan, A. (2016). The link between types of attachment and childhood trauma. *Universal Journal of Educational Research*, *4*(5), 1071–1079. https://doi.org/10.13189/ujer.2016.040517

Fang, A., Hoge, E. A., Heinrichs, M., & Hofmann, S. G. (2014). Attachment style moderates the effects of oxytocin on social behaviors and cognitions during social rejection. *Clinical Psychological Science*, *2*(6), 740–747. https://doi.org/10.1177/2167702614527948

Feldman, R., Gordon, I., & Zagoory-Sharon, O. (2010a). Maternal and paternal plasma, salivary, and urinary oxytocin and parent-infant synchrony: Considering stress and affiliation components of human bonding. *Developmental Science*, *14*(4), 752–761. https://doi.org/10.1111/j.1467-7687.2010.01021.x

Feldman, R., Gordon, I., Schneiderman, I., Weisman, O., & Zagoory-Sharon, O. (2010b). Natural variations in maternal and paternal care are associated with systematic changes in oxytocin following parent–infant contact. *Psychoneuroendocrinology*, *35*(8), 1133–1141. https://doi.org/10.1016/j.psyneuen.2010.01.013

Feldman, R., Weller, A., Zagoory-Sharon, O., & Levine, A. (2007). Evidence for a neuroendocrinological foundation of human affiliation. *Psychological Science*, *18*(11), 965–970. https://doi.org/10.1111/j.1467-9280.2007.02010.x

Felitti, V. J., Anda, R. F., Nordenberg, D., Williamson, D. F., Spitz, A. M., Edwards, V., Koss, M. P., & Marks, J. S. (1998). Relationship of childhood abuse and household dysfunction to many of the leading causes of death in adults. *American Journal of Preventive Medicine*, *14*(4), 245–258. https://doi.org/10.1016/s0749-3797(98)00017-8

Folayan, M. O., Oginni, O., Arowolo, O., & El Tantawi, M. (2020). Internal consistency and correlation of the adverse childhood experiences, bully victimization, self-esteem, resilience, and social support scales in Nigerian children. *BMC Research Notes*, *13*(1). https://doi.org/10.1186/s13104-020-05174-3

Forsling, M. L., Montgomery, H., Halpin, D., Windle, R. J., & Treacher, D. F. (1998). Daily patterns of secretion of neurohypophysial hormones in man: Effect of age. *Experimental Physiology*, *83*(3), 409–418. https://doi.org/10.1113/expphysiol.1998.sp004124

Frijling, J. L., van Zuiden, M., Nawijn, L., Koch, S. B., Neumann, I. D., Veltman, D. J., & Olff, M. (2015). Salivary oxytocin and vasopressin levels in police officers with and without

post-traumatic stress disorder. *Journal of Neuroendocrinology*, *27*(10), 743–751. https://doi.org/10.1111/jne.12300

Galbally, M., Lewis, A. J., van IJzendoorn, M., & Permezel, M. (2011). The role of oxytocin in mother-infant relations: A systematic review of human studies. *Harvard Review of Psychiatry*, *19*(1), 1–14. https://doi.org/10.3109/10673229.2011.549771

Gonzalez, A., Jenkins, J. M., Steiner, M., & Fleming, A. S. (2009). The relation between early life adversity, cortisol awakening response and diurnal salivary cortisol levels in postpartum women. *Psychoneuroendocrinology*, *34*(1), 76–86. https://doi.org/10.1016/j.psyneuen.2008.08.012

Gordon, I., Zagoory-Sharon, O., Leckman, J. F., & Feldman, R. (2010). Oxytocin and the development of parenting in humans. *Biological Psychiatry*, *68*(4), 377–382. https://doi.org/10.1016/j.biopsych.2010.02.005

Graustella, A. J., & MacLeod, C. (2012). A critical review of the influence of oxytocin nasal spray on social cognition in humans: Evidence and future directions. *Hormones and Behavior*, *61*(3), 410–418. https://doi.org/10.1016/j.yhbeh.2012.01.002

Gulpinar, M., & Yegen, B. (2004). The Physiology of Learning and memory: Role of peptides and stress. *Current Protein & Peptide Science*, *5*(6), 457–473. https://doi.org/10.2174/1389203043379341

Gunnar, M. R., & Quevedo, K. M. (2007). Early care experiences and HPA axis regulation in children: A mechanism for later trauma vulnerability. *Progress in Brain Research*, 137–149. https://doi.org/10.1016/s0079-6123(07)67010-1

Heim, C., Young, L. J., Newport, D. J., Mletzko, T., Miller, A. H., & Nemeroff, C. B. (2008). Lower CSF oxytocin concentrations in women with a history of childhood abuse. *Molecular Psychiatry*, *14*(10), 954–958. https://doi.org/10.1038/mp.2008.112

Heinrichs, M., Meinlschmidt, G., Wippich, W., Ehlert, U., & Hellhammer, D. (2004). Selective amnesic effects of oxytocin on human memory. *Physiology & Behavior*, *83*(1), 31–38. https://doi.org/10.1016/s0031-9384(04)00346-4

Hogan, B. E., Linden, W., & Najarian, B. (2002). Social Support Interventions. *Clinical Psychology Review*, *22*(3), 381–440. https://doi.org/10.1016/s0272-7358(01)00102-7

Julian, M. M., Rosenblum, K. L., Doom, J. R., Leung, C. Y., Lumeng, J. C., Cruz, M. G., Vazquez, D. M., & Miller, A. L. (2017). Oxytocin and parenting behavior among impoverished mothers with low vs. high early life stress. *Archives of Women's Mental Health*, *21*(3), 375–382. https://doi.org/10.1007/s00737-017-0798-6

Kagerbauer, S. M., Debus, J. M., Martin, J., Gempt, J., Jungwirth, B., Hapfelmeier, A., & Podtschaske, A. H. (2019). Absence of a diurnal rhythm of oxytocin and arginine-vasopressin in human cerebrospinal fluid, blood and Saliva. *Neuropeptides*, *78*, 101977. https://doi.org/10.1016/j.npep.2019.101977

Kennison, S. M., & Spooner, V. H. (2020). Childhood relationships with parents and attachment as predictors of resilience in young adults. *Journal of Family Studies*, 1–13. https://doi.org/10.1080/13229400.2020.1861968

Kirsch, P. (2005). Oxytocin modulates neural circuitry for social cognition and fear in humans. *Journal of Neuroscience*, *25*(49), 11489–11493. https://doi.org/10.1523/jneurosci.3984-05.2005

Kraemer, H. C., Stice, E., Kazdin, A, Offord, D., & Kupfer, D. (2001). How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *American Journal of Psychiatry, 158*, 848–856.

Lara-Cinisomo, S., Zhu, K., Fei, K., Bu, Y., Weston, A. P., & Ravat, U. (2018). Traumatic events: Exploring Associations with maternal depression, infant bonding, and oxytocin in Latina mothers. *BMC Women's Health*, *18*(1). https://doi.org/10.1186/s12905-018-0520-5

Levine, A., Zagoory-Sharon, O., Feldman, R., & Weller, A. (2007). Oxytocin during pregnancy and early postpartum: Individual patterns and maternal–fetal attachment. *Peptides*, *28*(6), 1162–1169. https://doi.org/10.1016/j.peptides.2007.04.016

Martins, D., Gabay, A. S., Mehta, M., & Paloyelis, Y. (2020). Salivary and plasmatic oxytocin are not reliable trait markers of the physiology of the oxytocin system in humans. *ELife*, *9*. https://doi.org/10.7554/elife.62456

McCullough, M. E., Churchland, P. S., & Mendez, A. J. (2013). Problems with measuring

peripheral oxytocin: Can the data on oxytocin and human behavior be trusted?

Neuroscience & Biobehavioral Reviews, *37*(8), 1485–1492.

https://doi.org/10.1016/j.neubiorev.2013.04.018

Miranda, A. M., Soares, C. N., Moraes, M. L., Fossaluza, V., Serafim, P. M., & Mello, M. F.

(2012). Healthy maternal bonding as a resilience factor for depressive disorder. *Psychology

& Neuroscience*, *5*(1), 21–25. https://doi.org/10.3922/j.psns.2012.1.04

Mizuki, R., & Fujiwara, T. (2015). Association of oxytocin level and less severe forms of

childhood maltreatment history among healthy Japanese adults involved with child care.

Frontiers in Behavioral Neuroscience, *9*. https://doi.org/10.3389/fnbeh.2015.00138

Mizushima, S. G., Fujisawa, T. X., Takiguchi, S., Kumazaki, H., Tanaka, S., & Tomoda, A.

(2015). Effect of the nature of subsequent environment on oxytocin and cortisol secretion

in maltreated children. *Frontiers in Psychiatry*, *6*. https://doi.org/10.3389/fpsyt.2015.00173

Mueller, I., & Tronick, E. (2019). Early life exposure to violence: Developmental consequences

on brain and behavior. *Frontiers in Behavioral Neuroscience*, *13*.

https://doi.org/10.3389/fnbeh.2019.00156

Murray, L., Fiori-Cowley, A., Hooper, R., & Cooper, P. (1996). The impact of postnatal

depression and associated adversity on early mother-infant interactions and later infant

outcome. *Child Development*, *67*(5), 2512–2526. https://doi.org/10.1111/j.1467-

8624.1996.tb01871.x

Muzik, M., Bocknek, E. L., Broderick, A., Richardson, P., Rosenblum, K. L., Thelen, K., &
    Seng, J. S. (2012). Mother–infant bonding impairment across the first 6 months
    postpartum: The primacy of psychopathology in women with childhood abuse and Neglect
    Histories. *Archives of Women's Mental Health*, *16*(1), 29–38.
    https://doi.org/10.1007/s00737-012-0312-0

Muzik, M., Umarji, R., Sexton, M. B., & Davis, M. T. (2016). Family social support modifies the
    relationships between childhood maltreatment severity, economic adversity and postpartum
    depressive symptoms. *Maternal and Child Health Journal*, *21*(5), 1018–1025.
    https://doi.org/10.1007/s10995-016-2197-4

Nissen, E., Lilja, G., Widström, A.-M., & Uvnás-Moberg, K. (1995). Elevation of oxytocin
    levels early post partum in women. *Acta Obstetricia Et Gynecologica Scandinavica*, *74*(7),
    530–533. https://doi.org/10.3109/00016349509024384

Niwayama, R., Nishitani, S., Takamura, T., Shinohara, K., Honda, S., Miyamura, T., Nakao, Y.,
    Oishi, K., & Araki-Nagahashi, M. (2017). Oxytocin mediates a calming effect on
    postpartum mood in primiparous mothers. *Breastfeeding Medicine*, *12*(2), 103–109.
    https://doi.org/10.1089/bfm.2016.0052

Oldfield, J., Stevenson, A., Ortiz, E., & Haley, B. (2018). Promoting or suppressing resilience to
    mental health outcomes in at risk young people: The role of parental and peer attachment
    and school connectedness. *Journal of Adolescence*, *64*(1), 13–22.
    https://doi.org/10.1016/j.adolescence.2018.01.002

Opacka-Juffry, J., & Mohiyeddini, C. (2011). Experience of stress in childhood negatively correlates with plasma oxytocin concentration in adult men. *Stress*, *15*(1), 1–10. https://doi.org/10.3109/10253890.2011.560309

Operario, D., Adler, N. E., & Williams, D. R. (2004). Subjective social status: Reliability and predictive utility for Global Health. *Psychology & Health*, *19*(2), 237–246. https://doi.org/10.1080/08870440310001638098

Osofsky, J. D., Osofsky, H. J., Frazer, A. L., Fields-Olivieri, M. A., Many, M., Selby, M., Holman, S., & Conrad, E. (2021). The importance of adverse childhood experiences during the perinatal period. *American Psychologist*, *76*(2), 350–363. https://doi.org/10.1037/amp0000770

Ozbay, F., Johnson, D. C., Dimoulas, E., Morgan, C. A., Charney, D., & Southwick, S. (2007). Social support and resilience to stress: from neurobiology to clinical practice. *Psychiatry (Edgmont (Pa. : Township))*, *4*(5), 35–40.

Pickreign Stronach, E., Toth, S. L., Rogosch, F., Oshri, A., Manly, J. T., & Cicchetti, D. (2011). Child maltreatment, attachment security, and internal representations of mother and mother-child relationships. *Child Maltreatment*, *16*(2), 137–145. https://doi.org/10.1177/1077559511398294

Rhyne, P. W., Wong, O. T., Zhang, Y. J., & Weiner, R. S. (2009). Electrochemiluminescence in Bioanalysis. *Bioanalysis*, *1*(5), 919–935. https://doi.org/10.4155/bio.09.80

Riem, M. M. E., Bakermans-Kranenburg, M. J., Pieper, S., Tops, M., Boksem, M. A. S.,

Vermeiren, R. R. J. M., van IJzendoorn, M. H., & Rombouts, S. A. R. B. (2011a).

Oxytocin modulates amygdala, insula, and inferior frontal gyrus responses to infant crying:

A randomized controlled trial. *Biological Psychiatry*, *70*(3), 291–297.

https://doi.org/10.1016/j.biopsych.2011.02.006

Riem, M. M., van IJzendoorn, M. H., Tops, M., Boksem, M. A., Rombouts, S. A., & Bakermans-

Kranenburg, M. J. (2011b). No laughing matter: Intranasal oxytocin administration

changes functional brain connectivity during exposure to infant laughter.

Neuropsychopharmacology, *37*(5), 1257–1266. https://doi.org/10.1038/npp.2011.313

Salimetrics. (2021). *Piloting salivary oxytocin – observations and opportunities for research*.

Salimetrics. Retrieved June 21, 2022, from https://salimetrics.com/piloting-salivary-

oxytocin-observations-and-opportunities-for-research/

Scatliffe, N., Casavant, S., Vittner, D., & Cong, X. (2019). Oxytocin and early parent-infant

interactions: A systematic review. *International Journal of Nursing Sciences*, *6*(4), 445–

453. https://doi.org/10.1016/j.ijnss.2019.09.009

Sharma, S. R., Gonda, X., Dome, P., & Tarazi, F. I. (2020). What's love got to do with it: Role of

oxytocin in trauma, attachment and resilience. *Pharmacology & Therapeutics*, *214*,

107602. https://doi.org/10.1016/j.pharmthera.2020.107602

Stacy, P. D. (2006). Early childhood attachments as a protective factor. *Journal of Evidence-

Based Social Work*, *3*(2), 49–65. https://doi.org/10.1300/j394v03n02_03

Strathearn, L., Fonagy, P., Amico, J., & Montague, P. R. (2009). Adult attachment predicts maternal brain and oxytocin response to infant cues. *Neuropsychopharmacology*, *34*(13), 2655–2666. https://doi.org/10.1038/npp.2009.103

Szymanska, M., Schneider, M., Chateau-Smith, C., Nezelof, S., & Vulliez-Coady, L. (2017). Psychophysiological effects of oxytocin on parent-child interactions: A literature review on oxytocin and parent-child interactions. *Psychiatry and Clinical Neurosciences*, *71*(10), 690–705. https://doi.org/10.1111/pcn.12544

Tops, M., van Peer, J. M., Korf, J., Wijers, A. A., & Tucker, D. M. (2007). Anxiety, cortisol, and attachment predict plasma oxytocin. *Psychophysiology*, *44*(3), 444–449. https://doi.org/10.1111/j.1469-8986.2007.00510.x

van Bussel, J. C., Spitz, B., & Demyttenaere, K. (2010). Three self-report questionnaires of the early mother-to-infant bond: Reliability and validity of the Dutch version of the mpas, PBQ and Mibs. *Archives of Women's Mental Health*, *13*(5), 373–384. https://doi.org/10.1007/s00737-009-0140-z

Wagner, A. C., Monson, C. M., & Hart, T. L. (2016). Understanding social factors in the context of trauma: Implications for measurement and Intervention. *Journal of Aggression, Maltreatment & Trauma*, *25*(8), 831–853. https://doi.org/10.1080/10926771.2016.1152341

Weisman, O., Zagoory-Sharon, O., & Feldman, R. (2012). Oxytocin administration to parent enhances infant physiological and behavioral readiness for Social Engagement. *Biological Psychiatry*, *72*(12), 982–989. https://doi.org/10.1016/j.biopsych.2012.06.011

White-Traut, R., Watanabe, K., Pournajafi-Nazarloo, H., Schwertz, D., Bell, A., & Carter, C. S. (2009). Detection of salivary oxytocin levels in lactating women. *Developmental Psychobiology*, *51*(4), 367–373. https://doi.org/10.1002/dev.20376

Wittkowski, A., Wieck, A., & Mann, S. (2007). An evaluation of two bonding questionnaires: A comparison of the mother-to-infant bonding scale with the postpartum bonding questionnaire in a sample of primiparous mothers. *Archives of Women's Mental Health*, *10*(4), 171–175. https://doi.org/10.1007/s00737-007-0191-y

Zelkowitz, P., Gold, I., Feeley, N., Hayton, B., Carter, C. S., Tulandi, T., Abenhaim, H. A., & Levin, P. (2014). Psychosocial stress moderates the relationships between oxytocin, perinatal depression, and maternal behavior. *Hormones and Behavior*, *66*(2), 351–360. https://doi.org/10.1016/j.yhbeh.2014.06.014

Zhang, J., Arbault, S., Sojic, N., & Jiang, D. (2019). Electrochemiluminescence imaging for Bioanalysis. *Annual Review of Analytical Chemistry*, *12*(1), 275–295. https://doi.org/10.1146/annurev-anchem-061318-115226

Tables

Table 1. Sample Characteristics

Variable	N	Min.	Max.	Mean	SD
Age					
Child Age (months)	33	5.49	20.52	10.95	4.12
Maternal Age (years)	33	19.47	39.07	31.45	5.62
Maternal Years of Education	33	11	20	16.06	2.87
Subjective Socioeconomic Status	33	1	8	5.52	2.11

Variable	N	% of Total Sample
Sex of Child		
Male	22	66.67
Female	11	33.33
Maternal Race/Ethnicity		
Non-Hispanic White	25	75.76
Other	8	24.24

Note. Subjective Socioeconomic Status was measured using the MacArthur Scale of Subjective

Social Status - Society Ladder (Operario et al., 2004)

Table 2. Descriptive Statistics for Study Variables

Variable	N	M	(SD)	Min.	Max.
ACE Score	33	2.91	2.67	0	9
PBQ Total	33	9.15	5.82	0	23
MPAS Total	33	29.26	5.32	20.40	39.90
Baseline Oxytocin	32	19.42	12.91	2.88	46.70
Winsorized Baseline Oxytocin	32	19.38	10.62	6.13	40.70

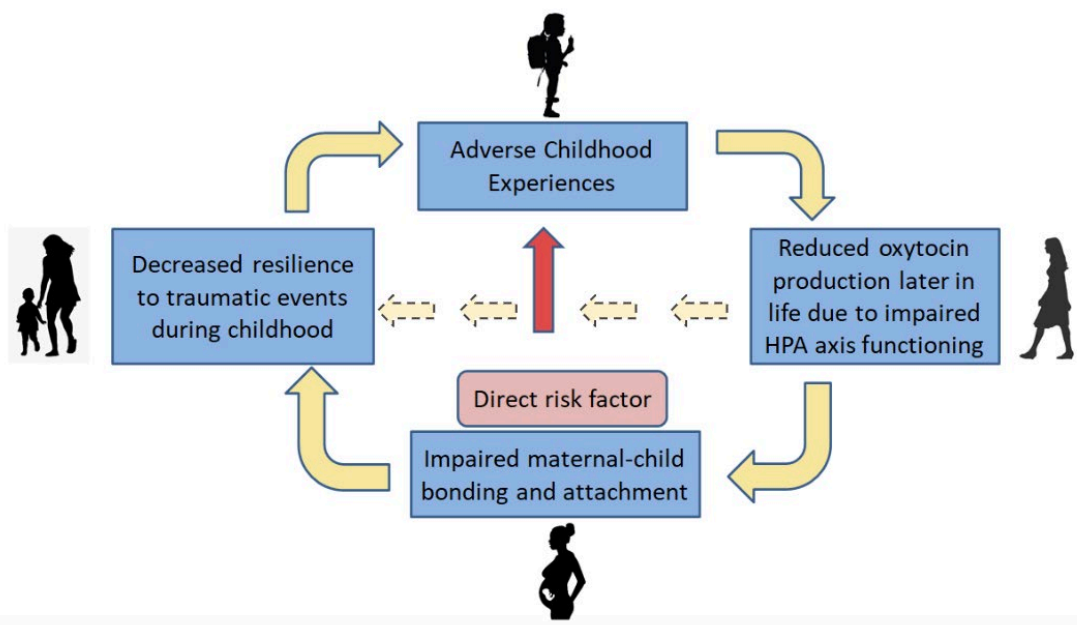*Note.* Winsorized baseline salivary oxytocin levels were used in multiple regression analysis

Table 3. Multiple Regression Models

MODELS		Regression Models						
	B	SE B	β	t	f^2	p	95% CI	
Model 1a: PBQ Total (outcome) ACE Score (predictor)	.38	.39	.18	.99	.03	.328	[-.40, 1.17]	
$R^2 = .03$, $F(1, 31) = .988$								
Model 1b: MPAS Total (outcome) ACE Score (predictor	.24	.36	.12	.67	.01	.511	[-.49, .96]	
$R^2 = .01$, $F(1, 31) = .442$								
Model 2a: PBQ Total (outcome) Baseline Oxytocin (predictor)	.04	.10	.07	.38	.01	.705	[-.17, .24]	
$R^2 = .01$, $F(1, 30) = .146$								
Model 2b: MPAS Total (outcome) Baseline Oxytocin (predictor)	.06	.09	.12	.66	0.01	.514	[-.13, .25]	
$R^2 = .01$, $F(1, 30) = .436$								

*Note.* CI = confidence interval; SE B = standardized coefficients of beta

Figures

Figure 1. Model of Oxytocin as a Pathway for the Intergenerational Impacts of Early Childhood Trauma

*Note.* The present study focuses on the correlations between adverse childhood experiences,

baseline oxytocin levels, and impaired maternal-child bonding and attachment specifically

Figure 2. Standard Curve for Baseline Salivary Oxytocin Assay
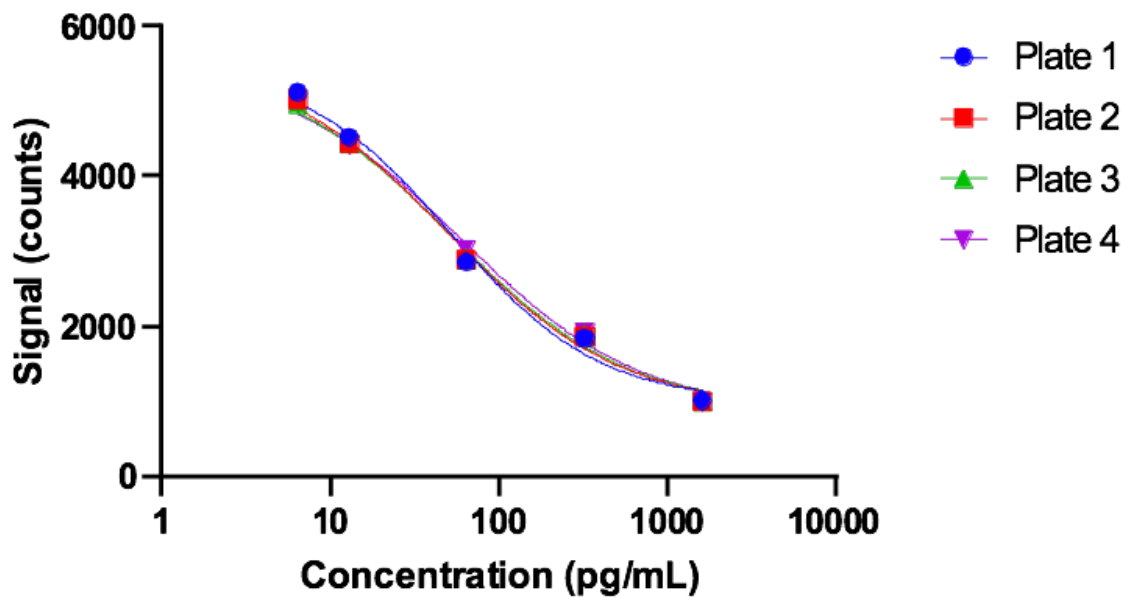
*Note*. Depiction of standard curve data for baseline maternal salivary oxytocin measured via a

novel electrochemiluminescence immunoassay method.

Figure 3. ACE Score's Prediction of Total PBQ Score
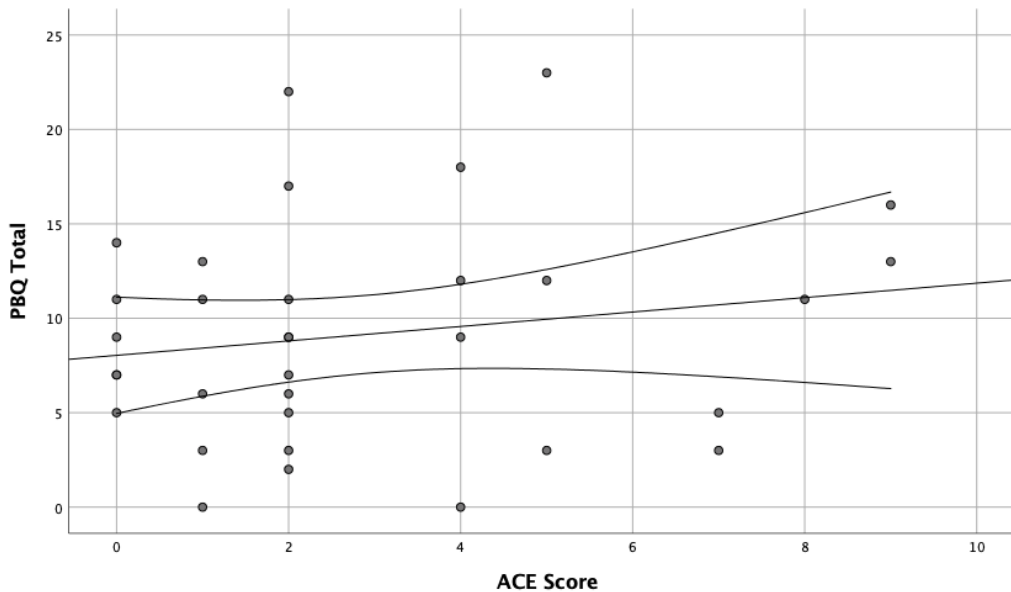
Note. Depiction of multiple regression model results for ACE Score's prediction of total PBQ score. Higher total PBQ scores indicate increased problematic maternal-child bonding as perceived by mothers.

Figure 4. ACE Score's Prediction of Total MPAS Score
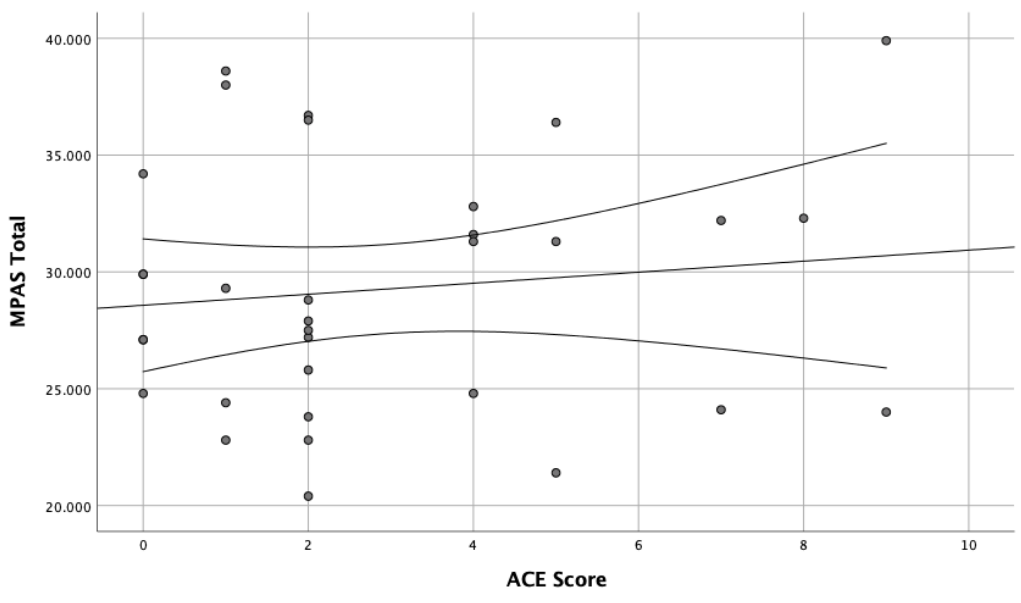
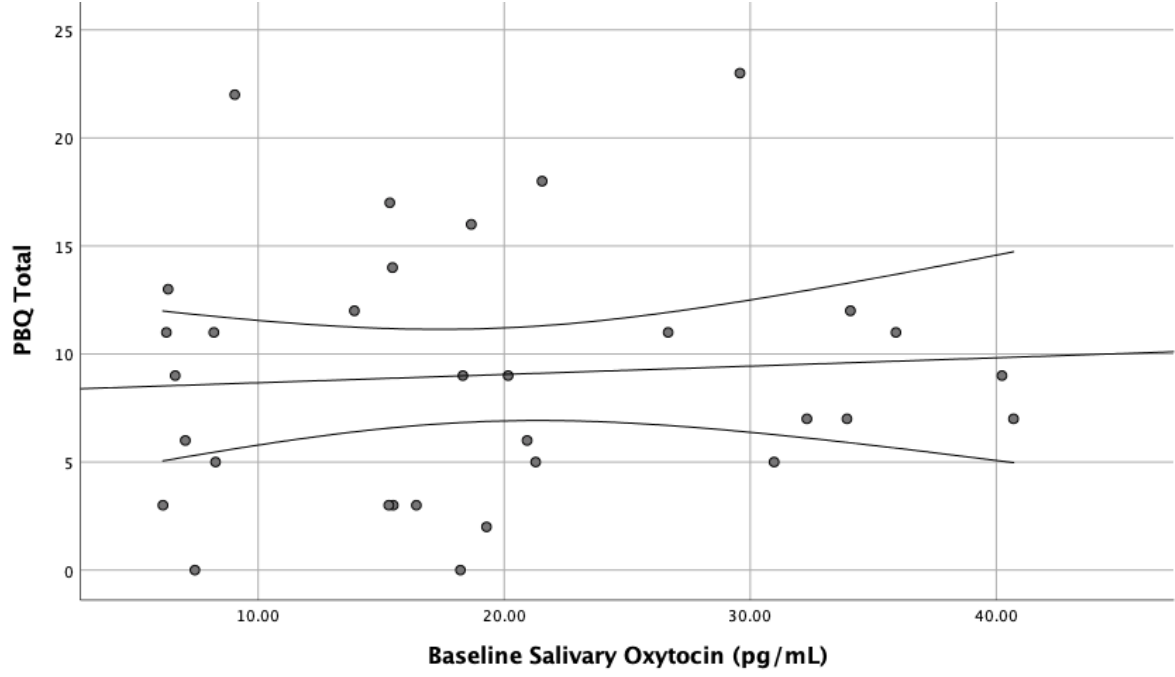*Note.* Depiction of multiple regression model results for ACE Score's prediction of total MPAS

score. Higher total MPAS scores indicate increased problematic maternal-child attachment as

perceived by mothers.

Figure 5. Baseline Salivary Oxytocin's Prediction of Total PBQ Score

Note. Baseline salivary oxytocin depicted here is based on winsorized average oxytocin values

for each participant.

Figure 6. Baseline Salivary Oxytocin's Prediction of Total MPAS Score
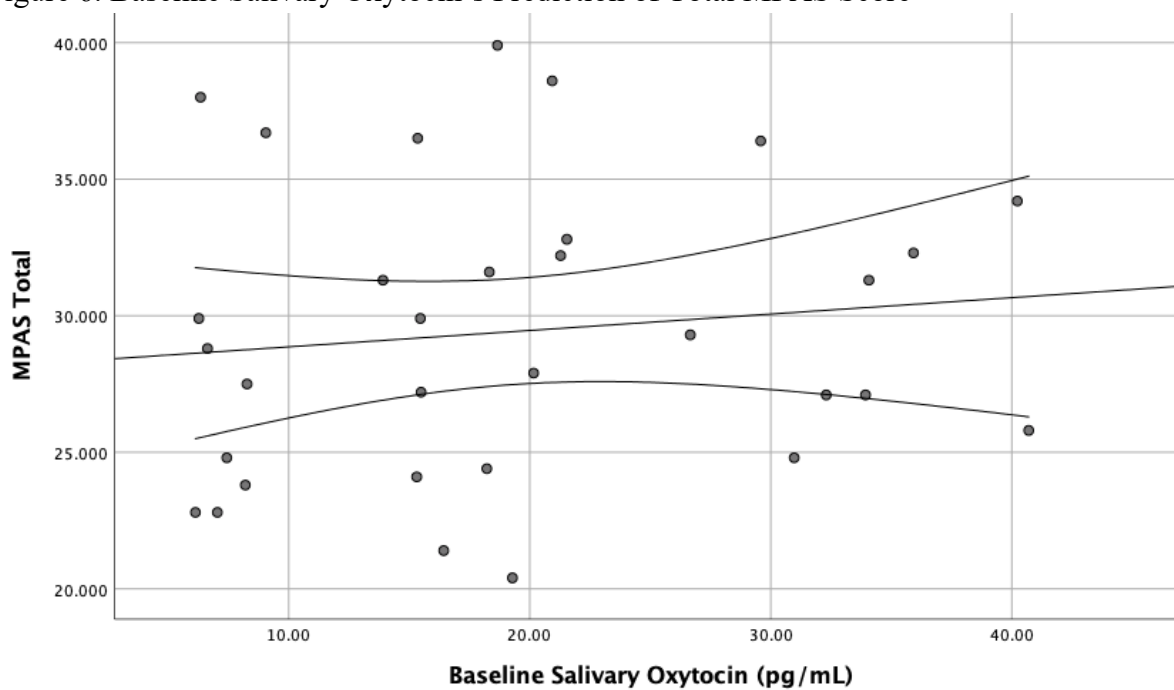
Note. Baseline salivary oxytocin depicted here is based on winsorized average oxytocin values for each participant.

The Relationship Between Students' Future Goals and Motivation to Learn in a Post-

Secondary Science Course

Madeleine G. Smith

University of Oregon

Abstract

Supporting student motivation in post-secondary introductory science courses is an important step to increasing student retention in science, technology, engineering, and mathematics (STEM) fields. Prior research has focused on a variable-centered approach to students' motivation to learn, but a person-centered approach to looking at the relationship between self-regulated learning and motivation variables is necessary. Additionally, research suggests that students' future goals and their perception of the future influences their performance in school. This study aimed to identify the relationship between future thinking and motivational and self-regulated learning profiles. 384 participants were recruited from an introductory science course at a large U.S. Northwest university and completed an online self-report survey on their motivation, goal orientation, self-regulation, knowledge building strategies, and future goals. K-means cluster analysis indicated that there are three motivational and self-regulated learning profiles: 1) Mastery-oriented, 2) Apathetic, and 3) Performance-oriented. A multinomial logistic regression was conducted and showed that students' actions towards their future goals are a stronger predictor of a mastery-oriented profile adoption than the clarity of their future goals. Understanding the relationship between students' future thinking and their adoption of learning profiles can help STEM instructors support student motivation and grade performance through focusing on mastery-oriented classroom practice.

Keywords: motivation, STEM, life project, future thinking, learning profiles

The Relationship Between Students' Future Goals and Motivation to Learn in a Post-Secondary Science Course

There has been a national focus on increasing the retention of science, technology, engineering, and mathematics (STEM) majors in post-secondary institutions (Olsen & Riordan, 2012); nearly 50% of STEM majors do not complete their degrees in STEM (Doerschuk et al., 2016). Students are more likely to persist in college if they have clear goals and value present activities (Ethington, 1990). Research has also found that students are more likely to be successful in class and persist in STEM majors if they use self-regulated learning strategies (strategies that help students monitor and evaluate their learning), focus on learning rather than performance goals, see the course as instrumental for their future goals, and have high self-efficacy (Wormington & Linnenbrink-Garcia, 2016). Students who, in general, value their future goals are more likely to engage in adaptive behavior and be motivated (Hilpert et al., 2012). It is important to understand how student perceptions of their future relate to other motivational variables in the context of STEM courses, in order to identify how student motivation and retention can be improved for STEM majors in post-secondary institutions. This study aims to identify the relationship between life project, a future time perspective variable, and motivational and self-regulated learning profiles for students in a post-secondary science course.

Education is inherently about the future (Husman & Lens, 1999) and a future orientation has motivational and behavioral consequences in the present (Lens et al., 2012). Individuals who are future-oriented and understand the value of present activities for the future are found to have enhanced motivation (Simons et al., 2000). Research suggests that facets of future thinking, such as perception of instrumentality, increases optimal learning (Husman et al., 2004). Therefore, it is important to investigate the role of future in STEM students' motivation.

Targeted interventions have been developed to widen the pipeline to post-secondary STEM education and careers by manipulating key aspects of students' motivation. Many interventions hope to increase student retention through improving students' perception of STEM utility value (e.g., Harackiewicz et al., 2016; Harackiewicz et al., 2012). These studies are built on the assumption that students' persistence in post-secondary STEM courses relates to their perception of instrumentality for the course. Perception of instrumentality, a future-oriented utility value, is an individual's belief about how their present activities will help them achieve their future goals (Husman et al., 2004). These utility-based interventions aim to decrease attrition in STEM majors by increasing students' perception of instrumentality; however, other motivational factors are simultaneously influencing students' performance and persistence. The complex relations between students' perceptions of instrumentality for their class and these other factors – including their goal orientation, self-efficacy, and self-regulated learning strategy use – requires a multivariate approach.

Prior research has focused on which motivational factors have the strongest relation to self-regulated learning (e.g., Zimmerman, 2008). A variable-centered approach is the dominant methodological approach to examine the relation between motivational and self-regulatory constructs. It is well established from a variable-centered approach that both goal orientation and utility value are related to self-regulation. A person-centered approach is helpful in understanding how these different motivational variables, in addition to utility value, influence students' motivation for self-regulated learning.

This study will use a student-centered approach (profiling approach) to identify and find the relation between patterns in students' motivation and self-regulated learning and performance in a general science course, as well as the relation between life project (future time orientation)

scores and the adoption of self-regulated and motivational learning profiles. I hypothesized that the motivational and self-regulation variables would form learning profiles and that adaptive profiles would relate to high levels of life project. This study addresses the following research questions:

RQ1: What motivational and self-regulated learning profiles do students adopt in a post-secondary science course?

RQ2: Does performance in a post-secondary science course differ by profile?

RQ3: Do life project scores predict profile adoption in a required science course?

Theoretical Framework

A Person-Centered Approach to Motivation

Research on motivational and self-regulated constructs have found a multivariate approach to studying the pattern of relations between perception of instrumentality, self-efficacy, self-regulation, and goal orientation is necessary (Nelson et al., 2015; Shell & Husman, 2008; Shell & Soh, 2013; Wormington et al., 2016). Profile adoption is dynamic, and students' preferred profile may change throughout the course (Entwistle & McCune, 2004). Students may gravitate toward a specific profile or adopt each of the multiple profiles depending on the class (Shell & Soh, 2013). The different profiles students adopt affect end-of-term grades and indicate that some profiles are adaptive, while others are maladaptive.

Using canonical correlations, Shell and Husman (2008) identified three bidirectional multivariate dimensions with different patterns of motivation and self-regulation in college students enrolled in upper division educational psychology courses. This study confirmed the multivariate nature of students' motivational and self-regulated learning and investigated potential relations between future time perspective and control. They found that patterns of self-

regulation are motivated by students' goal orientation, among other factors (Shell & Husman, 2008).

Motivational and self-regulated learning profiles were also studied using cluster analysis. Shell and Soh (2013) identified five different learning profiles for students taking a required post-secondary computer science class. Patterns of relations between self-regulation, goal orientation, perception of instrumentality, and future time perspective formed five cluster profiles. Students who were in a required course for their major were more likely to adopt strategic or knowledge building profiles than students taking a required, non-major course. Additionally, it was found that the different learning profiles adopted can affect academic performance. Students who adopted apathetic or surface learning profiles received lower grades in the class. The five profiles identified in this study can be mapped onto the multivariate dimensions identified by Shell and Husman (2008).

Nelson et al. (2015) investigated the pattern of relations between motivational and self-regulated variables in college students. In a post-secondary foundational engineering course, it was found that students adopt either adaptive or maladaptive learning profiles. Similar to Shell and Soh (2013), cluster analysis was used and five learning profiles were identified. Strategic and knowledge building profiles are considered adaptive, while apathetic, surface learning, and learned helplessness are maladaptive learning profiles. These studies indicate that motivational and self-regulated variables are best explained with profile analysis.

Profile Variables

Self-regulation Variables

Self-regulation variables can be divided into self-regulation and knowledge building strategies. Self-regulated students actively engage in their learning through self-imposed tasks,

skills, and strategies. Students are constantly monitoring and evaluating their learning and

understanding. Knowledge building strategies are when students engage in deeper-level learning

which is supported by making connections between new and old information (Hilpert et al.,

2012).

Motivation Variables

Self-efficacy, perceptions of instrumentality, and goal orientation are motivational factors

in educational settings. Self-efficacy is a student's confidence in their ability to complete a

specific task (Bandura, 1986). Perception of instrumentality is an individual's belief about how

their present activities will help them achieve their future goals (Husman et al., 2004).

Endogenous perception of instrumentality (PIEN) is when students are motivated to learn and are

engaged in present activities because they are interesting and valued (i.e., taking a class because

students need to learn the content to reach their future goals). In contrast, exogenous perception

of instrumentality (PIEX) is when students are engaged in present activities because they need to

earn a specific grade to move forward towards their future goals (i.e., taking a required course to

fulfill major or graduation requirements).

Goal orientation theory refers to how goals affect motivation. There are two main types

of goals: mastery and performance. With a mastery goal orientation, students perceive failure as

motivating. In performance goal orientation, failure is considered negative. In a trichotomous

framework of achievement goal orientation, performance goals are split into two

branches: approach and avoid goals. Performance-approach goals are when students act in order

to receive good grades. In contrast, performance-avoid goals are when students will engage in

self-sabotaging behavior because they feel incapable of achieving their goals (Lee et al., 2010,

Midgley et al., 2000). Students with a mastery approach are motivated to learn and understand

the material. In the trichotomous framework, mastery goals are not further divided into mastery-approach and mastery-avoid goals. Mastery goals are associated with adaptive patterns of learning. Students with performance-approach goals are motivated to receive a good grade, these goals may be considered adaptive or maladaptive. Finally, students with performance-avoid goals are motivated to avoid looking unintelligent or incapable and these goals are maladaptive to learning and motivation (Midgley et al., 2000).

Life Project

Motivation can be examined through students' goals and personal projects (Marttinen & Salmela-Aro, 2012). Life project is a trait-level variable that is composed of future goals and plans and how those representations of the future shape self-concept and current actions. Life project is a future time perspective variable.

Life project has two components: identification and involvement (Coscioni, 2021). Identification refers to how future goals define the individual, similar to an identity narrative (McAdams, 2001). Individuals with high identification scores are more likely to have a clear idea of their future goals. Involvement is the extent to which present activities and actions are guided by future goals.

In sum, students adopt motivational and self-regulated learning profiles. Learning profiles demonstrate the relations between patterns of motivational and self-regulated variables, such as: self-efficacy, perception of instrumentality, goal orientation, and self-regulation strategies. Learning profiles may be adaptive or maladaptive. Adaptive profiles are high in mastery goals and endogenous perception of instrumentality, while maladaptive profiles are often characterized by high extrinsic motivation and performance goals (Nelson et al., 2015). Prior research has indicated that three or five learning profiles are often found (Shell & Husman, 2008; Shell &

Soh, 2013; Nelson et al., 2015), but the relation between learning profiles and life project has not

been researched.

Method

Participants

Participants were recruited from a required, non-major introductory science course for

non-chemistry science majors. The sample consisted of 384 students (68% female, 60% White)

who completed an online self-report survey for extra-credit points, see Table 1. For students who

did not want to participate, alternative extra credit opportunities were available. The course was

offered online and students attended synchronous lectures twice a week in addition to watching

asynchronous lecture videos. Two exams were given during the term and students completed a

cumulative final exam at the end of the course.

Table 1

Demographics of Participants

	N	$\%$
Gender		
Male	116	30.2
Female	262	68.2
Prefer not to say	6	1.6
Race and Ethnicity		
Asian	53	13.8
Hispanic	38	9.9
White	232	60.4
Multiracial/multi-ethnic	31	8.1

Other   15   3.9

Note. Total *N* = 384. Age ranged from 18 to 45 years

old (*M* = 19.31, *SD* = 2.13). Due to low numbers, we

do not report Native American, Black, and Pacific

Islander students' individually. The number of

students across these groups are 15.

Data Sources

The motivation and self-regulation measures used in this study are described below and

correspond to the motivation and self-regulation variables defined previously. See Table 2 for a

summary of variable measures and reliability.

Self-regulation Measures

Self-regulation and knowledge building strategy use was measured using the Student

Perception of Classroom Knowledge-Building scale (SPOCK; Shell & Husman, 2008). Students

were asked to rate their strategy use with a 5-point Likert-type scale.

Self-regulation strategy use was measured with 4-item subscale (α = .86). Example item:

"In [course], I make plans for how I study."

Knowledge-building strategy use was measured with a 4-item subscale (α = .77).

Example item: "As I study the topics in [course], I try to think about how they relate to the topics

I am studying in other classes."

Motivation Measures

Perceptions of Instrumentality. Perceptions of instrumentality (PI) were measured with

the Future Time Perspective scale (FTPS; Husman & Shell, 2008). Students were asked about

their perceived value of the content in the course for their future using the following two

subscales:

Perception of instrumentality exogenous was measured with a 4-item subscale (α = .91).

Example item: "The only aspect of [course] that will affect my academic future is my grade."

Perception of instrumentality endogenous was measured with a 4-item subscale (α = .88).

Example item: "What I learn in [course] will be important for my future occupational success."

Self-efficacy was measured with the eight-item subscale from the Motivated Strategies for

Learning Questionnaire (MSLQ; Pintrich et al., 1991). Students were asked to rate their

confidence in their ability to do well in the course using a 7-point Likert scale (α = .95). An

example item: "I'm confident I can do an excellent job on the assignment and tests in [course]."

Goal Orientation was measured using a trichotomous framework derived from the

Patterns of Adaptive Learning Scales (PALS; Midgley et al., 2000). Students rated themselves

with a 5-point Likert-type scale for the following constructs:

Mastery goals were measured with a 5-item subscale (α = .89). Example item: "It's

important to me that I thoroughly understand my class work."

Performance-approach goals were measured with a 5-item subscale (α = .90). Example

item: "It's important to me that other students in [course] think I am good at my class work."

Performance-avoid goals were measured with a 4-item subscale (α = .84). Example item:

"It's important to me that I don't look stupid in [course]."

Measure of Future

Future was measured with the 8-item Life Project Scale (LPS; Coscioni, 2021). Students

were first asked to write down their future goals and the age at which they hoped to achieve them

before rating themselves using a 7-point Likert-type scale. The scale was comprised of the

following two constructs:

  Involvement was measured with a 4-item subscale ($\alpha = .89$). Example item: "I'm engaged

in activities to achieve my future goals."

  Identification was measured with a 4-item subscale ($\alpha = .89$). Example item: "I have

already decided what to do with my life in the future.

Table 2.

Variable Measures and Reliability

Measure	Abbr.	# of Items	*M* (*SD*)	Skewness	Kurtosis	α
Self-efficacy	SE	8	4.25 (1.35)	-0.18	-0.60	.95
Perception of Instrumentality Endogenous	PIEN	4	3.72 (0.84)	-0.55	0.31	.88
Perception of Instrumentality Exogenous	PIEX	4	2.80 (1.00)	0.29	-0.48	.91
Mastery Goals	MG	5	4.00 (0.79)	-0.87	1.06	.89
Performance-Approach Goals	PAP	5	2.36 (1.00)	0.51	-0.28	.90
Performance-Avoid Goals	PAV	4	2.70 (1.04)	0.16	-0.64	.84
Knowledge-building	KB	4	3.57 (0.85)	-0.63	0.74	.86
Self-regulation	SR	4	3.36 (0.81)	-0.32	0.09	.77
Involvement	IN	4	5.50 (1.10)	-0.85	0.84	.89
Identification	ID	4	5.15 (1.30)	-0.74	0.30	.86

Note. Mean and standard deviation for complete sample. MSLQ (Pintrich et al., 1991), Husman et al.

(2004), PALS (Midgley et al., 2000), SPoCK (Shell & Husman, 2008).

Procedure

      Participants were first presented with an informed consent before being prompted to

reflect on and write down their life goals. This was followed by a randomized presentation of the

life project scale items. Next, participants received a randomized administration of the remaining

scales and their items. The order in which the scales were presented to the participant were

randomized, and the order of the items within each scale were also randomized. The final portion

of the survey asked participants to provide demographic information about age, gender, and

race/ethnicity.

<div align="center">

Results

</div>

      To determine what motivational and self-regulated learning profiles (profiles) were

present in this class, hierarchical and K-means cluster analysis was used. First, hierarchical

cluster analysis was used to determine the number of clusters used in the K-means analysis. A

dendrogram and icicle plot were extracted using Ward's method and squared Euclidean distance.

The data converged around five and three clusters. Three clusters best described the differences

in the data; see Table 3 and Figure 1. A statistically significant difference was found between the

three profiles ($p < .01$). Bonferroni post-hoc comparisons indicated that cluster 2 significantly

differed from clusters 1 and 3 across all variables ($p < .05$), except there was no significant

difference between clusters 1 and 2 for performance-approach goals. Clusters 1 and 3 were

significantly different for the following variables ($p < .05$): self-efficacy, perception of

instrumentality endogenous, perception of instrumentality exogenous, performance-approach,

and performance-avoid goals. Students in cluster 1 have the highest average scores across all

variables except in performance-approach, performance-avoid, and perception of instrumentality

exogenous ($N = 126$); this cluster was labeled as the mastery-oriented profile. Students in cluster

2 have the lowest average scores across all variables except perception of instrumentality

exogenous ($N = 135$), therefore this cluster was named the apathetic profile. Finally, students in

cluster 3 have on average the highest performance-approach and performance-avoid scores ($N =$

124), so this cluster is named the performance-oriented profile.

Table 3.

K-means Cluster Analysis—Cluster centers and standard deviations

	Profiles			F
	Mastery	Apathetic	Performance	
Self-efficacy	5.15(.90)	3.03(.91)	4.70(1.13)	168.33
PI Endogenous	4.18(.60)	3.09(.80)	3.95(.64)	92.11
PI Exogenous	2.15(.75)	3.24(.95)	2.96(.96)	51.31
Mastery	4.39(.52)	3.42(.77)	4.27(.59)	90.03
Performance-approach	1.80(.59)	1.94(.05)	3.37(.76)	190.13
Performance-avoid	2.10(.72)	2.35(.08)	3.70(.76)	157.41
Self-regulation	3.87(.67)	3.09(.85)	3.81(.76)	41.84
Knowledge building	3.73(.63)	2.76(.71)	3.64(.67)	83.93

Note. $p < .01$ for differences between cluster centers.

Figure 1.

Motivational and Self-Regulated Learning Profiles: Final Cluster Centers
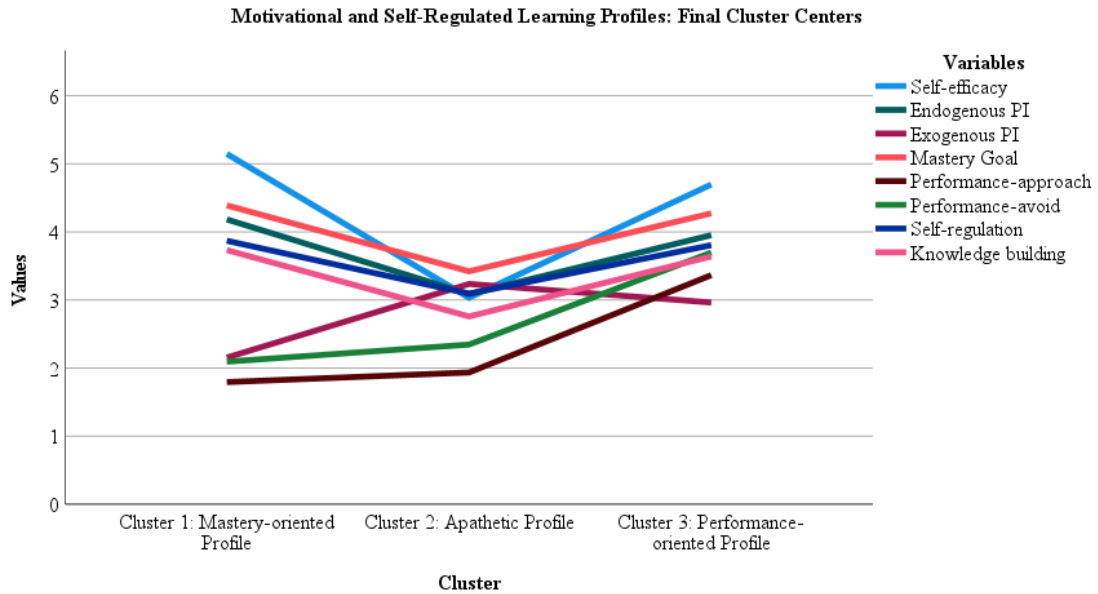
A one-way ANOVA was used to test cluster profile differences in end of term grades, see Table 4. A statistically significant difference was found between the three profiles, F (2, 382) = 29.60, $p < .01$. Post-hoc comparisons using Bonferroni's analysis indicated that students with a mastery-oriented profile earned significantly better end-of-term grades on average than students with a performance-oriented or apathetic profile ($p < .01$). Students with a performance-oriented profile had, on average, significantly better grades than the students with an apathetic profile but did not score as high as their mastery-oriented peers ($p < .01$).

Table 4.

One-Way ANOVA—Grade Differences Across Profiles

Profile	Profile Comparison	Mean Difference	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	0.78	.101	.000	0.53	1.02
	3	0.37	.103	.001	0.13	0.62

2	1	-0.78	.101	.000	-1.02	-0.53
	3	-0.40	.101	.000	-0.64	-0.16
3	1	-0.37	.103	.001	-0.62	-0.13
	2	0.40	.101	.000	0.16	0.64

Note. 1 = Mastery-oriented, 2 = Apathetic, 3 = Performance-oriented

A multinomial logistic regression was used to analyze the relationship between future thinking (involvement and identification) and the three profiles, see Table 5. For every one-unit increase in involvement score on the life project scale, the likelihood of a student with an apathetic profile being in the mastery-orientated profile increases by a factor of .44. Similarly, for every one-unit increase in involvement score on the life project scale, the likelihood of a student with a performance-orientated profile being in the mastery-orientated profile increases by a factor of .59. In contrast, for every one-unit increase in identification score, students in the apathetic and performance-oriented profiles decrease their likelihood of being in the mastery-oriented profile by factors of 1.27 and 1.55, respectively. The fit between the model containing only the intercept and data improved with the addition of the predictor variables, χ^2 (df) = 38.30 (4), $p < .01$, Nagelkerke R^2 = .11.

Table 5.

Multinomial Logistic Regression—Relationship Between Life Project Scale and Profiles

	Apathetic Profile				Performance-oriented Profile			
	B	*SE*	Wald	Exp.*B*	*B*	*SE*	Wald	Exp.*B*
Intercept	3.39**	.73	21.71		.64	.78	.68	
LPS_ID	.24	.14	2.72	1.27	.44**	.15	8.25	1.55
LPS_IN	-.83**	.18	20.21	.44	-.53**	.19	7.64	.59

Note. * *P* < .05 ** *p* < .01. Mastery-oriented profile was used as the reference category.

LPS_ID is involvement and LPS_ID is identification from the life project scale.

Discussion

The three profiles found, mastery-oriented, apathetic, and performance-oriented, are similar to profiles and dimensions found in prior goal-orientation and person-centered research (Shell & Husman, 2008; Shell & Soh, 2013; Nelson et al., 2015). The mastery-oriented students are highly efficacious, engage in self-regulated and knowledge building strategies, are motivated by being personally interested in the material they are learning and have a mastery approach to learning the course work. Students in the apathetic profile have low self-efficacy, engage significantly less in self-regulation and knowledge building strategies compared to their peers, and have the highest measures of performance-avoid goals and perception of instrumentality exogenous. Finally, the students in the performance-oriented profile engage in self-regulated and knowledge building strategies but have significantly higher performance-oriented and perception of instrumentality exogenous values than their mastery-oriented peers. The difference in end-of-term grades indicate that a mastery-oriented profile is adaptive, while the performance-oriented and apathetic profiles are maladaptive motivational profiles.

Students' future orientation is related to their motivational and self-regulated learning profiles. Higher involvement scores (i.e., students' engagement in activities that help students work toward future goals) increased the likelihood of students having an adaptive profile. Students with higher identification scores (the clarity of their future goals) were more likely to have a maladaptive profile. Future goals may support students being motivated to learn and value present activities, but students may also feel pressured by future goals to increase their performance (de Bilde et al., 2011).

      Consequently, classrooms that support students' mastery goals, strategy use, self-efficacy, perception of instrumentality endogenous, and exploration of possible futures, may facilitate adaptive motivation. Students may be more likely to engage in a mastery goal orientation and have an endogenous perception of instrumentality if they are supported to make connections between their future values and present activities (Husman & Lens, 1999). By creating an environment that encourages the development of mastery goals and decreases the emphasis on performance goals, students may be more academically successful and interested in the course (Wormington & Linnenbrink-Garcia, 2016).

Implications

      STEM instructors may be able to support student grade performance through focusing on mastery-oriented classroom practice (Anderman & Anderman, 2008). Students who are open to exploring their future careers may, in fact, be more mastery oriented in their approach to learning in introductory STEM classes. Supporting students' exploration of possible future may be one way to support mastery goal orientation. Additionally, knowing the different profiles students adopt during an introductory science course may help students identify what they may need to do to enhance their self-regulation and knowledge building strategies and understand how their future goals influence present activities.

Limitations

      This study collected data at one time-point during the term in an introductory science course. Future studies may look at how profile adoption may vary over the course of a term and if a similar three profiles are found in other introductory STEM classes. Future studies may also investigate how alternative course structure (with a decreased emphasis on grades) affects student motivation.

Conclusion

There is a national need for more students to major in STEM. Supporting student motivation and self-regulation in introductory science classes may decrease student attrition. By analyzing students' self-efficacy, perception of instrumentality, goal orientation, and self-regulation and knowledge-building strategy use in an introductory science course, three different motivational profiles were found. There were significant differences in end-of-term grades across the three profiles, indicating that mastery-oriented motivation is adaptive, while the other two profiles are maladaptive. What students do in the present to work towards their future is a strong predictor of having a mastery-oriented profile, while the clarity of future goals is predictive of an apathetic or performance-oriented profile. More work needs to be done to support students' mastery goal orientation and attend to students' thinking about the future in and outside of the classroom.

Acknowledgments

References

Anderman, L.H. & Anderman, E.M. (2008). Oriented towards mastery: Promoting positive

      motivational goals for students. In R. Gilman, E.S. Heubner, & M. Furlong (Eds.),

      Handbook of Positive Psychology in Schools. Mahwah, NJ: Erlbaum.

Coscioni, Vinicius. (2021). *A comprehensive theory of life projects* (Doctoral dissertation,

      00500:: University of Coimbra).

de Bilde, J., Vansteenkiste, M., & Lens, W. (2011). Understanding the association between

      future time perspective and self-regulated learning through the lens of self-determination

      theory. *Learning and Instruction, 21,* 332-344.

Entwistle, N., & McCune, V. (2004). The conceptual bases of study strategy inventories.

      Educational Psychology Review, *16*(4), 325-345.

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2016). Closing

      achievement gaps with a utility-value intervention: Disentangling race and social class.

      Journal of Personality and Social Psychology, 111(5), 745.

Harackiewicz, J. M., Rozek C.S., Hulleman, C. S., and Hyde, J.S. Helping parents to motivate

      adolescents in mathematics and science: An experimental test of a utility-value

      intervention. *Psychological Science, 23*, 899-906.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The elements of statistical learning: data

      mining, inference, and prediction. *Springer Science & Business Media.*

Hilpert, J. C., Husman, J., Stump, G. S., Kim, W., Chung, W.-T., & Duggan, M. A. (2012).

      Examining students' future time perspective: Pathways to knowledge building. *Japanese*

      Psychological Research, *54*(3), 229–240. https://doi.org/10.1111/j.1468-

      5884.2012.00525.x

Husman, J., Derryberry, W. P., Crowson, H. M., & Lomax, R. (2004). Instrumentality, task

    value, and intrinsic motivation: Making sense of their independent interdependence.

    Contemporary Educational Psychology, *29*, 63–76.

Husman, J., & Hilpert, J. (2007). The intersection of students' perceptions of instrumentality,

    self-efficacy, and goal orientations in an online mathematics course. *Zeitschrift für*

    Pädagogische Psychologie, 21(3/4), 229-239.

Husman, J. & Lens, W. (1999). The role of the future in student motivation. *Educational*

    Psychologist, 34(2), 113-125.

IBM Corp. Released 2020. IBM SPSS Statistics for Windows, Version 27.0. Armonk, NY: IBM

    Corp.

Lens, W., Paixão, M. P., Herrera, D., & Grobler, A. (2012). Future time perspective as a

    motivational variable: Content and extension of future goals affect the quantity and

    quality of motivation. *Japanese Psychological Research*, *54*(3), 321–333.

    https://doi.org/10.1111/j.1468-5884.2012.00520.x

Marttinen, E., & Salmela-Aro, K. (2012). Personal goal orientations and subjective well-being of

    adolescents. *Japanese Psychological Research*, *54*(3), 263–273.

    https://doi.org/10.1111/j.1468-5884.2012.00521.x

Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., &

    Urdan, T. (2000). Manual for the patterns of adaptive learning scales. Ann Arbor:

    University of Michigan.

Nelson, K. G., Shell, D.F., Husman, J., Fishman, E.J. & Soh, L. (2015). Student approaches to

    learning in a foundational engineering course: A motivational and self-regulated learning

profiles perspective. *Journal of Engineering Education, 104*(1), 74-100. DOI: 10.1002/jee.20066.

Olson, S., & Riordan, D. G. (2012). Engage to Excel: Producing One Million Additional College Graduates with Degrees in Science, Technology, Engineering, and Mathematics. Report to the President. *Executive Office of the President*.

Pintrich, P.R., Smith, D., Garcia, T., & McKeachie, W. (1991). A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ), The University of Michigan, Ann Arbor, MI.

Pulfrey, C., Buchs, C., & Butera, F. (2011). Why grades engender performance-avoidance goals: The mediating role of autonomous motivation. *Journal of Educational Psychology, 103*(3), 683–700. https://doi.org/10.1037/a0023911

Shell, D. F., & Husman, J. (2008). Control, motivation, affect, and strategic self-regulation in the college classroom: A multidimensional phenomenon. *Journal of Educational Psychology, 100*(2), 443.

Shell, D. F., & Soh, L. (2013). Profiles of motivated self-regulation in college computer science courses: differences in major versus required non-major courses. *Journal of Science Education and Technology*, 22, 899–913.

Simons, J., Dewitte, S., & Lens, W. (2000). Wanting to have vs. wanting to be: The effect of perceived instrumentality on goal orientation. *British Journal of Psychology*, *91*(3), 335–335.

Simons, J., Vansteenkiste, M., Lens, W. et al. (2004). Placing Motivation and Future Time Perspective Theory in a Temporal Perspective. *Educational Psychology Review, 16*, 121–139. https://doi.org/10.1023/B:EDPR.0000026609.94841.2f

Wormington, S. V., & Linnenbrink-Garcia, L. (2016). A new look at multiple goal pursuit: The

    promise of a person-centered approach. *Educational Psychology Review*, *29*(3), 407-445.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background,

    methodological developments, and future prospects. *American Educational Research

    Journal*, *45*(1), 166-183.