OBJECTIVE BEHAVIOURAL COMPARISON OF YOUTH AND ADULT ANXIETY: A MOBILE SENSING APPROACH

by

Nicholas Edward Murray

Submitted in partial fulfilment of the requirements
for the degree of Master of Science

at

Dalhousie University
Halifax, Nova Scotia
July 2024

Dalhousie University is located in Mi'kma'ki, the
ancestral and unceded territory of the Mi'kmaq.
We are all Treaty people.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

Background: Clinicians and researchers often use subjective measures of anxiety at isolated timepoints (e.g., retrospective self-report questionnaires), which are useful for capturing meaningful clinical differences between people and across time. However, such subjective measures are limited by recall biases and an inability to capture real-time fluctuations in anxiety-related behaviours. Smartphone mobile sensors (e.g., GPS; call logs; screen usage logs) have emerged as a method for measuring mental health-related behaviours in real-time. In adults, smartphone mobile sensors have been found to be useful for predicting symptom course and outcomes for internalizing disorders (e.g., depression; anxiety). However, research in youth with anxiety remains scarce, particularly amongst samples with clinically significant anxiety collected after the COVID-19 pandemic.

Objective: The present study sought to use the Predicting Risks and Outcomes of Social InTeractions (PROSIT) app to track anxiety-related behaviours in a sample of youth and adults with clinically significant anxiety, collected after the COVID-19 pandemic.

Methods: 65 youth (aged 15-21 years) and 96 adults (aged 26-40 years) with clinically significant anxiety (as assessed at baseline) were recruited via online advertising for a two-week study. Participants first were asked to complete baseline assessments of anxiety, depression, substance use, and demographics. Then, they were asked to use the PROSIT app for 14 days, including passive sensing (i.e., without them having to do anything) of their mobility, sociability, physical activity, screen usage, and sleep behaviours, as well as answering daily ratings of anxiety and weekly questions about sociability, physical activity, and sleep behaviours. Features were engineered from the PROSIT app and summarized into two weekly timepoints. Mobile sensing features were used in Spearman correlations, structural equation models, and random forest machine learning models to predict four outcomes: momentary anxiety (i.e., weekly summaries of the daily anxiety ratings), momentary anxiety variance (variance in weekly summaries of daily anxiety ratings), baseline anxiety, and baseline depression.

Results: Overall, mobile sensing features related to internalizing symptoms, with the associations being outcome-specific and sometimes different between age groups: for example, at the bivariate level, screen usage features predicted greater momentary and baseline anxiety in youth only; greater mobility in a structural in a structural equation model predicted greater momentary anxiety variance in youth, and lesser momentary anxiety variance in adults. Random forest models indicated that mobile sensing features assessing all five behavioural categories offered some degree of utility in predicting anxiety in youth and adults.

Conclusions: The present study indicates that mobile sensing features can predict anxiety and anxiety variance in youth and adults, and that there are measurable differences in the expression of anxiety between youth and adults, many of which are consistent with prior literature. Future studies may benefit from larger sample sizes, as the present study was too small to rigorously assess the contribution of demographic moderators, and some effects may have failed to approach statistical significance due to small sample size.

LIST OF ABBREVIATIONS USED

ASSIST	Alcohol, Smoking, and Substance Involvement Screening Test
BSTAD	Brief Screening Instrument for Adolescent Tobacco, Alcohol, and Drug Use
CAD	Canadian Dollars
CES-D	Centre for Epidemiological Studies Depression Scale
CES-DC	Centre for Epidemiological Studies Depression Scale for Children
CFI	Comparative Fit Index
CI	Confidence Interval
COVID	Coronavirus Disease
DSM	Diagnostic and Statistical Manual of Mental Disorders
GPS	Global Positioning System
MSE	Mean Squared Error
N	Number of Participants
PROSIT	Predicting Risks and Outcomes of Social InTeractions
RADAR-MDD	Remote Assesment of Disease and Relapse–Major Depressive Disorder
REB	Research Ethics Board
REDCap	Research Electronic Data Capture
RF-RFE	Random Forest Recursive Feature Elimination
RMSEA	Root Mean Squared Error of Approximation
SCAARED	Screen for Adult Anxiety Related Emotional Disorders
SCARED	Screen for Child Anxiety Related Emotional Disorders
SD	Standard Deviation
SRMR	Standardized Root Mean Square Residual
SSL	Secure Sockets Layer

ACKNOWLEDGEMENTS

I like to think of my life in four domains: academics, extracurriculars, family, and creative outlets. Therefore, I wish to thank those who are important to me in all these pursuits.

I want to thank everyone in my academic life who made these two years possible, starting with my supervisor, Dr. Sandra Meier. Sandra, I will always be grateful for your patience, mentorship, guidance, and the opportunities afforded to me working with such excellent colleagues in the PROSIT Lab. Speaking of, the PROSIT Lab is a huge place, so I think it is best if I extend my thanks to all my amazing lab mates. I am thankful as well to my co-supervisor, Dr. JianLi Wang, and my thesis committee members, Dr. Alexa Bagnell and Dr. Patricia Lingley-Pottie. I truly felt inspired and supported by all of you in pursuing this project which we each share an infectious passion for. Thank you to the Department of Psychiatry, the Faculty of Medicine, and the DeWolfe Research Fund, both for the generous scholarship support and the gift of working with the most honest, rigorous, curious scientists anywhere in the world. Thanks to Dr. Thomas Trappenberg for your excellent machine learning course and advice. Thank you to this study's participants for their effort and enthusiasm which drove me to pursue this research.

In terms of extracurriculars, I have three special thanks: to Dr. Raymond Klein, thank you for being my first scientific mentor and for always encouraging my learning in research, whether it be on topics related to attention or to psychiatry. Thank you to my supervisors in Dalhousie Mental Health Peer Support, Dr. David Pilon and Joanne Mills, as well as my colleague Ankana Boruah, who have all taught me so much which I will carry forward on my journey into clinical psychological practice. Thank you to my colleagues in the Canadian Positive Psychology Association, especially Rémi Thériault, Andrea LeFebvre, and Mihaela Zlatanovska, who have always been great joys to work with as we build up the CPPA Student Ambassador Program.

My most personal thanks go to my family, especially my parents, Rob and Joy Murray; my sister, Heidi Murray; my identical twin brother, Zachary Murray; my niece, Willow Murray, and my grandparents, Bob and Joan Murray. Thank you for always loving me in a way that leads me to aim up, even in the moments I may be unsure which way "up" will bring me.

For my creative endeavours, I thank my bandmates, Michael Butz, Jakob Hall, and Justin Heisler; John Lindsay-Botten (Music Minister, Knox United Church); our friends, fellow performers, and all who have hired us for many fun gigs. Thank you as well to the DalKings Swing Dance Society and the Halifax West Coast Swing Society. My love for jazz and blues has been renewed by the friends there with whom I love to dance lindy hop, blues fusion, and modern west coast swing.

Finally, a shoutout to two caretakers in the Tupper building: Juanita Champion and Ronald Spears. Your wit and humour were always fun, and I'll miss chatting with you both.

Are those a lot of acknowledgements? Yes. And for that I'm grateful. The story of this past two years goes beyond the halls of the Tupper building and lives on in the relationships I've developed simply by being here, at this time and place. I leave you now with two favourite quotes, from musicians whose work I always turn to for personal guidance. First, a lyric from a fellow Newfoundlander, Séan McCann, in the song "Good People" by Great Big Sea:

> *The world today can be a scary place, hard to keep your faith in the human race. We're runnin' out of trees and we're runnin' out of space, but we'll never run out of good people.*

And second, a quote from my favourite musician of all time, Chet Baker:

> *I'm definitely a romantic, I don't think life is really worth all the pain and effort and struggling if you don't have somebody that you love very much.*

CHAPTER 1. INTRODUCTION

1.1. Anxiety Disorders: An Overview

In the most general terms, anxiety disorders are united by excessive fear, while varying by their contextual specificity, behavioural avoidance patterns, physical symptoms, and cognitive styles or characteristics. Contextually, people may experience excessive fear throughout many of their daily activities (as in the case of generalized anxiety disorder), or in more restricted kinds of settings (as is true for specific phobia). Behaviourally, people may pursue avoidance by engaging in behaviours that maladaptively and only temporarily ease anxiety, such as by avoiding social situations (e.g., social anxiety disorder), or avoiding novel places altogether for fear of inescapability (e.g., agoraphobia). Physically, people may be so gripped by anxiety that they shake, get dizzy, faint, vomit, pace around, have trouble sleeping, and so on (as in the case of panic disorder). Cognitively, people may be or become intolerant of uncertainty, leading them to be excessively stressed, worried, frustrated, and even incapable of coping in response to novel or uncertain situations; they may also learn to be helpless, concluding that there is no way to help themselves or be helped in dealing with anxiety problems. In the *Diagnostic and Statistical Manual of Mental Disorders – 5th Edition*, there are 11 identified anxiety disorders, with the three most common being generalized anxiety disorder, social anxiety disorder, and agoraphobia (Szuhany & Simon, 2022). Despite their rigorous classification, anxiety disorders have high comorbidity with other anxiety and mental disorders: upwards of 60% of people with an anxiety disorder meet the criteria for another disorder (Goldstein-Piekarski et al., 2016).

1.1.1. Prevalence and Risk Factors

The lifetime prevalence of anxiety disorders is between 32.4% and 33.7%, being more common in females (38.3% to 40.4%) than males (26.4.% to 26.8%). Anxiety disorders are also generally more prevalent than mood disorders, which have a lifetime prevalence between 14.4% and 21.4% (Kessler et al., 2012). The average age of onset for an anxiety disorder is 21.3 years old – an estimate which varies widely depending on the diagnosis: Social Phobia, for example, onsets around 10.6 years, whereas generalized anxiety disorder tends to onset around 34.9 years (de Lijster et al., 2017).

Risk factors for anxiety disorders can be roughly categorized biologically, psychologically, and socially: examples of biological risk factors include female sex (Farhane-Medina et al., 2022), genetic heritability (estimated at 30-50%; Shimada-Sugimoto et al., 2015), and endocrinal dysfunction of the hypothalamic-pituitary-adrenal axis (Faravelli et al., 2012). Psychological risk factors include high trait-neuroticism (i.e., vulnerability to negative emotion) on the five factor model of personality (He et al., 2021), exposure to adverse childhood events and/or other traumas or moral injuries (Elmore & Crouch, 2020; Koenig & Al Zaben, 2021; Laugharne et al., 2010), and an external locus of control (Graham et al., 2022). Social risk factors include exposure to racism and discrimination (MacIntyre et al., 2023), low socioeconomic status (Azizabadi et al., 2022), and maladaptive parenting practices (e.g., overprotection; Brook & Schmidt, 2008). Of course, these mechanisms are not mutually exclusive, and often interact across levels of analysis: for example, dysfunction of the hypothalamic-pituitary-adrenal axis may be explained, at least in part, by exposure to adverse childhood events (Faravelli et al., 2012).

1.1.2. Anxiety Burdens, Outcomes, and Interventions

Estimates of the burden of anxiety disorders in Canada are based primarily on generalized anxiety disorder: roughly half of Canadians with generalized anxiety disorder also experience depressive symptoms, and the treatment needs of half of those people are, according to their own self-report, incompletely met or completely unmet (Louise et al., 2017). Approximately 208,000 Canadians are hospitalized each year for a mental health disorder, with the five most frequent hospitalizing diagnoses including anxiety disorders, among others such as substance use disorders and personality disorders (Canada, 2019).

It is challenging to create an exhaustive and/or wholly accurate list of the outcomes of anxiety disorders, for two reasons: first, anxiety disorders are neither sufficiently treated nor diagnosed (Vermani et al., 2011; Weisberg et al., 2007). Second, the outcomes measured in the literature are heterogenous, in large part because of variance in the inclusion and exclusion criteria of available studies: Goldstein-Piekarski et al. (2016) demonstrated that upwards of 92% of people pursuing clinical intervention for their anxiety may be excluded from clinical trials. The heterogeneity introduced by such highly selective (and inconsistent across studies) inclusion/exclusion criteria is exacerbated by false assumptions about mental disorder co-occurrence. The authors give an example showing three times as many posttraumatic stress disorder patients have major depressive disorder compared to major depressive disorder patients who have posttraumatic stress disorder. It is common for researchers to select people exclusively with a mood disorder, and people exclusively with an anxiety disorder, to estimate clinical outcomes by comparison. Those studies are, unfortunately, selecting quite a non-representative sample of anxiety disorder patients to recruit.

With the aforementioned limitations in mind, there are numerous negative outcomes associated with anxiety, including: increased risk of social isolation (Narchal, 2017), worsened sleep, cognitive function, and depressive symptoms (Xiao et al., 2023), higher occurrence of cardiovascular disease (Celano et al., 2016), and even increased risk of suicide (possibly for all anxiety disorders aside from obsessive compulsive disorder; Kanwar et al., 2013). Furthermore, when anxiety disorders co-occur with other disorders, their respective symptoms can be mutually exacerbating: for example, 18% of people with a substance use disorder have a comorbid anxiety disorder, and 15% of people with an anxiety disorder have a comorbid substance use disorder (Grant et al., 2004). In their co-occurrence, anxiety and substance use problems exacerbate each other, perhaps through substance use inducing an anxious state, which people then try to cope with by using more substances (Stewart & Conrod, 2008). People with anxiety disorders are more likely to die both naturally and unnaturally than people without anxiety disorders, and the risk of unnatural death is especially inflated in the presence of a comorbid depressive disorder (Meier, et al., 2016a). People with anxiety disorders are also nine times more likely to have bipolar disorder (Meier, et al., 2016) – an effect which may be driven, at least in part, through co-occurrence of attention deficit hyperactivity disorder (Meier et al., 2018).

Fortunately, a variety of psychotherapeutic and pharmacological interventions for anxiety are effective. Canadian clinical practice guidelines indicate several lines of treatment. First-line treatments include cognitive behavioural therapy techniques. Psychotherapeutic approaches generally emphasize progressive fear exposure to treat avoidance, as well as techniques such as cognitive restructuring (i.e., reframing cognitions which otherwise induce anxious states) and reducing negative reinforcement habits like checking behaviours. Ironically and despite the empirical evidence to the contrary, even many psychologists have been shown to think exposure therapy worsens anxiety symptoms by inducing a transient anxious state (Becker et al., 2004; Pittig et al., 2019). There are also a number of first-line pharmacotherapies, such as antidepressants (selective serotonin reuptake inhibitors and selective norepinephrine reuptake inhibitors; Katzman et al., 2014). Second- and third-line treatments include other pharmacotherapies which tend to have worse side effects and/or less evidence for effectiveness than first-line treatments: for example, benzodiazepines are recommended starting as a second-line treatment, particularly because of risks of addiction and withdrawal side effects (Starcevic, 2022). Third-line pharmacotherapies also include some antidepressants (e.g., fluoxetine) and anticonvulsants, which current guidelines claim have very minimal evidence for effectiveness (Katzman et al., 2014). In general, pharmacological interventions for anxiety are more effective the earlier in disease onset patients obtain treatment (Altamura et al., 2008).

1.1.3. Youth Anxiety

Youth anxiety is on the rise. During the COVID-19 pandemic, global estimates of youth with anxiety problems doubled to 1/5 youth (Racine et al., 2021). The average age of onset of anxiety disorders is 21.3 years, but this estimate varies depending on the particular disorder: some diagnoses (e.g., social anxiety) typically onset around 15 years of age, whereas others onset during the transition to young adulthood (around 21.1 years) and during adulthood (around 34.9 years; e.g., agoraphobia; generalized anxiety; (Lijster et al., 2017)).

Explanations of the present rise in youth anxiety invoke biological, social, and psychological causes: youth are biologically vulnerable to anxiety disorders, in part because of brain changes during puberty which increase sensitivity to fear stimuli; they are more vulnerable to anxiety via socialization, as they are biased towards imitating fear responses in adults; psychologically, youth are higher in trait-neuroticism, making them generally more susceptible to negative emotion (Vasey et al., 2014). Youth anxiety comes with many burdens, such as poor performance in school, withdrawal from enriching opportunities, and even suicide (Klaufus et al., 2022). Due to their characteristically chronic course, long-term negative outcomes associated with anxiety are exacerbated by earlier onset: anxiety onset in adolescence is associated with later substance use problems, suicidality, and other mental health problems such as depression

(Woodward & Fergusson, 2001). Earlier studies using the DSM-IV classifications found that 50% or more of youth who experience subclinical anxiety may meet the classification for another mental disorder in adulthood (Goodwin et al., 2013).

When anxiety disorders are diagnosed and treated early, outcomes are significantly better. Cognitive-behavioural interventions have, in meta-analyses, been demonstrated to be useful for children as young as preschool age (Fisak et al., 2023; Howes Vallis et al., 2020).

1.1.4. Comparing Youth and Adult Anxiety

As discussed earlier, anxiety disorders vary by their contextual specificity, behavioural avoidance patterns, physical symptoms, and cognitive styles or characteristics. Youth and adult anxiety problems are comparable in most of these domains as well. Contextual differences in youth and adult and anxiety are apparent from factor analyses of clinical assessment questionnaires: the Screen for Child Anxiety Related Emotional Disorders (SCARED), for example, has a social anxiety subscale, as well as a subscale for school avoidance behaviours (e.g., "I am scared to go to school"; Behrens et al., 2019). The Screen for Adult Anxiety Related Emotional Disorders (SCAARED), however, is missing an analogous subscale for school avoidance (Angulo et al., 2017). These factor analytic differences are consistent with other findings that Social Anxiety Disorder symptoms may be more school-specific in youth (Beidel et al., 1999), whereas they are more broadly found in work and communal contexts in adults (Ruscio et al., 2008). Behaviourally, in youth with anxiety disorders who are presented with something they fear, avoidance is exacerbated in participants who are high in trait anxiety (Lebowitz et al., 2015). Such moderation of avoidance behaviour by trait anxiety may be more specific to youth than adults, since trait anxiety may moderate avoidance learning in youth but not in adults (Klein et al., 2020). In contrast, youth and adult anxiety seem to be characterized by the same cognitive styles (e.g., rumination; Garnefski et al., 2002). Nevertheless, it remains likely there are cognitive differences in youth and adult anxiety, given that a third of youth do not seem to respond to cognitive behavioural therapy for anxiety (Kendall & Peterman, 2015), while around half of adults in treatment are non-responsive (Loerinc et al., 2015). In addition, though youth with anxiety do not respond as quickly to cognitive behavioural therapy as adults, a meta-analysis suggests anxiety problems are less severe at follow-up in youth than in adults (Barry et al., 2018).

1.2. Measuring Mental Health-Related Symptoms and Behaviours

Mental constructs can be measured subjectively and objectively. Subjective measures are limited to self-report methods, such as standardized questionnaires. Objective measures are harder to define, in part because they may include all things that do not fall under the self-report umbrella. As this thesis will combine both approaches, it will be useful to review the theoretical foundations and limitations of specific self-report and objective behavioural measurement tools to be used, including retrospective self-report, ecological momentary assessment, and mobile sensing/digital phenotyping.

1.2.1. Retrospective Self-Report Questionnaires

Self-report is a foundational methodological school in psychology and psychiatry, responsible for the development of clinical questionnaires (e.g., the SCARED; Behrens et al., 2019), personality inventories (e.g., the five-factor model; He et al., 2021) and more. The guiding intuition of self-report is captured within the factor analytic statistical tools used to validate self-report questionnaires: people vary in how they answer self-report questionnaires, and this variance between individuals can be explained by positing psychological constructs which are derived as correlational patterns between questionnaire items. For example, a person high in trait-neuroticism is likely to answer “Agree” or “Strongly Agree” to the items “I often feel blue” and “I am typically unhappy” while being unlikely to “Agree” with the item “I seldom feel blue”.

The most common self-report tools used in psychology and psychiatry are retrospective self-report questionnaires. For example, on the SCARED, youth are asked to recall the last three months, and rank survey items (e.g., “I don’t like to be with people I don’t know well”) from zero (“Not True or Hardly Ever True”) to two (“Very True or Often True”). Scores across items are summed, with the maximum score on the SCARED being 82. Psychometricians determine the cut-offs for clinical relevance on retrospective self-report questionnaires by observing what scores best predict some clinically useful outcome. The SCARED cut-offs were determined by observing the scores typical of children who met DSM-IV classifications for an anxiety disorder, compared to children who met classifications for non-anxiety disorders. Differences between anxiety disorders were captured through factor analyses revealing subscales that could discriminate specific anxiety classifications, such as social phobia (Birmaher et al., 1997).

The clinical utility of retrospective self-report comes from a statistical fact underlying the validation process: these questionnaires quantify individual differences well while maintaining minimal within-subjects variance. Low within-subjects variance relative to higher between-subjects variance is necessary for producing reliable measures of individual differences (Hedge et al., 2018), and as a consequence, for predicting clinical outcomes which rely on estimations of individual differences. This statistical fact is what makes longitudinal changes in the same individual(s) on retrospective self-report questionnaires meaningful.

Despite their clinical utility for classification and for detecting longitudinal changes in symptomatology, retrospective self-report questionnaires can be incomplete and biased measures of mental health. First, since they are retrospective assessments, they do not produce within-subjects variance in real time. If anxious episodes are time-dependent such that there is a symptom severity threshold for detecting their onset, then measures which do not assess symptomatology in real-time are unlikely to detect it. Detection of real-time changes in symptomatology could be useful for targeting patients for treatment when they need it most. Deep learning models have been used, for example, to detect symptom changes in real-time in adolescents with anxiety disorders using actigraphy (Jacobson et al., 2021). Second, retrospective self-reports are biased and incomplete. The most well-documented example of bias in retrospective self-report is the recall bias: when people are asked to recall their mood (e.g., in the last two weeks), they reliably recall more negative mood than they may have reported in real-time mood assessments (Sato & Kawahara, 2011). There are also other biases, such as social desirability bias, with some research indicating that people who are high in social desirability may underreport their anxiety and depression in retrospective self-report questionnaires (Komarahadi et al., 2004).

In sum, retrospective self-reports are clinically useful because they can reliably estimate individual differences and longitudinal changes in symptomatology, both of which are valuable for diagnosis and treatment planning. However, they fail to assess within-subjects variance in relatively short time frames and are prone to biases such as recall bias. Real-time assessments of mental health are therefore promising for supplementing retrospective self-reports.

1.2.2. Ecological Momentary Assessments

Ecological momentary assessments (sometimes referred to as “experience sampling”, or “contemporaneous self-report”) are rapid, in-the-moment assessments of mental health symptoms and behaviour. Often, they are implemented in smartphone apps, which notify patients and research subjects to complete an assessment on their smartphone device. Ecological momentary assessments may be “event-contingent” so that they can be completed only when something relevant to the assessment has occurred; they may also vary in their frequency (how often people are prompted to complete them) and what time of day they must be completed (Hall et al., 2021).

Ecological momentary assessments address certain limitations of retrospective self-report questionnaires. First, they can produce within-subjects variance by sampling multiple time points at closer, more frequent intervals than retrospective self-report questionnaires. Fluctuations in daily mood ratings can be used to classify mental health problems, such as depression (Kim et al., 2019), as well as to predict longitudinal changes in depressive severity measured via retrospective self-report (Bai et al., 2021). Second, ecological momentary assessments minimize recall bias, which makes them useful for assessing how much a current mood state may influence retrospective self-reports of mood (Mengelkoch et al., 2024). Third, in addition to assessing within-subjects variance, they can also assess within-event variance in mental health: by adjusting the “event contingency” and frequency of assessment (Hall et al., 2021), researchers can assess how mental health problems change across days, and throughout different times of the day, which has been especially useful in analyzing how within-day fluctuations of anxiety relate to same-night sleep quality (Cox et al., 2018).

Despite addressing these limitations of retrospective self-report, ecological momentary assessments have their own limitations as subjective assessments of mental health. Ecological momentary assessments generally converge more with objective assessments of mental health-related behaviour than retrospective self-reports, especially for substance use behaviours which can be biologically substantiated (Shiffman, 2009). However, ecological momentary assessments still may diverge from objective measures: for example, the more physically active people are (according to accelerometer measures), the less physical activity they report in ecological momentary assessments (Bruening et al., 2016). Ecological momentary assessments may also produce "reactivity", such that participants' mood and/or substance use patterns can improve through regular self-monitoring (Isaacs et al., 2021). It may be useful, therefore, to supplement retrospective self-report and ecological momentary assessments with objective measures of mental health.

1.3. Mobile Sensing and Digital Phenotyping

In sum of previous sections: retrospective self-report questionnaires are useful for assessing clinical outcomes but are limited by bias and a lack of real-time, within-subjects variability. Ecological momentary assessments are less biased than retrospective self-reports and obtain real-time, within-subjects variability, but they still have some bias and diverge from objective assessments of mental health. Mobile sensing and digital phenotyping have recently emerged to supplement subjective assessments of mental health, by measuring mental health-related behaviours in real-time.

"Phenotypes" are sets of defining characteristics belonging to biological organisms. Richard Dawkins argued in *The Extended Phenotype* that phenotypes are not just limited to biological characteristics, but also to behavioural ones. So, just as a beaver may express an “extended phenotype” through its characteristic dam-building activities (Dawkins, 1999), people may express extended phenotypes digitally, such as in how they use social media. Digital phenotypes, therefore, can be assessed with mobile sensing (Montag & Baumeister, 2023). The terms “digital phenotyping” and “digital phenotype”, then, may be used either as a verb or as a proper noun: digital phenotyping can involve, for example, using a GPS to track the mobility of patients with depression. They may also express a digital phenotype, such that they are more prone to certain maladaptive social media use behaviours like doom scrolling (passively viewing social media in a ruminative state). For the sake of clarity, “mobile sensing” will herein be used as a synonym for the verb form of “digital phenotyping”. Furthermore, the present thesis will focus on smartphone mobile sensing, wherein smartphone sensors (e.g., GPS; light/noise sensors) are used to digitally phenotype mental health-related behaviours, such as sleep and physical activity.

In general, mobile sensing technologies are well-tolerated by people with mental health problems: Nicholas et al. (2019) found adults conceptualized mobile sensing data as being either “health information” (e.g., sleep and exercise features) and “personal data” (e.g., call logs). Participants overall were more hesitant to share personal data than health information, but these effects did not vary with anxiety or depression symptoms. Orr et al. (2023) extended Nicholas et al.'s (2019) findings, showing that parents and adolescents were more comfortable sharing health information and that both groups become less comfortable the more intricate the data collected are, with this effect being stronger for adolescents than for their parents. With these findings in mind, it is unlikely for the present thesis that symptom severity will be a limiting factor in recruitment given comfort with mobile sensing seems to vary with the kind of data collected, but not with anxiety severity.

Mobile sensing avoids the recall biases of retrospective self-report and supplements ecological momentary assessments with objective, real-time behavioural data. However, it is not a panacea: just as ecological momentary assessments can produce assessment reactivity, mobile sensing may produce a “Hawthorne effect”, whereby people change their behaviour in response to being monitored or measured. Some authors (e.g., Lan et al., 2022) have argued that passive sensing likely reduces Hawthorne effects, but evidence on this front is wanting. The present study will not assess Hawthorne effects, so it is worth noting upfront as a potential limitation.

1.3.1. Mobile Sensing in Youth and Adults

Mobile sensing has been used to objectively phenotype various mental health disorders: for example, Bai et al. (2021) used an Android smartphone mobile sensing app to track mental health symptoms and behaviours in 334 major depressive disorder outpatients in Bejing. Participants completed standardized assessments of depression at multiple time points across 12 weeks and answered the Patient Health Questionnaire biweekly in the app. They answered a daily mood rating on a visual analogue scale in the phone app, which also passively (i.e., without participants having to do anything) measured de-identified call, text, and app usage log data, as well as their location (via GPS), and screen time. From these mobile sensing measures, the researchers engineered various behavioural features, such as the average number of phone calls made per day and the average amount of screen usage per day. The researchers found that the mobile sensing features could be used to predict changes in depressive symptom severity across 12 weeks with accuracy as high as 82.67%. Other studies have shown promising results in predicting depressive symptoms using mobile sensing data (e.g., Asselbergs et al., 2016) and some have shown that personalizing the prediction algorithm improves model fit (Cho et al., 2019).

Literature on mobile sensing with youth samples is generally sparser than mobile sensing with adult samples. There is certainly a lack of studies directly comparing youth and adult samples using the same mobile sensing app, which is the primary limitation of the available literature the present thesis seeks to address. It is worth briefly reviewing, then, the available literature on mobility, sociability, screen usage, physical activity, and sleep studied in both youth and adults and some of the scant findings analyzing age-related differences in mobile sensing for mental health problems.

1.3.2. Mobility Features

Mobility features quantify how frequently people visit novel locations, and/or how diversely they spread their time across the various locations they visit. Mobility features (engineered from GPS data) are certainly the most well-established features in the available mobile sensing-mental health literature. In the European Union, there is a research program called Remote Assessment of Disease and Relapse–Major Depressive Disorder (RADAR-MDD). This program has demonstrated in adults with MDD that longitudinally, worsened depression is associated with less smartphone-measured mobility (e.g., more time spent at home; lower likelihood of travelling to novel locations, etc.; Zhang et al., 2022). The picture for anxiety is more complicated. For example, in college aged youth with social anxiety problems, more locations visited is associated with less social anxiety (Boukhechba et al., 2018). Other studies with youth, however, have found that anxiety is associated with more locations visited whereas depression was associated with fewer locations visited in the same sample (MacLeod et al., 2021).

There are several candidate explanations for the inconsistent mobility findings. One explanation highlights that associations of anxiety with increased mobility were discovered during the height of the COVID-19 pandemic (e.g., MacLeod et al., 2021). Given that mobility during a pandemic implies increased risk of disease exposure, it may be unsurprising that mobility at that time was associated with anxiety. Another explanation says that the effect of mobility may depend on the kind of locations people with anxiety are visiting. People with social anxiety, for example, may visit grocery stores more frequently so that they can purchase food to eat at home alone (Boukhechba et al., 2017). However, the importance of the type of location (e.g., spiritual, work, or school locations) may not be useful for classifying depression or generalized anxiety, as authors in this field have noted that location categories may not be sufficiently relevant to mental health (Saeb et al., 2017). Finally, authors such as Renn et al. (2018) argue mobile sensing researchers conflate mobility and physical activity. Mobility features quantify participants' tendencies to visit few or many locations, and/or their tendency to spend time diversely amongst the locations they visit. Physical activity features, however, quantify the intensity and diversity of the participants' activities. A highly mobile person may not be very physically active (e.g., someone who drives a lot but does not exercise), and a highly physically active person may not be very mobile (e.g., someone who works from home and who uses a home exercise gym). Perhaps, then, the inconsistent effects are due to authors not considering both behavioural categories in their analyses (e.g., MacLeod et al., 2021; Saeb et al., 2015, 2017). In short, depression is probably associated with less smartphone-measured mobility, and the effects for anxiety are too inconsistent to justify a strong prediction. Furthermore, predictions for direct comparisons of youth and adults are complicated: if youth are more mobile than adults, then associations of internalizing symptoms and smartphone-measured mobile may be stronger in youth. However, if youth visit fewer locations per day than adults, the effect could be the opposite (perhaps youth tend to have more "hangout" spots where they spend a lot of time). Furthermore, it may be that youth with social anxiety are less mobile and have higher screen usage. Given this lack of clarity, exploratory analyses in a direct comparison of youth and adults are necessary.

1.3.3. Sociability Features

Sociability features in mobile sensing typically include call logs, amongst other less common features such as app usage and text logs (Bufano et al., 2023). Increased sociability detected via mobile sensing has been associated with trait-extraversion, such that people who are lower in extraversion engage in fewer sociability behaviours detectable via a smartphone (e.g., phone call frequency; Harari et al., 2020). In some studies, fewer calls have been associated with worsened internalizing symptoms in youth (e.g., Cao et al., 2020; MacLeod et al., 2021), whereas other studies have shown worsened internalizing symptoms with greater calls and phone usage in youth (e.g., Messner et al., 2019). Mixed findings may be explained, for example, by the fact that that some youth may experience addictive behaviours in relation to their smartphone use, which may lead them to be less sociable offline. In adults with bipolar disorder, fewer and shorter phone calls have also been associated with depression (Gillett et al., 2021). Similar results have been found using machine learning models to predict depression with incoming and outgoing call frequency, but the age of the sample was unspecified (Tlachac et al., 2021). Given these findings, youth and adults with anxiety may show associations of internalizing symptoms and less sociability measured by call features, with the effect(s) probably being driven by depression.

1.3.4. Screen Usage Features

Screen usage is typically measured with total screen time, and/or total screen unlock events. In general, recent reviews and meta-analyses agree that greater screen usage predicts greater internalizing symptoms in youth (Eirich et al., 2022) and adults (Santos et al., 2024). There are, however, some age-related caveats: Sharifian et al. showed youth may unlock their phone more frequently to check social media. So, associations of screen unlocks to internalizing symptoms in youth could depend on mental health effects of social media use. Furthermore, if youth are more likely to check social media as an addictive behaviour, then weaker or less robust effects of screen usage to anxiety in adults might be reasonable to expect (2021).

1.3.5. Physical Activity Features

As the present study will work exclusively with a smartphone app, it is worth noting that there are a wide variety of methods for assessing physical activity via mobile sensing. Some of these include smartphone apps (e.g., fitness apps) which label activities based on various algorithms, accelerometry (either wrist-worn or smartphone-recorded), GPS, and so forth. It is beyond the scope of the present thesis to analyze all available modalities of physical activity assessment, but interested readers are encouraged to see Bufano et al. (2023) for a review.

Physical activity has been repeatedly associated with improved anxiety and depression in youth and in adults (including in randomized trials), though authors may disagree about how intense the physical activity should be (Awick et al., 2017; Klemmer et al., 2023). During the COVID-19 pandemic and independently from mobility features, increases in sedentary time as assessed by mobile sensing predicted worsened mental health in college students (Mack et al., 2021). Samples of youth and adults collected before, during, and since the COVID-19 pandemic who are more sedentary are more likely to be depressed, and experience more voltatile changes in their self-reported mood (Chikersal et al., 2021; Mack et al., 2021; Meyerhoff et al., 2021; Wang et al., 2018; Xu et al., 2022). Beyond sedentary time, there are some studies showing the variety of physical activities participants pursue may relate to anxiety: youth college students with larger variance in the speed of their stride, for example, self-report more social anxiety than their counterparts with more steady stride patterns across two weeks of data collection (Jacobson et al., 2020). Boukhechba et al. (2017) also found that in college students, greater variance in the types of physical activities participants engaged in predicted lower social anxiety. The literature investigating variance in the kinds of physical activities participants pursue independently of overall sedentary time is sparse, and so for present purposes a strong hypothesis on this front is unjustifiable. Likely, youth and adults who are more sedentary will report greater anxiety and depression.

1.3.6. Sleep Features

A recent review found screen events (screen-active/on vs screen-inactive/off) were the most common feature for detecting sleep interruptions, since people may check their phone during a wakeful moment throughout the night. Other common features include ambient light and noise sensors, as well as GPS sensors, which measure sleep as the times with the lowest light, noise, and physical activity (Alamoudi et al., 2023). Another recent review showed that sleep-related mobile sensing research in broad populations with anxiety, depression, or psychosis is dominated by pilot studies, and most of these studies involve smartphone-based self-report of sleep rather than real-time behavioural measurement of sleep states or sleep disruptions (Aledavood et al., 2019). There is, however, consistency amongst the limited studies with mobile sensing data in this subfield: in adults being monitored with a mobile sensing ring, greater time spent sleeping and/or lying in bed has been associated with higher depressive symptoms, whereas waking up more frequently throughout the night has been associated with higher anxiety symptoms (Moshe et al., 2021).

Sleep features are generally strong at predicting both mood and self-reported sleep quality, however even amongst studies which include youth and adults (e.g., Niemeijer et al., 2022; age range of 17-32 years), comparisons of sleep and mental health features across ages are lacking. Generally, more sleep interruptions predict higher anxiety and depression (Choi et al., 2024). However, some authors emphasize that there are large individual differenes in associations of sleep features and mental health. For example, Melcher et al. (2023) found large positive correlations between sleep time and depressive symptoms in some participants, and equally large negative correlations between the same variables in others. In a prospective cohort study, adolescents were more likely to develop sleep-related problems (e.g., restlessness; difficulties getting to sleep/sleep onset latency) if they checked their phone throughout the night (Foerster et al., 2019). Given that addictive smartphone behaviours are more prevalent in youth than in adults (Ratan et al., 2022) and that smartphone addiction is associated with worsened sleep (Sohn et al., 2021), it is likely that youth may check their phone more throughout the night, so this requires exploration. This may explain why college students who text more during the night tend to self-report higher anxiety (Mendu et al., 2020). Therefore, associations of sleep interruption assessed by screen usage features and mental health symptoms in youth are likely stronger, particularly amongst female youth who may be more likely than males to check their smartphone throughout the night (Maurya et al., 2022). Other features assessing sleep interruption, however, may show opposite associations, as some studies have found more exposure to darkness at night predicts greater anxiety, or that more anxious people are in the dark less frequently (Fukazawa et al., 2019; Matteo et al., 2021).

1.4. The Present Study

The present study seeks to directly compare youth and adults with clinically significant anxiety, using the Predicting Risks and Outcomes of Social InTeractions (PROSIT) mobile sensing app. The PROSIT app is being developed on iOS and Android smartphones, and includes a variety of sensors to proxy mobility, sociability, sleep, physical activity, and screen usage behaviours. The PROSIT app has been used to monitor youth with anxiety, depression, and attention deficit hyperactivity disorder symptoms (MacLeod et al., 2021); youth sleep during the COVID-19 pandemic (Marin-Dragu et al., 2023); healthy youth social media users (ongoing), amongst other projects. Though the development of the PROSIT app has been focused on youth populations, most mobile sensing studies are with adults. Therefore, it will be useful to directly compare youth and adults with the same app, using features that are common across mobile sensing studies – specifically, those assessing sleep, sociability, and physical activity. Youth are defined here as people between the ages of 15 and 21 years old, and adults as people between the ages of 26 and 40. These age ranges were chosen for two reasons: first, anxiety disorders related to brain maturation and development typically onset before 25 years old (de Lijster et al., 2017). Second, anxiety disorders are most prevalent in middle age (Bandelow & Michaelis, 2015), and begin to decline in prevalence thereafter (Ramsawh et al., 2009). Therefore, the chosen age ranges of 15-21 and 26-40 are conservative estimates of adolescence and adulthood, while selecting from an adult population most likely to have clinically significant anxiety.

There are a variety of qualities that make this study unique and strong as a thesis seeking to contribute to mobile sensing research. First, this study is (to our knowledge) one of the first studies to directly compare youth and adults with clinically significant anxiety problems, using the same mobile sensing app. Second, it is the first to perform such a comparison in a sample collected after the COVID-19 pandemic. Third, in a literature dominated by adult studies, the PROSIT mobile sensing app is unique as an app co-designed by youth, with considerations such as youths' comfort with various types of mobile sensing data being included in the codesign of the app (Orr et al., 2023). And fourth, we are the first to attempt to replicate and extend the methods (and potentially the findings) of MacLeod et al. (2021) using the PROSIT app.

1.4.1. Aims, Questions, and Hypotheses

Given that this is likely one of the first studies to directly compare youth and adults with anxiety problems using the same mobile sensing app, the overarching aim is to explore comparisons of youth and adult anxiety using mobile sensing features. Following from this theme are three analytical aims, each with complimentary research questions and hypotheses. Our first analytical aim is a replication attempt: common results from mobile sensing studies will be investigated in the present sample of youth and adults. MacLeod et al. (2021) demonstrated that the PROSIT app significantly predicts baseline anxiety and depression in a sample of youth with anxiety, depression, and attention deficit hyperactivity disorder symptoms. We will attempt to replicate and extend their findings, now with a sample of youth and adults selected specifically for clinically significant anxiety problems using the PROSIT app. Second, a theory test: it is hypothesized that the various features derived from the PROSIT app's mobile sensors will cluster into latent factors representing mobility, sociability, sleep, screen usage, and physical activity. Therefore, we will use factor analyses to derive whichever factors we can, and structural equation models to test each factor's relation to anxiety in youth and adults. Third, an atheoretical approach with machine learning: using random forest models, we will determine the best set of features for predicting momentary anxiety in youth and in adults, and which features are most capable of classifying youth and adults based on age-related behavioural differences. To better guide our aims, we have several research questions and complimentary hypotheses:

RQ1. Does mobile sensing predict anxiety in youth and in adults?

H1. Mobile sensing data will predict baseline anxiety (as measured by retrospective self-report) and momentary anxiety (as measured by ecological momentary assessment).

H2. Mobile sensing data will improve the fit of models predicting anxiety in both youth and in adults, on top of contributions of demographic covariates.

RQ2. Are there differences in how youth and adults express anxiety, as measured by mobile sensing features?

H3. There will be bivariate correlational differences and path differences relating anxiety to mobile sensing features, as tested via Spearman correlation and structural equation analysis.

RQ3. What are the most useful mobile sensing features for predicting anxiety in youth and in adults?

CHAPTER 2. METHODS

2.1. Participants

2.1.1. Ethical Approval

This thesis project was approved by the Dalhousie Health Sciences Research Ethics Board, under two ethics filings: REB#2022-6440 (for the adult sample) and REB#2023-6634 (for the youth sample). All procedures were conducted in accordance with the Declaration of Helsinki and the Canadian Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans – TCPS 2.

2.1.2. Recruitment

Inclusion and Exclusion Criteria. Youth between 15-21 years old and adults between 26-40 were recruited for this study between May 2023 and May 2024. In that timeframe, we aimed to recruit 130 participants in each group, as a power analysis indicated that was sufficient to account for expected attrition of 20% with sufficient statistical power ($\beta = 0.80; \alpha = .05$) to detect a small-medium effect size in a structural equation model. We invited youth who were between the ages of 15-21 years old, and adults between the ages of 26-40. The only other inclusion criterion by which the samples differed was the anxiety score cut-off: youth participants must have scored ≥ 25 on the SCARED, whereas adult participants must have scored ≥ 23 on the SCAARED. These are the validated clinical cut-offs for the respective youth and adult versions of the questionnaires. Otherwise, all participants were required to meet the following inclusion/exclusion criteria: have access to an iOS or Android smartphone (to download the PROSIT app); no serious cognitive delay, suicidality, or language disorders identified during screening (to avoid confounding the mobile sensing data with behavioural effects associated with these conditions); no medical incident/accident which would prevent participants' engaging in their routine activities for the two weeks of mobile sensing data collection.

Study Advertising. Participants were recruited with paid online advertising, via the Meta platforms Facebook and Instagram. Advertisements were initially boosted within Canada. However, they could not be restricted only to viewers within Canada, so they were viewable in other countries. Advertisements were paid to boost reach, with a chosen daily reach of 500-2000 people. As it is forbidden in Canada to target youth populations with paid advertising on their basis of their expressed interest in mental health topics, Meta ads for youth and adults were initially separate: for adults, we were able to restrict the ads to adults interested in mental-health related Facebook and Instagram content. However, this specification did not obtain adult participants any more effectively than the broader, age-based advertising for youth. Therefore, we later combined the ads for youth and adults simply to target Meta platform users between 15 and 40 years of age.

Prevention of Robot Sign-Ups. We programmed a small library of functions in the *R* programming language to deal with robot sign-ups to the study through the "REDCap" (Garcia & Abrahão, 2021) application programming interface (API). Sign-ups were automatically prevented from further participation (i.e., not sent any questionnaires nor app download instructions) if they met any of the following criteria: repeat sign-ups with inconsistent demographic information (e.g., signing up first with a birth year outside the required age range, and then again with a birth year in the required age range); missing signatures on the consent form; impossibly fast response times (particularly, less than a one minute response time to answer the consent form); multiple sign-ups in a row with identical contact information (e.g., email) or anxiety questionnaire responses. We also implemented a non-automatic screening method, whereby a human researcher (N.M., the author) screened virtual consent form signatures for obvious forgery (e.g., many sign-ups in a row with identical signatures but different contact and demographic information).

2.1.3. Procedure

Interested ad viewers were directed to sign-up and provide consent via “REDCap”, which is a secure online platform hosted on a server at Dalhousie University (Garcia & Abrahão, 2021). On the landing page and throughout the consent survey, respondents read about the study procedure, risks and benefits, and purpose of the research, and were given multiple opportunities to ask questions before proceeding with signing up for the study. Ineligible respondents were automatically notified of their ineligibility and thanked for their time. As the Health Sciences Research Ethics Board required, we assessed baseline anxiety in the consent form (to save participants time, since anxiety was a screening factor), eligible respondents were sent a link upon completing consent to baseline assessments of demographics, depressive symptoms, and substance use habits. Upon completing these baseline assessments, participants were sent instructions for using the PROSIT app for 14 days. The REDCap surveys were brief (15-20 minutes to complete in total) and active use of the PROSIT app was less than two minutes per day. Participants were compensated upon completing study requirements with an Amazon e-gift card in the amount of $20 CAD.

2.2. Materials

2.2.1. Demographics

We assessed a variety of mental health-related demographic covariates, including age, biological sex (i.e., sex assigned at birth; Maeng & Milad, 2015), gender identity (Stoyanova & Hope, 2012), sexual orientation (Waller et al., 2023), ethnicity, location (urban/suburban/rural/remote; Romans et al., 2011), mental health diagnoses (van Santvoort et al., 2015), and socioeconomic status. For youth, socioeconomic status was proxied using parental education, as is commonly recommended in the youth psychiatric literature (Hauser, 1994). For adults' socioeconomic status, we used a question assessing annual household income from the Canadian Longitudinal Study on Aging (Raina et al., 2009).

2.2.2. Retrospective Self-Report Questionnaires

Baseline Anxiety. Youth participants completed the Screen for Child Anxiety Related Emotional Disorders (SCARED). The SCARED is a 41-item questionnaire with total scores ranging from 0 to 82. Participants rate how often statements are true of them on a three-point scale from 0 ("Not True or Hardly Ever True") to 2 ("Very True or Often True"). The SCARED has five subscales, each screening for a different anxiety symptom category: social anxiety (e.g., "It is hard for me to talk with people I don't know well"), generalized anxiety ("I worry about things working out for me"), somatic anxiety ("When I get frightened, I sweat a lot"), separation anxiety ("I don't like to be away from my family), and school avoidance ("I am scared to go to school"). A minimum score of 25 is associated with clinically significant anxiety. The SCARED has good test-retest reliability (r = .62, N = 1092 children aged 7-18 in the United States; Behrens et al., 2019) and high internal consistency (α = .43-.89, N = 1559 children aged 8-16 in

China; Su et al., 2008). It has also been validated for remission prediction in youth with anxiety disorders in the United States (Caporino et al., 2017).

Adult participants completed the Screen for Adult Anxiety Related Emotional Disorders (SCAARED). The SCAARED is a 44-item, self-report questionnaire with factors for generalized anxiety, social anxiety, separation anxiety, and somatic anxiety symptoms. Like the youth version (SCARED), participants rate how often statements were true of them in the past three months on a three-point scale (0 = "Not True or Hardly Ever True"; 2 = "Very True or Often True"), with a maximum score of 88. A minimum score of 23 is associated with clinically significant anxiety. For data analysis purposes, SCAARED scores were multiplied by 41 and divided by 44 to correct for the difference in the number of items compared to the SCARED. The SCAARED has good internal consistency (α = .86-.97, N = 336 participants aged 18-27 in the United States; Angulo et al., 2017).

Baseline Depression. Youth participants completed the Centre for Epidemiological Studies Depression Scale for Children (CES-DC). The CES-DC is a 20-item questionnaire, with scores ranging from 0 to 60. Participants rate how accurately statements describe them during the past week on a four-point scale, from 0 ("A lot") to 3 ("Not at all"). Scores greater than or equal to 15 are good evidence for clinically relevant depression. Example items include "I was bothered by things that usually don't bother me", "I was happy", and "I felt like crying" (Garber & Weersing, 2010). The CES-DC has been validated for depression in adolescents in the United States (Faulstich, 1986), and has shown good internal consistency (α = .86), inter-rater reliability (ICC = 0.82), and test-retest reliability (r = .85; N = 367 children aged 10-17 in Rwanda; Betancourt et al., 2012).

Adult participants completed the Centre for Epidemiological Studies Depression Scale (CES-D), which has been validated in adults in the United States (Radloff, 1977) and China (Chin et al., 2015). Like the CES-DC, the CES-D has 20 items to which respondents rate how accurately statements describe them during the past week on a four-point scale, from 0 ("Rarely or none of the time (less than 1 day)" to 3 ("Most or all of the time (5-7 days)"; Radloff, 1977). The CES-D has good test-retest reliability (ICC = .91) and internal consistency (ωH = .434-.738; N = 3686 adults, mean age 49.4 years; Chin et al., 2015).

Baseline Substance Use. Youth participants completed the Brief Screening Instrument for Adolescent Tobacco, Alcohol, and Drug Use (BSTAD). The BSTAD is a questionnaire for identifying and quantifying substance use in adolescents. There are at least seven questions, with a maximum of 39 questions if participants indicate "Yes" to having used any of the substances listed in the questionnaire. Essentially, participants indicate what substances they have used (e.g., marijuana, heroin, alcohol) and how frequently they have used the indicated substance in the past 30 days, 90 days, and year. The BSTAD has been validated for detecting and quantifying substance use with high sensitivity (0.95) and specificity (0.97; N = 525 youth aged 12-17 in the United States; Kelly et al., 2014).

Adult participants completed the Alcohol, Smoking, and Substance Involvement Screening Test (ASSIST). The ASSIST consists of eight questions, each with 10 sub-items identifying particular substances. Like the BSTAD does for youth, the ASSIST assesses the type and frequency of substance use in adult populations (Humeniuk et al., 2008). The ASSIST has good test-retest reliability (κ = .58-.90; N = 236 adults, mean age 34 years; tested in Australia, Brazil, India, Thailand, the United States, the United Kingdom, and Zimbabwe; WHO ASSIST Working Group, 2002).

2.2.3. Ecological Momentary Assessments

Participants were prompted on the PROSIT app to complete ecological momentary assessments of anxiety, sociability, physical activity, and sleep. Anxiety was rated on a reverse Likert-item, from 1-7 ("Very anxious" to "Not anxious at all"). The anxiety item was reversed to ensure that participants would not default only to an extreme answer as a form of response bias (Baumgartner & Steenkamp, 2001). Sociability items included one yes/no response ("Did you have friends visit your house in the past week?") and a rating item ("How often did you go out with friends in the past week?") with six answers ranging from "Never" to "More than four times". Items assessing physical activity were borrowed from the International Physical Activity Questionnaire, asking participants how many days and hours/minutes per day of the last week they did "vigorous" (e.g., "heavy lifting") and "moderate" (e.g., "bicycling at a regular pace") activities, and how often they walked for 10 minutes or more (Craig et al., 2003). Items assessing sleep were borrowed from the Sleep Habits Survey, assessing bedtime, wake time, and how frequently participants awoke each night (Wolfson et al., 2003).

2.2.4. The PROSIT App

App Data Collection. Participants were required to download and use the PROSIT app for 14 days, including answering the daily anxiety ratings as well as the weekly ecological momentary assessments and leaving the app running in the background to passively (i.e., without participants doing anything) collect mobile sensing data. The PROSIT app collects data from numerous sensors and uploads the data to a server at Dalhousie University through highly secure encrypted secure sockets layer (SSL) connections. Though the PROSIT app collects data from numerous sensors, we were interested in this study in five behavioural categories: mobility, sociability, screen usage, sleep, and physical activity. Therefore, for the present study our analyses focused on GPS, ambient light, screen logs (i.e., phone lock and unlock states), call logs, and activity logs. From each of these sensors, we engineered features measuring various characteristics of the anxiety-related behaviour categories, which we describe in the next section.

Choosing the PROSIT App. There are a wide variety of mobile sensing apps and analytical frameworks available which could be utilized in mental health research. Mendes et al. (2022) reviewed 31 mobile sensing mental health apps. The four most frequently monitored behavioural feature sets amongst these apps are mobility, sociability, physical activity, and sleep. Three well-established apps, for example, are the AWARE framework (Nishiyama et al., 2020), the Effortless Assessment of Research System (EARS; Lind et al., 2018, 2023), and the mindLAMP project (Vaidyam et al., 2022; see Table 1 for a comparison of these platforms to the PROSIT app). Each of these apps has sensors and available features from the five behaviours we aimed to track, and they are all available for iOS and Android devices. Alongside alternative mobile sensing apps, there are available analysis pipelines, such as the Reproducible Analysis Pipeline for Data Streams (RAPIDS) project, which has mobile sensing features for iOS and Android smartphones, as well as devices such as Fitbit and Empatica (Vega et al., 2021). The AWARE framework, for instance, uses the RAPIDS project (Nishiyama et al., 2020).

We chose to use the PROSIT app specifically for several reasons. First, the PROSIT app was codesigned by youth (Orr et al., 2023), which makes it uniquely strong to use in a sample of youth with anxiety problems. Second, it was a central aim of the present study to attempt to replicate and extend the methods employed by MacLeod et al. (2021), who related PROSIT app features to internalizing symptoms (especially anxiety) in youth during the COVID-19 pandemic. Third, though there are other apps available for mobile sensing, we chose to use the PROSIT app with features engineered based on standards set in other studies, with other apps and analytical approaches. For example, many of the features (especially those generated from GPS) are based on features found to be useful in clinical samples of adults by the Remote Assessment of Disease and Relapse–Major Depressive Disorder research program (Saeb et al., 2015, 2017; Zhang et al., 2022). It is a strength, therefore, that our study comparing youth and adults compromises between the youth and adult mobile sensing literature by using an app codesigned by youth, with features shown robustly to correlate with anxiety and depression symptoms in adults.

PROSIT App Security. The PROSIT app is highly secure. Participants are assigned an anonymized username to login to the app. The most sensitive information the app collects are longitude and latitude coordinates, which are encrypted and were only decrypted temporarily during the calculation of mobility features. Data from the PROSIT app are uploaded first to a public domain server, then transferred to a non-public domain, secure server protected by a firewall and stored securely at Dalhousie University. Participants were extensively informed of the PROSIT app's security features during consent, including that data will be kept securely for a total of seven years, after which it is destroyed. In short, the PROSIT app follows a data isolation protocol for transferring data from the app, to a public domain, and then immediately to secure storage in the private server, as recommended in available literature from experts in mobile sensing security (Jagesar et al., 2021).

2.3. Data Preprocessing

All data preprocessing and feature engineering was performed using the *R* programming language, version 3.6.3 (R Core Team, 2021). At the beginning of data preprocessing, participants were kept only if they provided at least 70% (10 days) of the requested 14 days of data collection. As some participants forgot to delete the app for a couple of days after being notified of study completion, only data which was collected during their 10-14 days of anxiety rating responses was kept for analysis. All features – including anxiety ratings – were summarized into weekly averages for a total of two timepoints (week one and week two).

2.3.1. Mobility Feature Engineering

Stay Points. GPS data were logged and timestamped as longitude and latitude coordinate pairs. These logs were then clustered into unique locations, or “stay points”, which were locations where participants spent at least ten minutes in each week. Following MacLeod et al.'s (2021) work with the PROSIT app, coordinate pairs within a 20-meter haversine radius of the stay points were clustered and labelled as belonging to that stay point – a process which was repeated until a total number of weekly stay points was derived for every participant. Larger numbers of stay points imply greater mobility, at least in terms of unique locations visited.

Location Variance. Following Saeb et al. (2017), we derived location variance as the combined variance of latitude and longitude coordinate logs. This was calculated using the following formula: $\log(\sigma_{lat}^2 + \sigma_{long}^2)$, where σ^2 denotes the variance. Larger values imply larger variance in the coordinate pairings of participants’ GPS data, and therefore higher mobility.

Total Haversine Distance. As is common in the mobile sensing literature, we calculated the total distance participants travelled each week using the haversine formula, which quantifies the distance between two points on a sphere (Sardar et al., 2022). Larger values therefore imply greater mobility, purely in terms of distance rather than the number of unique locations or variance in time spent at visited locations.

Location Entropy. We extracted two entropy features: raw and normalized location entropy. Raw location entropy is calculated with the formula $-\Sigma_{\mathrm{i}=1}^{N} p_i \log(p_i)$, where p_i is the proportion of time a participant spent at a given stay point in each week. Normalized location entropy is similar; however it is adjusted to be invariant to the number of stay points, according to the formula $-\frac{1}{\log(N)}\Sigma_{\mathrm{i}=1}^{N} p_i \log(p_i)$. Therefore, lower values of these variables imply greater mobility, in the sense that participants spread their time diversely across stay points. Normalized entropy restricts the range of possible values between 0-1, so that a value of 1 indicates that participants spent exactly the same proportion of time at every single stay point in a given week (MacLeod et al., 2021; Saeb et al., 2017).

2.3.2. Sociability Feature Engineering

Total Call Duration. As iOS and Android vary slightly in the labels they use for call logs, only connected calls were considered for sociability features, where a connected call could have been answered either as an incoming call or connected as an outgoing call. Total call duration was calculated as the sum of weekly incoming and outgoing call durations, with higher values therefore implying greater time spent on phone calls in each week.

Total Incoming and Outgoing Calls. For each week, the total number of incoming calls and total number of outgoing calls each participant received were calculated. Therefore, higher values of either variable imply more phone calls of the respective type.

2.3.3. Screen Usage, Physical Activity, and Sleep Feature Engineering

Screen usage, physical activity, and sleep features overlap strongly in terms of measurement properties. Evening and nightly screen time and unlocks, for example, are associated with insomnia and sleep deprivation in adolescents (Foerster et al., 2019). Likely, people who have high screen time in the evening and night are more sedentary. Therefore, overall screen usage variables and overall physical activity variables were calculated, as well as screen use and physical activity variables restricted to the evening hours and nighttime hours to proxy sleep.

Total Screen Time. Screen usage was extracted in the form of timestamped phone lock and unlock events. Therefore, total screen time was taken as the sum of time between successive unlock and lock states. To proxy sleep disruption, we also calculated evening (5pm-11pm) and nightly (11pm-7am) screen time. Higher values therefore suggest greater time spent with the screen in an unlocked state, presumably using the smartphone in some way (MacLeod et al., 2021).

Total Screen Unlocks. Total screen unlocks were taken as the sum of unlock events in each week. As for screen time, we also calculated evening and nightly screen unlocks to proxy sleep disruption. Higher values therefore suggest a greater frequency of smartphone checking behaviours – specifically, checking behaviours where participants unlocked their device (MacLeod et al., 2021).

Proportion of Time Spent Sedentary. iOS and Android have labels for physical activity, which are derived from their respective operating systems' classification algorithms based on accelerometry. As the categories differ slightly (e.g., Android has a category called "tilting" which iOS does not have), we only investigated common feature labels provided by both devices: "stationary", "in a vehicle", "cycling", "walking", and "running". From this, we calculated the proportion of time participants spent sedentary (i.e., stationary) for a given week, as well as in the evening and nightly hours in case those variables correlate with evening and nightly screen usage features. Higher values therefore imply more sedentary time.

Physical Activity Entropy. Recall that for mobility, location entropy quantifies variance in locations visited and/or the time spent at those locations. For physical activity, entropy was calculated using the proportion of time participants spent engaged in each activity category common to iOS and Android. Higher values therefore indicate little variance in the kind of physical activities participants engage in.

Ambient Light. Ambient light intensity was logged continuously by the PROSIT app. Following MacLeod et al. (2021), we calculated ambient light intensity between the hours of 11pm-7am, with higher values presumably indicating greater exposure to light (and therefore, a lower probability a participant is asleep) in the nighttime.

2.4. Analysis Plan

First, we planned Spearman rank correlations of each mobile sensing feature by two-week summaries of daily anxiety scores (i.e., momentary anxiety), and baseline anxiety scores for both youth and adult participants. As some participants were missing data for specific features, Spearman correlations were calculated using complete cases respective to each feature. We also planned six linear modles with baseline and momentary anxiety as the outcomes, which we tested across all participants, youth participants only, and adult participants only. We planned to compare base models only including demographic covariates to models including mobile sensing features, using F-statistics from an analysis of variance and ΔR^2 statistics to check for changes in model fit. The Spearman and linear model analyses were planned to address our first research aim (to replicate and extend the PROSIT app findings of MacLeod et al., 2021), as well as to address our first research question and associated hypotheses (that mobile sensing will predict anxiety, while improving model fit on top of covariates). The Spearman analyses also assist in addressing our second research question and hypothesis (that there will be differences in how mobile sensing may associate to anxiety in youth and in adults).

Second, we planned factor analyses and structural equation models, to investigate our second research aim about differences in youth and adult anxiety as measured by mobile sensing. For factor analyses, we planned to fit latent factors for mobility, sociability, sleep, screen usage, and physical activity with their respective features. Part of these analyses involve assessing measurement invariance of these constructs between youth and adults, as well as across time (from the first to the second week). We also planned structural equation models, to investigate whether there are differences in the paths relating latent behavioural factors to momentary and baseline anxiety. We were interested in whether such effects exist at the purely between-subjects level, at the within-subjects level, and/or longitudinally (i.e., between the first and second week of measurement; limited to momentary anxiety).

Third, we addressed our final research question investigating what mobile sensing features are best for anxiety prediction. Specifically, we were interested in two questions: first, which set of features best predict momentary anxiety and momentary anxiety variance across populations? And second, which set of features best classify participants as being a youth or an adult? Following Bai et al. (2021) who found random forest models were most useful in classifying changes in depression. Random forest models are extensions of decision trees, calculating classification decision criteria and producing a ranking of feature importance based on the predictive value each feature has in performing the classification.

CHAPTER 3. RESULTS

3.1. Sample Descriptives

3.1.1. Demographic Covariates

A total of 65 youth and 96 adults provided usable (10-14 days) mobile sensing data for our analyses in the timeframe allotted to this thesis. See Table 2 for sample descriptive characteristics. Several things are notable for present analytic purposes: regarding ethnicity, the three most common ethnicities were Asian (27.7% of youth; 9.4% of adults), Black/African (24.6% and 18.8%), and White (46.2% and 70.8%). Given that only 6 youth and 6 adults reported an ethnicity not within these three categories, for analysis purposes, the ethnicity variable was recoded as an ordered three-factor with "Non-(Asian or Black/African)" as the reference category. Notably, the three most common countries of origin for participants were Canada (N = 22 youth and 48 adults), the United States (8 youth; 20 adults), and Kenya (12 youth; 19 adults), and one youth participated in Nigeria. Given that the majority of participants did not report a mental health diagnosis (73.85% of youth and 66.67% of adults), and that the reported diagnoses were almost entirely anxiety or depression diagnoses, the diagnostic status variable was recoded as an ordered two-factor variable with "No Diagnoses" as the reference category. Due to research ethics considerations, socioeconomic status was measured via annual income in adults, and proxied via parental education in youth. Since the median household income in Canada is $98,390 (*Canada: Median Annual Family Income by Province | Statista*, n.d.) and adult participants reported nearly equally often an income below $100k or above $100k, the income variable was recoded into a two-factor variable. Furthermore, since the majority of youth reported having a parent who finished university (78.5% of youth participants), the parental education and income variables were combined. Therefore, a value of "1" on the socioeconomic status proxy variable indicated an income greater than $100k for adults, and a parent with a non-university education for youth. Remaining covariates were recoded to reflect strong sample biases in each of them, including sexual orientation (two levels; “Straight” reference category) and location (two levels; “Urban” reference category). Due to very little gender diversity in the sample, gender was not included as a covariate in our analyses.

3.1.2. Baseline and Momentary Anxiety

Youth participants had an average baseline anxiety (SCARED) score of 47.48 (*SD* = 10.91). Since the youth and adult versions of the baseline anxiety questionnaire have different numbers of items (41 versus 44), we adjusted for the difference in the number of questionnaire items by multiplying the adult participants’ scores by 44, then dividing by 41. After adjusting for the difference in the number of questionnaire items, adult participants had an average baseline anxiety (SCAARED) score of 47.82 (*SD* = 14.89). Given the high baseline anxiety scores, we calculated subscale scores to see which subscale cutoffs participants in each group met. For youth participants, 60 (92.71%) met the cutoff for likely panic disorder, 52 (80%) for generalized anxiety, 54 (83.08%) for separation anxiety, 47 (72.31%) for social anxiety, and 45 (69.23%) for school avoidance. For adult participants, 92 (95.83%) met the cutoff for panic, 86 (89.58%) for generalized anxiety, 78 (81.25%) for separation anxiety, and 76 (79.17%) for social anxiety.

For each participant, momentary (i.e., daily) anxiety ratings were summarized into two timepoints – one for each week of assessment. The momentary anxiety item was reversed to ensure participants did not default to one of two extreme answers (1 or 7). Therefore, for participants, 7 meant “Not anxious at all” and 1 meant “Very anxious”. For analytical purposes, we removed the reverse scoring prior to calculating models. So, whereas a score of 7 meant “Not anxious at all” for participants, the present results can be interpreted as 7 indicating the highest value of momentary anxiety ("Very anxious"). We also calculated momentary anxiety variance, which was defined as each participant's variance in momentary anxiety ratings across a given week.

3.1.3. Baseline Depression and Substance Use

To roughly compare baseline substance use across groups, we extracted questions from the BSTAD and ASSIST questionnaires assessing which substances participants had used in the past 90 days, if any. Unsurprisingly, the most common substance used in both groups in the past 90 days was alcohol (36.9% of youth versus 64.6% of adults). Tobacco was the next most frequent but was used considerably more frequently in adults (32.3%) than youth (4.6%). Youth did not report any use in the past 90 days of cannabis (versus 35.4% for adults), cocaine (2.1% of adults), amphetamines (4.2% of adults), hallucinogens (2.1% of adults), nor opioids (2.1% of adults). Therefore, to make the models most comparable, only past 90-day alcohol use was included as a covariate.

Youth participants had an average baseline depression score of 32.1 (*SD* = 10.6) compared to slightly lower average of 26.7 (*SD* = 9.98) for adults. However, comparable proportions of youth scored in the clinically significant depressive symptom range (89.23%; minimum score of 15 on the CES-DC) as adults (89.58%; minimum score of 16 on the CES-D). Therefore, both groups are well within the range of severe depressive symptoms.

3.1.4. Ecological Momentary Assessment Descriptives

Each participant's weekly ecological momentary assessments of sociability, physical activity, and sleep behaviours were summarized into total (i.e., across two weeks) descriptives. Youth reported spending more time with friends than adults did: 66.15% of youth had a friend visit their house in the past two weeks, and 43.08% had a friend over at least once per week. For adults, these statistics were 47.92% and 18.75%, respectively. A total of 22.92% of adults reported never having had a friend over nor having spent time with a friend in the past two weeks, compared to just 3.08% of youth.

Adults reported engaging in more vigorous and moderate physical activity than youth: youth engaged in vigorous physical activities an average of 1.6 days per week, during which they spent an average of 64.85 minutes engaged in a vigorous activity. Adults reported engaging in vigorous physical activities an average of 1.76 days per week, for 70.72 minutes. Differences were similar for moderate activities: youth spent 48.6 minutes on 1.87 days per week in moderate activity compared to 71.67 minutes on 1.94 days per week for adults. Both groups spent approximately equal minutes walking per week (50.75 minutes for youth vs 51.35 minutes for adults) and walked for at least 10 minutes on similar numbers of days (4.68 and 4.15 days for youth and adults, respectively).

Youth self-reported sleeping an average of 10.04 hours per night (*SD* = 4.85), with a median of 8.5 hours. Adults self-reported sleeping an average of 9.37 hours per night (*SD* = 4.18), with a median of 8.5 hours. In case this somewhat high average sleep time was related to clinically significant depression, we checked whether relatively high sleep time (> eight hours) participants were more likely to have clinically significant depression. However, this seems unlikely, as youth with clinically significant depression slept an average of 9.9 hours per night whereas youth with non-clinically significant depression scores slept an average of 11.55 hours per night. For adults, these numbers were 9.35 hours and 9.54 hours, respectively. Furthermore, 83.72% of youth who slept more than eight hours per night had clinically significant depression, compared to 94.59% of youth who slept less than or equal to eight hours per night. For adults, these numbers were 88.14% and 90.74% of their respective sub-groups. Youth may have awoken less frequently during the night than adults: youth awoke a median of 1 time per night, or 1.33 times on average ($SD = 0.89$); adults awoke a median of 2 times per night, or 2.12 times on average ($SD = 1.69$).

3.1.5. Mobile Sensing Feature Descriptives

A summary of descriptive statistics for each mobile sensing feature can be found in Table 3. To see if the mobile sensing indicators for sociability, sleep, and physical activity were consistent with the ecological momentary assessment findings, we ran Spearman correlations between a select few features and the ecological momentary assessments. Specifically, we ran Spearman correlations beteen all the sociability features and sociability assessments, as well as all the features and assessments for physical activity and sleep. For physical activity and sleep, after correcting for multiple tests with the false discovery rate correction method, none of the Spearman correlations between any feature and an ecological momentary assessment were related. For sociability, youth who reported having had a friend visit their house at least once in the past two weeks had significantly higher total call durations, $r = 0.28$, $p < .05$, as well as more incoming calls, $r = 0.33$, $p < .001$, and more outgoing calls, $r = 0.38$, $p < .001$. The ecological momentary assessments therefore only seem to have related to mobile sensing indicators of sociability in youth, and perhaps not at all to physical activity and sleep in either sample.

3.2. Spearman Correlations of Momentary and Baseline Anxiety

To assess whether our momentary anxiety item was tapping a similar anxiety construct as the baseline anxiety questionnaires, we correlated baseline anxiety scores to mean momentary anxiety ratings. Momentary anxiety weakly (and non-significantly) positively associated with baseline anxiety in youth (*r* = 0.12), and negatively associated with baseline anxiety in adults (*r* = -0.10). Given this finding, we performed followup correlations of momentary anxiety by each item on the youth and adult baseline anxiety questionnaires (SCARED and SCAARED), in case some items of the respective questionnaires are more related to momentary experiences of anxiety. Statistically significant correlations of momentary anxiety with items on the SCARED and SCAARED are shown in Table 4. At face value, the momentary anxiety measure appears to tap more relevant items on the youth SCARED rather than the adult SCAARED. For example, an item on the youth SCARED such as "I get shaky" makes sense to positively correlate with momentary anxiety, *r* = 0.24, *p* = .006. However, the same item on the adult SCAARED does not seem to correlate with momentary anxiety in adults, *r* = -0.04, *p* = .59.

3.3. Spearman Correlations of Mobile Sensing Features and Internalizing Symptoms

To assess whether mobile sensing features predict baseline and momentary anxiety (RQ1, H1) and whether there are differences in predictions between youth and adults (RQ2, H3), we ran Spearman correlations between mobile sensing features, baseline anxiety scores, and momentary anxiety scores in both samples. For exploratory purposes, we also ran Spearman correlations between baseline depression scores and mobile sensing features. A sensitivity analysis in G*Power showed that with 80% power and 95% confidence, we would need Spearman correlations between mobile sensing features and internalizing symptoms of 0.33 in the youth sample, and 0.28 in the adult sample. Even with the two samples combined, the minimum correlation would be 0.22 for significance. Given that MacLeod et al.'s (2021) largest Spearman correlation was 0.293, it is unlikely *a priori* that we will have sufficient power to achieve statistical significance in a Spearman analysis. Therefore, we calculated Spearman correlations between mobile sensing features and internalizing symptoms with 95% bootstrapped confidence intervals, using the bias-corrected and accelerated (BCa) bootstrap interval.

3.3.1. Mobility Features

As shown in the top and bottom left panels of Figure 1, the correlations of mobility features collected from the PROSIT app with momentary and baseline anxiety show similar trends to MacLeod et al.'s (2021) Spearman findings with baseline anxiety scores as the outcome: in youth, momentary and baseline anxiety are positively associated with haversine distance travelled ($r = 0.15$; $r = 0.15$), location entropy ($r = 0.18$; $r = 0.31$), and the number of unique stay points ($r = 0.21$; $r = 0.30$). For adults, momentary anxiety positively associates with location entropy ($r = 0.19$) and normalized entropy ($r = 0.20$), whereas baseline anxiety positively associates with distance travelled ($r = 0.17$) and possibly location variance ($r = 0.18$). In short, when expressed in terms of the number of locations visited (youth) or the distance travelled (youth and adults), greater mobility predicts greater anxiety. However, participants of both age groups who spread their time less diversely between locations visited also tend to have higher anxiety.

3.3.2. Sociability Features

As shown in Figure 2, the correlations between internalizing symptoms and sociability features tend towards null (as in MacLeod et al., 2021), perhaps with greater outgoing calls weakly negatively associating with baseline anxiety in adults ($r = -0.12$). In terms of variance in momentary anxiety, however, youth with less variance may have longer phone calls ($r = -0.16$), as well as more incoming ($r = -0.13$) and outgoing ($r = -0.23$) phone calls. The opposite appears true for adults, who have higher momentary anxiety variance with longer phone calls (*r* = 0.20), as well as more incoming (*r* = 0.18) and outgoing (*r* = 0.27) phone calls.

3.3.3. Screen Usage Features

As shown in Figure 3, the screen usage feature findings of MacLeod et al. (2021) have been extended: screen time positively associates with momentary anxiety (*r* = 0.23), whereas phone unlocks positively associate with baseline anxiety in youth (*r* = 22) and perhaps more weakly in adults (*r* = 0.12).

3.3.4. Sleep Features

As shown in Figure 4, MacLeod et al.'s (2021) finding that baseline anxiety in youth positively associates with nighttime ambient light exposure intensity is replicated (*r* = 0.38). The opposite was true for adults (*r* = -0.23), however the adults also had a much lower mean ambient light intensity than youth (1745 versus 7002, respectively), and ambient light intensity correlated with ecological momentary assessment self-reported sleep time in youth (*r* = -0.23), but not adults (*r* = 0.03). Evening screen time and phone unlocks also trended towards positive associations with baseline anxiety in youth (*r* = 0.19; *r* = 0.18), as did nighttime phone unlocks (*r* = 0.17). Notably, the percent of time youth spent stationary at night was negatively associated with baseline anxiety (*r* = -0.24), despite that the percent of time spent stationary of night was positively associated with self-reported sleep time (*r* = 0.32), and positively associated with the number of self-reported sleep disruptions (*r* = 0.25).

3.3.5. Physical Activity Features

As shown in Figure 5, there is a trend consistent with MacLeod et al. (2021) that youth with higher momentary anxiety spend more time stationary (*r* = 0.14). The trends for entropy show that youth who vary more in the physical activities they engage in have higher momentary anxiety (r = -0.10 for activity entropy, and r = -0.19 for normalized activity entropy).

3.4. Tests of Model Fit: Covariate-Base Models Compared to Mobile Sensing Models

To assess whether mobile sensing features improve the fit of models predicting baseline and momentary anxiety symptoms (RQ1, H2), we tested differences of model fit between linear models containing only demographic covariates, and models with mobile sensing features added. We therefore hierarchically tested six models, with the outcomes being baseline and momentary anxiety for models including all participants, youth participants only, and adult participants only. Demographic covariates included biological sex, ethnicity, mental health diagnostic status, socioeconomic status, sexual orientation, location, and phone device type.

For baseline anxiety, the only model that was improved to statistical significance by mobile sensing features was the model including all participants. However, ΔR^2 was in the expected direction for model improvement for all three models, in descending order from the youth model, ΔR^2 = 0.30, F(18, 43) = 1.81, p = .056; the adult model, ΔR^2 = 0.19, F(17, 127) = 1.64, p = .06, and the all-participants model, ΔR^2 = 0.13, F(18, 197) = 1.72, p = .04. The effects of demographic covariates in each model are shown in Table 5.

For momentary anxiety, none of the model fit improvements with the addition of mobile sensing features were statistically significant. However, all ΔR^2 values were similar in magnitude as those for baseline anxiety, and they were all in the direction indicating model improvement with the addition of mobile sensing features. In descending order, the results are as follows: youth participants, ΔR^2 = 0.22, F(18, 43) = 0.88, p = .60, adult participants, ΔR^2 = 0.16,

$F(18, 126) = 1.40$, $p = .14$, and all participants included, $\Delta R^2 = 0.10$, $F(18, 197) = 1.36$, $p = .15$. The effects of demographics covariates in each model are shown in Table 6.

3.5. Factor Analyses of PROSIT Mobile Sensing App Features

As explained in our introduction, the features we derived from the PROSIT app were derived based on theories and findings in past mobile sensing, anxiety, and mental health-related behavioural literature. We were interested in whether a five-factor solution uniting the features could be derived, with factors for mobility, sociability, physical activity, screen usage, and sleep. Therefore, we performed factor analyses on each of the feature sets, with an exploratory and a confirmatory step. For exploratory factor analyses, we fit oblique/oblimin rotation exploratory factor models. For confirmatory factor analyses, we fit a confirmatory model using full-information maximum likelihood estimation for missing data, and robust maximum likelihood estimation for parameter estimation. All input features were standardized prior to factor analyses and subsequent structural equation modelling.

Exploratory factor analyses were conducted in the 'psych' package for *R* (Revelle, 2024). Confirmatory factor analyses and structural equation models were performed in 'lavaan' (Rosseel, 2012).

3.5.1. Exploratory Factor Analyses: A Possible Four-Factor Solution

Exploratory factor analyses revealed possible factors for mobility, sociability, physical activity, and screen time features. For mobility, the features with a loading greater in magnitude than 0.40 were the number of unique stay points (loading = 0.55), location entropy (loading =

1.04)[1], and normalized entropy (loading = 0.67). For sociability, the features with a loading greater than 0.40 were total call duration (loading = 0.83), total incoming calls (loading = 0.87), and total outgoing calls (loading = 0.93). For physical activity, these features included activity entropy (loading = -0.89), normalized activity entropy (loading = -0.83), percent of time spent stationary (loading = 0.95). Two features also originally extracted for sleep loaded instead onto physical activity, including percent of time spent stationary in the evening (loading = 0.74) and percent of time spent stationary at night (loading = 0.77). Furthermore, the remaining sleep features were derived from screen usage features, which collapsed into a single factor containing the following features: total screen time (loading = 0.72), total unlocks (loading = 0.85), total evening screen time (loading = 0.51), total evening unlocks (loading = 0.72), total nightly screen time (loading = 0.47), and total nightly unlocks (loading = 0.62). We therefore moved onto confirmatory factor analyses in search of the optimal solution, beginning with a proposed four-factor structure including mobility, sociability, physical activity, and screen usage.

3.5.2. Confirmatory Factor Analyses

We started by fitting a four-factor model for mobility, sociability, physical activity, and screen usage. The initial model fit was bad (*CFI* = 0.81; RMSEA = 0.16). Therefore, we followed modification indices suggested by ‘lavaan’, focusing on adjusting the model in a stepwise fashion as per the largest modification index on a given step. All the modification indices from the initial model and successive steps focused on constraining error variances of

[1] The reader will notice that the factor loading for location entropy indicates an ultra-Heywood case, as the error variance of the loaded variable is left to be negative. One way to test if this actually invalidates the factor solution or if it is due to sampling error is through confidence intervals in a structural model (Kolenikov & Bollen, 2012): if the confidence interval for the error variance include positive values, then sampling error cannot be eliminated as a candidate explanation. We concluded this did not invalidate the mobility factor, as the confidence intervals of the error variance of both groups included non-negative values (youth: [0.095, 0.723]; adults: [-0.037, 0.930]).

physical activity variables, and then of screen usage variables. Unfortunately, these modification indices led the model to become unidentified. Therefore, we were forced to experiment with simplifying the factor structures for physical activity and screen usage.

In the end, two features had to be removed from the physical activity factor: percent of time spent stationary in the evening, and percent of time spent stationary at night. The screen usage factor also had to be removed, as leaving it in the model despite the adjustments to physical activity left the model without convergence. The final baseline confirmatory model had acceptable fit (*CFI* = 0.97; *RMSEA* = 0.09; *SRMR* = 0.08), with a three-factor solution including: mobility (number of stay points, location entropy, and normalized location entropy; error variances of the latter two items constrained), sociability (total call duration, total incoming calls, and total outgoing calls), and physical activity (activity entropy, normalized activity entropy, and percent of time spent stationary). We attempted to add in ecological momentary assessments of sociability and physical activity to the model. However, the model failed to converge, so they were left out of the factor definitions.

3.6. Measurement Invariance Testing

3.6.1. Between-Group, Non-Longitudinal Invariance Testing

With a baseline structural model in place (see Figure 6), we moved on to test for measurement invariance. Establishing measurement invariance is necessary for group comparisons involving latent factor models, for reasons which are apparent from the process of invariance testing. First, configural invariance must be established, such that a factor structure is defined which holds with good fit across all groups. If the model fits only for a single group, then the proposed factor structure cannot be applied consistently in both groups, and group comparisons therein are meaningless. Second, metric invariance must be established by constraining factor loadings across groups to equality. Metric invariance is different from configural invariance, then, in that the same overlying factor structure may be already established, but the structure within factors (i.e., extent to which the factor is defined by each input) may vary across groups. If the model fit remains satisfactory upon constraining factor loadings, then the factors can be soundly compared to the extent that the internal factor definitions are the same. Third, scalar invariance must be established by constraining item intercepts across groups to equality. If groups differ in the intercepts of their observed items, group differences in path analysis may be attributable simply to measurement problems rather than actual group differences.

It is important to note, then, that a structural model which does not satisfy complete measurement invariance can still be used to make group comparisons: for example, we could observe that the number of stay points is more relevant to the mobility factor for youth with anxiety than adults with anxiety. However, we would be limited in our capacity to infer whether the effect of mobility on, say, momentary anxiety as an outcome, is due to measurement inconsistencies or actual group differences.

With these considerations in mind, we proceeded to test for measurement invariance of the defined mobility, sociability, and activity factors derived from the PROSIT mobile sensing app. Criteria for determining measurement invariance are somewhat controversial, as the extent to which a given criterion is overly liberal or overly conservative depends heavily on sample size: for example, a significant difference in model fit on a chi-square test has been criticized as being overly conservative (particularly in small samples), whereas a reduction of < 0.01 in the *CFI* may be overly liberal (especially in large samples with many group comparisons; Putnick &

Bornstein, 2016). As our sample size is relatively small for a structural equation model, we employed several combinations of the following rejection criteria: $\Delta CFI < -.01$, $\Delta RMSEA < .015$, $\Delta SRMR < .030$ (for metric invariance, and 0.015 for scalar invariance), and a non-significant chi-squared difference test. If a model has a significant chi-squared difference test and one of the three fit index criteria are unmet, or a model has two or more unmet fit index criteria, then we conclude the model is noninvariant at a given test step.

The results of our measurement invariance testing are shown in Table 7. The proposed model passed metric invariance, but did not achieve scalar invariance. Following Sharifian et al. (2021), we released intercepts sequentially according to the largest modification index suggested by 'lavaan'. We discovered that youth and adults differ in their intercepts for two features: total call duration and activity entropy. Therefore, regardless of overall sociability and physical activity, youth make shorter phone calls (standardized intercept = -0.227) than adults (standardized intercept = 0.087), and youth vary the kinds of physical activities they participate in (standardized intercept = -0.019) more than adults do (standardized intercept = 0.018). We went on to test for any differences in the means of latent factors, by constraining the latent means to be equal across groups (after accounting for partial scalar noninvariance).

Given the partial scalar noninvariance of our model, we further followed Sharifian et al. (2021) in calculating regression outcomes with the proposed three-factor solution: all outcome models were calculated with partial invariance constraints, and under the assumption of full measurement invariance. If the pattern of results differed across the partially invariant and the assumed full invariance model, then we decided not to interpret them beyond measurement noninvariance. Otherwise, we planned to interpret them if the findings were the same across models.

3.6.2. Between-Group, Longitudinal Measurement Invariance Testing

Prior to fitting any cross-lagged panel models, we first assessed for measurement invariance across weeks (week one versus week one) with our baseline structural model, following guidelines proposed by Mackinnon et al. (2020). Similar to group comparisons, testing for longitudinal effects requires measurement invariance so that findings cannot be attributed simply to changes in the stability and/or meaning of measurements.

Our configural model left factor loadings and intercepts unconstrained, with a single modification index specifying covariance between location entropy and normalized location entropy to achieve adequate model fit. The initial configural model fit was adequate (*CFI* = 0.95, *RMSEA* = 0.08, *SRMR* = 0.09). As shown in Table 8, our model achieved full longitudinal measurement invariance, with acceptable changes in fit indices upon constraining factor loadings and intercepts across weeks.

3.7. Structural Equation Models of Anxiety and Depression Symptoms

With partial measurement invariance established, we moved on to investigate differences in how mobile sensing features may predict anxiety (and depression, for exploratory purposes) in youth and adults. Therefore, we attempted to construct eight structural equation models: four purely between-subjects, non-longitudinal models (momentary anxiety, momentary anxiety variance, baseline anxiety, and baseline depression); two repeated measures, cross-lagged panel models (momentary anxiety and momentary anxiety variance), and two multilevel models (momentary anxiety and momentary anxiety variance).

3.7.1. Between-Subjects Structural Equation Models

The final models included mobility, sociability, and physical activity latent factors as predictors. For momentary anxiety and momentary anxiety variance, baseline anxiety and baseline depression were included as predictors. For baseline anxiety, baseline depression was included as a predictor, and vice versa for the baseline depression model. Unfortunately, the addition of any other predictors either made model fit completely unacceptable beyond the utility of modification indices or prevented the model from converging altogether.

Both partial and full measurement invariance models for all outcomes had overall acceptable or adequate fit (*CFI* = 0.93-0.96; *RMSEA* = 0.08-0.10; *SRMR* = 0.08-0.09). The pattern of findings also did not differ between the partial and full measurement invariance models, so findings discussed here will reference numbers from the full invariance models (see Tables 7-10 for partially invariant model statistics). Due to low sample size, reports here will also focus on trends demonstrated by 95% confidence intervals, specifically those that do not cross zero and which might be useful in comparing trends across groups by constraining regression coefficients to equality.

As shown in Table 9, baseline depression was significantly positively associated with momentary anxiety in adults (b* = 0.22 [0.03, 0.40]), but not youth (b* = 0.02 [-0.26, 0.30]). This difference was not significant, however, when the regression coefficients were constrained to equality across groups, $\chi^2(df_{diff} = 4)$ = 8.43, p = .08.

Table 10 shows that for adults, mobility was significantly negatively associated with momentary anxiety variance (b* = -0.20 [-0.38, -0.02]), whereas the trend was positive for youth (b* = 0.18 [-0.08, 0.44]), and this difference was statistically significant, $\chi^2(df_{diff} = 4)$ =

10.93, p = .03. Sociability was significantly negatively associated with momentary anxiety variance in youth (b* = -0.18 [-0.33, -0.04]), whereas the trend was positive for adults (b* = 0.11 [-0.10, 0.31]), which was also a statistically significant difference, $\chi^2(df_{diff} = 4)$ = 12.41, p = .01. These trends in the structural models are consistent with the trends shown in Figures 1 (mobility features) and 2 (sociability features) of bootstrapped Spearman correlations. Baseline anxiety was significantly negatively associated with momentary anxiety variance only in youth (b* =-0.19 [-0.33, -0.04] versus b* = -0.03 [-0.23, 0.16]) – a difference which was presumably significant given that constraining the coefficients prevented the model from converging. The model also would not converge when the coefficients for baseline depression were constrained to equality, as baseline depression was significantly positively associated with momentary anxiety variance only in youth (b* = 0.23 [0.05, 0.41] versus b* = 0.09 [-0.13, 0.31]).

For the models of baseline anxiety and depression, baseline depression significantly positively predicted baseline anxiety in both groups (b* = 0.47 [0.33, 0.61] and b* = 0.43 [0.30, 0.57]), and baseline anxiety significantly positively predicted baseline depression in both groups (b* = 0.44 [0.32, 0.57] and b* = 0.47 [0.32, 0.61]).

3.7.2. Cross-Lagged Panel Models

To assess whether latent factors for mobility, sociability, and physical activity can predict momentary anxiety longitudinally (i.e., from week-to-week), we fit cross-lagged panel models. The fit of the baseline model (i.e., not including outcomes for momentary anxiety or momentary anxiety variance) was poor (*CFI* = 0.86, *RMSEA* = 0.14, *SRMR* = 0.22) and was immune from improvement despite following modification indices. We therefore will not discuss these models further, as our sample size is likely too small to attain sufficient fit for any interpretation to be coherent or meaningful.

3.7.3. Multilevel Models

We attempted to fit multilevel models to parse out within- and between-subjects effects of the latent factors on momentary anxiety and momentary anxiety variance. However, 51 unique participants had insufficient variance in their mobile sensing indicators and/or outcomes (momentary anxiety and momentary anxiety variance). As such, there was insufficient within-subjects variation for the structural equation model to converge, even upon attempting to force convergence via the expectation maximization algorithm.

3.7.4. Simple(r) Cross-Lagged and Multilevel Models: An Analytical Compromise

The cross-lagged panel models failed due to poor model fit, likely because our sample size is too small to specify such a complicated model with three latent factors. The multilevel models failed due to insufficient within-subjects variance in various predictors and outcomes, leading to too many excluded cases to estimate multilevel effects with three latent factors. We looked to the available literature for guidance, and found inspiration in Zhang et al. (2022), who used simpler methods to estimate cross-lagged and multilevel effects in a mobile sensing study of people with depression. We therefore fit simple cross-lagged models for every mobile sensing feature, estimating the lagged and cross-lagged effects of each feature and outcome (momentary anxiety and momentary anxiety variance) across age groups. We then fit simple multilevel models of each feature, with two limitations: first, we had to combine both samples to get sufficient within-subjects variation to estimate a multilevel model. Second, despite these efforts, there remained some features for which models of either type would not converge. We therefore will focus on the features for which both model types converged, noting this limitation of measurement for other available features.

The cross-lagged model results are shown in Figures 7-10, and the multilevel model results are shown in Figures 11 and 12. Generally, trends for lagged effects were positive, showing that mobile sensing features in the first week are related to mobile sensing features in the second week, to varying degrees of statistical significance in our study (see the right columns of Figures 7 and 9). Momentary anxiety (and perhaps to a lesser extent, momentary anxiety variance) also tended to produce lagged effects across weeks (see the right columns of Figures 8 and 10).

While other features produced effects in some models (e.g., total screen time in both age groups had a significant lagged effect upon itself across weeks), it is most notable that out of all the features, only four of the five mobility features allowed for convergence in all model types: location variance, the number of unique stay points, location entropy, and normalized location entropy. Their patterns of results may then be more robust, as they clearly produced sufficient within-subjects and across week variance to achieve convergence in each model type. As shown in Figure 7, there appears to be an age-related effect of location variance, such that location variance in the first week of measurement predicts higher momentary anxiety in the second week for youth (b^* = 0.85 [0.48, 1.22] versus b^* = 0.12 [-0.18, 0.42] for adults). Aside from this finding, remaining cross-lagged results appear either null or limited to features which did not produce convergence across all model types.

3.8. Random Forest Models of Mobile Sensing Feature Importances

For our final set of analyses, we performed random forest models to see if we could predict momentary anxiety, momentary anxiety variance, baseline anxiety, and baseline depression from mobile sensing features. Random forest regression methods make use of decision trees, whereby subsamples of the data are used to make multiple decision trees corresponding to multiple predictions. At the end, a prediction is made based on the mean or the mode of the predictions offered by all decision trees. Random forest classifiers work in a similar fashion, offering a classification based on the most frequent class label produced by the decision trees. We chose random forest models for two reasons: first, past mobile sensing research has found them to be broadly the best for classification accuracy (e.g., Bai et al., 2021). Second, random forest models produce feature importance values, which quantify and rank the contribution of each input feature to a given prediction. Therefore, random forest models are suitable for addressing our third research question regarding the set of mobile sensing features which best predict internalizing symptoms in youth and in adults.

All random forest modelling was conducted in the Python programming language, version 3.11.4 (Van Rossum & Drake, 2009), using the 'Scikit-learn' package (Pedregosa et al., 2011).

3.8.1. Random Forest Regressions of Momentary Anxiety and Momentary Anxiety Variance

When the entire sample was used to construct a model of momentary anxiety, model fit was improved beyond a baseline mean model substantially by the addition of mobile sensing features (R^2 = 0.12, *MSE* = 1.04 versus baseline *MSE* = 1.23; see the feature importances in Figure 13). Fit was similar for a model of momentary anxiety variance features (R^2 = 0.14, *MSE* = 2.97 versus baseline *MSE* = 3.45; see Figure 14). The primary difference between the two models was that for momentary anxiety variance, the most important features were clearly dominated by physical activity features. When the sample was split into youth and adults, however, model fit was substantially poorer for momentary anxiety (for youth, R^2 = -0.14, *MSE* = 1.14 versus baseline *MSE* = 1.09; for adults, R^2 = 0.07, *MSE* = 1.32 versus baseline *MSE* = 1.46; see Figures 15 and 16). Models of momentary anxiety variance were substantially improved when splitting the participants across age groups (youth: R^2 = 0.20, *MSE* = 1.78 versus baseline *MSE* = 2.54; adults: R^2 = 0.43, *MSE* = 2.92 versus baseline *MSE* = 5.19; see Figures 17 and 18).

3.8.2. Random Forest Recursive Feature Elimination

The earlier random forest models included the entire set of mobile sensing features and ecological momentary assessments in all predictions. Given it is logically possible that some features are useful for predicting momentary anxiety and momentary anxiety variance, and that others may be counterproductive, we employed the random forest recursive feature elimination algorithm (RF-RFE) with ten-fold cross-validation. The RF-RFE algorithm begins with the entire set of features, and then progressively removes features until a pre-specified model fit index (in our case, R^2) stops improving. Like classic random forest models, the RF-RFE algorithm can be used for continuous prediction, as well as for classification. We therefore refined our third research question into three extended questions: (1) What minimal set of features is best at predicting momentary anxiety and momentary anxiety variance? (2) What minimal set of features is best at classifying a participant as a youth or as an adult, given information only about their behaviours (i.e., mobile sensing features only; no inputs for momentary anxiety)? (3) What minimal set of features is best at classifying a participant as a youth or as an adult, given information about their behaviours and self-reported anxiety symptoms (i.e., mobile sensing features, momentary anxiety, and momentary anxiety variance)? For classification, we balanced the youth and adult training sets using random oversampling.

The model we constructed for the first question indicated that predictions of momentary anxiety and momentary anxiety variance did not substantially improve by limiting the feature set via the RF-RFE algorithm. Fit was best when the entire sample was included (R^2 = 0.12; see

Figure 19 for feature importances). When only youth were included, fit was still worse than a baseline mean model (R^2 = -0.15; see Figure 20 for feature importances). When only adults were included, fit was still worse than the overall sample model, but was better than the baseline mean model (R^2 = 0.08; see Figure 21). For momentary anxiety variance, model fit was better when splitting into youth (R^2 = 20; see Figure 22) and adults (R^2 = 0.35; see Figure 23).

For the second and third questions, we constructed four more RF-RFE models, this time setting them up to classify participants as youth or adults. The best performing model included mobile sensing features and momentary anxiety variance (accuracy = 80.52%; youth recall = 90%, adult recall = 74%; see Figure 24 for feature importances). The next best models included mobile sensing features only (accuracy = 79.22%; youth recall = 77%, adult recall = 83%; see Figure 25), followed by a model including mobile sensing features, momentary anxiety scores, and momentary anxiety variances (accuracy = 77.92%; youth recall = 87%, adult recall = 72%). Classification accuracy was lowest (61.04%) when only momentary anxiety scores and momentary anxiety variances were included (youth recall = 73%, adult recall = 53%).

CHAPTER 4. DISCUSSION

4.1. Summary of Research Findings

We were interested in extending the PROSIT mobile sensing app to measure anxiety and anxiety-related behaviours in both youth and adults, specifically those with clinically significant anxiety, and in a post-pandemic sample.

A previous study with the PROSIT app did not screen for clinically significant anxiety in youth, and was conducted completely during the COVID-19 pandemic (MacLeod et al., 2021). Furthermore, there are very few studies comparing youth and adult populations with mobile sensing, and none to our knowledge which select participants based on clinically significant anxiety. We were also interested in measuring anxiety both at baseline and using ecological momentary assessments. Past studies using the PROSIT app only assessed anxiety at baseline, thereby missing out on potentially useful fluctuations of momentary anxiety (MacLeod et al., 2021). In these regards, this thesis successfully addressed several methodological limitations in past mobile sensing literature.

Beyond methodology, we had several research aims and questions. We aimed to address these with a diversity of analytical methods, ranging from simple correlation analyses to structural equation models and machine learning models. We will review our findings according to each research question.

4.1.1. Does Mobile Sensing Predict Anxiety in Youth and in Adults?

To address our first research question and hypotheses, we conducted Spearman regressions of each extracted mobile sensing feature, to see if the features predicted baseline anxiety, momentary anxiety, and momentary anxiety variance, with predictions based on week-long summaries of two weeks of data collection (H1). We also fit linear models to see if mobile sensing features would improve the fit of models of anxiety symptoms, over and above the contributions of demographic covariates (H2).

While the specific results of our Spearman correlations will be discussed later, we did find a variety of associations between mobile sensing features and internalizing symptoms, which were essentially consistent with the findings of MacLeod et al. (2021). Further consistent with that earlier study, we showed that mobile sensing features increased the proportion of variance explained in models of anxiety and depression symptoms, in addition to the contributions of demographic covariates. Though not all models achieved statistically significantly better fit, their improvements were all comparable to models that achieved statistical significance. Therefore, we conclude that we have successfully replicated and extended MacLeod et al. (2021), showing that mobile sensing features collected via the PROSIT app predict anxiety in a post-pandemic sample of youth and adults with clinically significant anxiety problems.

4.1.2. Are There Differences in How Youth and Adults Express Anxiety Through Mobile Sensing?

We hypothesized that there would be differences in how youth and adults' behaviours related to anxiety, both at the bivarate level with Spearman correlations, and in structural equation models (H3).

In this study, we measured five behavioural categories which have been related to anxiety: mobility, sociability, physical activity, sleep, and screen usage. We used exploratory factor analyses to see if features loaded on their respective hypothesized categories, which many of them did, indicating a potential four-factor solution. We were successful in fitting a three-factor solution modelling mobility, sociability, and physical activity behaviours in youth and in adults. Three out of five mobility features loaded onto a common factor, and so did all three sociability features, all three physical activity features, and both screen usage features. The sleep features, however, loaded onto physical activity and screen usage factors in an exploratory analysis, likely because the sleep features were estimated from common measurements (e.g., total evening screen time). In a confirmatory factor analysis, we were able to derive a three-factor solution including features for mobility, sociability, and physical activity. We could not include measures or factors of screen usage and sleep in our structural models without sacrificing model fit or model convergence, which we attribute to our low sample size.

We saved specific comparisons for discussion of clinical implications later. For now, it will suffice to say that we found age-related differences in how mobile sensing features relate to baseline anxiety, momentary anxiety, and momentary anxiety variance, in both our bivarate Spearman correlations and our structural equation models. Only four out of five mobility features produced sufficient within-subjects and longitudinal variance so as to make simple cross-lagged and multilevel models converge. This made us unable to fit cross-lagged panel models or multilevel models using our entire three-factor structure. Thus, our small sample size and little across-week variance made anything beyond between-subjects comparisons across each age group impossible. It is interesting to note that past studies have found mobility features to be the most robust for predicting internalizing symptoms, both in terms of multilevel and longitudinal or cross-lagged effects, especially in two week time frames such as ours (Bai et al., 2021; Moshe et al., 2021; Saeb et al., 2015, 2017; Zhang et al., 2022).

4.1.3. What Are the Most Useful Features for Predicting Anxiety in Youth and in Adults?

We were interested in using machine learning analyses to identify differences in feature importances for predicting anxiety in youth and adults, for several reasons. Some of these reasons are related to limitations of structural equation models, and others to inherent strengths of random forest machine learning models.

Though structural equation models are easier to intrepret given their simple latent factor solutions, they are extremely sensitive to sample size. We observed this limitation in a variety of barely non-significant effects, model fit issues, and convergence issues. Structural equation models also have strong distributional assumptions pertaining to input data. We used robust maximum likelihood estimators to avoid potential distributional violations biasing our results, but even robust methods do not eliminate the constraints of such distributional assumptions. So, though regression coefficients in structural models can be used for quantifying differences in predictor relevance across groups, models with less strict distributional assumptions may be better suited for this task. Finally, structural equation models privilege explanatory parsimony by reducing the number of predictors to a set of common factors. These common factors are not necessarily based on the combination of features which explain the most variance in a predicted outcome. Random forest models address these limitations well: as they do not involve the estimation of latent factors, they may be less sensitive to issues with our small sample size. They also provide a list of feature importances, which rank the predictive utility of all the input features. Using the recursive feature elimination algorithm, we also fit random forest models which started from the entire set of mobile sensing features and eliminated features until the fit of the models stopped improving.

Generally, mobile sensing features were more useful for predicting momentary anxiety variance than momentary anxiety scores in youth and adults. Model fit was best for the youth model of momentary anxiety variance, next best for the adult model of momentary anxiety variance, next best for the adult model of momentary anxiety scores, and worst by far for the youth model of momentary anxiety scores. Though the feature importance plots generally show age-related differences that are consistent with differences shown in Spearman regressions (e.g., screen time is more related to momentary anxiety in youth than in adults), those are not the most important results here. Rather, it is more useful to note that there is variation in the degree to which each feature explains variance anxiety, and these results do not necessarily follow importances according to the features which were implemented in the structural equation models. When we tried to identify a minimally important feature set using the recursive feature elimination algorithm, the only difference between the groups was that a smaller list of features optimally predicted momentary anxiety variance in adults than in youth. Otherwise, all models included some number of features assessing mobility, sociability, physical activity, screen usage, and sleep, and restricting the features to a minimally important set did not substantially change the order of importances.

Finally, we fit four classification models to classify participants as youth or adults based on their mobile sensing features alone, momentary anxiety scores and momentary anxiety variances alone, features with momentary anxiety scores, and features with momentary anxiety variances. To evaluate model performance, we focused on two metrics: accuracy and recall. Recall is the proportion of cases of a given class which were correctly classified out of the total number of observations with that class label. There are therefore two recall statistics per model:

the percent of youth correctly classified as youth, and the percent of adults correctly classified as adults.

Generally, model accuracies were comparable, except for the model including only momentary anxiety scores and momentary anxiety variances which had the lowest accuracy by a nearly 20% margin. The best performing model included mobile sensing features and momentary anxiety variances. Most interestingly, however, was that the addition of momentary anxiety variances in any model tended to substantially increase recall for both classes, especially youth. So, while momentary anxiety scores may not be particularly useful for distinguishing youth and adults with anxiety problems, momentary anxiety variances may be useful, especially when paired with mobile sensing features.

4.2. Clinical Implications

From our comparisons of youth and adults, there emerge a variety of associations between behaviours and anxiety symptoms which clinicians may find useful to attend to or consider. We will therefore spend some time reviewing these associations for each set of mobile sensing features and place them in the broader context of the clinical literature.

Increased mobility associated with worsened anxiety in both youth and adults in the Spearman correlations, as well as with higher variance in youth anxiety and lower variance in adult anxiety in the structural equation models. While the effects of mobility were not significantly associated with worsened anxiety in a structural equation model, the trend in youth was positive and with an effect size similar to that of significant effects for other structural equation findings. Decreased mobility has been associated with worsened depression, but findings vary for anxiety (Boukhechba et al., 2017; Zhang et al., 2022). It may be that for anxiety, changes in mobility could represent fluctuations of avoidance behaviours, but the measurement of such effects would probably be more useful had we obtained information not just quantifying mobility, but qualifying the types of locations people were visiting. Past research has shown that worsened anxiety is associated with less mobility, specifically when decreases in mobility are explained by a less visits to friends and religious sites (Boukhechba et al., 2018; Huang et al., 2016). Our findings show that associations of increased mobility and worsened anxiety found with the PROSIT app are not unique to samples collected during the COVID-19 pandemic (MacLeod et al., 2021). In sum, it is hard to make a strong case for specific clinical implications regarding mobility and anxiety, so we recommend future researchers investigate these associations further, with a focus on quantifying how mobile participants are alongside qualifying the types of locations they are visiting.

Physical activity has repeatedly been found to be protective against anxiety in mobile sensing studies (Bufano et al., 2023; Mack et al., 2021), and physical activity is sometimes useful in the treatment of internalizing symptoms (Wanjau et al., 2023). We did not find significant associations of physical activity and anxiety in our structural equation models, perhaps due to our small sample size. However, in our Spearman correlations, there were trends such that youth who spent more time stationary expressed more momentary anxiety. Interestingly, youth who varied the kinds of physical activities they engaged in more expressed more momentary anxiety. We are unaware of other studies which have included features for activity entropy based on classifications provided by smartphone devices, so it may be worth studying in the future whether there is a useful distinction to be made between the quantity of physical activity people engaged in, versus the quantity of unique activities they engage in. The set of features we used to assess physical activity was small in comparison to other mobile sensing studies, particularly those which have used mobile sensing methods alongside smartphone sensors, such as accelerometery (Bufano et al., 2023). Therefore, future studies may benefit from including a wider array of physical activity features, particularly for constructing structural equation models. Our results also suggest that physical activity is a separate, independent factor from mobility, validating a critique we mentioned earlier that past studies in mobile sensing and anxiety have inappropriately conflated these two constructs (Renn et al., 2018). Nonetheless, there are intuitive reasons for supposing these two constructs may interact in clinically relevant respects: for example, a person who is highly mobile because they are highly engaged in a specific physical activity (e.g., running) probably differs from a person who is highly mobile because they are constantly moving to avoid certain locations. The latter person may or may not be highly physically active. Such potential test cases highlight the importance of assessing these two constructs independently.

As for screen usage, our findings show that in youth, greater screen usage (both in terms of phone unlocks and screen time) predict worsened anxiety. There are a variety of candidate mechanisms for explaining this association. For example, the traditional model of problematic behaviours states that people may engage in problematic behaviours (such as substance use) to cope (e.g., to distract themselves from anxiety or depression), to conform (e.g., to cave into peer pressure), for social reasons (e.g., to meet more people), or for enhancement (e.g., for the inherent excitement of engaging in the behaviour). For general screen usage, it may be that youth who use screens to cope (i.e., to distract themselves from anxiety or depression), or to conform (i.e., due to fear of missing out) experience worsened internalizing symptoms, just as youth who use alcohol or other problematic substances as driven coping or conformity motives (Brailovskaia & Margraf, 2024; Cooper et al., 1995; Elmquist & McLaughlin, 2018). We also found trends such that evening and nighttime-specific screen usage may predict worsened anxiety and/or more variance in momentary anxiety in youth. These results are consistent with past studies, which have found screen usage in youth to predict poor sleep (Foerster et al., 2019), which may explain the association with increased anxiety (Alshoaibi et al., 2023; Santiago et al., 2022). We also found that in youth, exposure to higher amounts of ambient light at night predicted higher baseline anxiety, whereas the trend was opposite for adults. Youth also had a much higher average ambient light exposure at night than adults, indicating that there may be a threshold for ambient light exposure which is disruptive to sleep. The associations of nightly screen usage and anxiety also lead us to think that youths' nighttime ambient light exposure may be influenced by their nighttime screen usage behaviours. Still, youth and adults self-reported sleep times within a seemingly healthy range. So, these findings may be spurious, or it could be that self-reported sleep times do not capture disruptions to sleep throughout the night, and rather just reflect time spent in bed. Future studies should therefore extend the nighttime screen usage and ambient light features with validated measures of sleep and insomnia in youth.

While sociability features did not associate with anxiety scores in either group, youth who were higher in sociability (in terms of total call duration and the number of phone calls) experienced less variance in momentary anxiety whereas adults experienced more variance in momentary anxiety. It would be a stretch to draw any clinical conclusions from this, especially given that findings for associations of both anxiety and depression with sociability features are mixed (Cao et al., 2020; Gillett et al., 2021; Messner et al., 2019; Stamatis et al., 2024). Further investigation is necessary, perhaps in studying age-related differences in how people experiencing mental health problems may use phone calls, for example, to seek social support, or as a form of smartphone addiction. Future studies could also benefit from investigating features associated with text messages, which may reveal semantic patterns (e.g., excessive personal pronoun use) that could predict internalizing symptoms (Funkhouser et al., 2023).

4.3. Limitations

Aside from limitations already discussed, there are some other methodological limitations of the present study.

First, our sample size was insufficient not just for deriving complicated structural equation and cross-lagged panel models, but also for performing rigorous moderation analyses with the inclusion of demographic factors. We were aiming for 130 participants in each group, but could achieve only 65 youth and 96 adults in the timeframe allotted to this as a Master's thesis project. Future studies with larger sample sizes could take a variety of analytical directions, such as explorations of age effects within the youth group. For example, in Canada (the country of origin for most of the study participants), youth can obtain their license at 16 years old, and legally drink at 19 years old. It would be reasonable to expect, then, that emerging adults are more mobile than younger adolescents, as the former may travel more for work, school, and social opportunities. Emerging adults probably also are more likely to engage in substance use. The youth sample in this study was strongly biased towards older adults, with only five youth participants below 18 years old, and the youngest participant being 16. Aiming for a wider age range as well as a more representative spread of ages within the defined youth age range could be helpful in future studies.

Two other demographic factors we were unable to assess the importance of are residential location and ethnicity. The majority of participants (72.3% of youth and 59.4% of adults) were urban dwellers, and only a few (3.1% of youth; 13.5% of adults) were rural dwellers. This could have affected data quality, particularly for the GPS sensors which have been shown to be noisier and less accurate in cities due to the presence of buildings blocking the GPS signal (DXOMARK, 2023). Furthermore, some differences in the mental health of urban and rural youth have been magnified since the COVID-19 pandemic. For example, anxiety and externalizing behaviours became more prevalent particularly amongst urban youth, whereas depression remained stably more prevalent in rural youth compared to urban youth (Figas et al., 2022). As for ethnicity, the samples had strong representation of Asian (27.7% of youth; 9.4% of adults) and Black/African participants (24.6% of youth; 18.8% of adults). Furthermore, many of the Black/African participants were residents of African countries, with the majority being from Kenya (18.46% of youth; 19.79% of adults). Notably, the baseline questionnaires we used have not been validated in Kenyans, and we were unable to test for measurement invariance nor ethnic differences in the questionnaires and mobile sensing data. Youth and adults in sub-Saharan Africa are extremely underserved for their mental health (Sequeira et al., 2022). Other researchers have speculated that sub-Saharan African adolescents may be more strongly affected in some areas of performance (e.g., school; overall health) given higher rates of poverty, disease, violence, and other health risks in their regions of residence (Mabrouk et al., 2022). Given the strong interest from Kenyan participants in this study, remote assessment tools employed through mobile sensing may offer fruitful ways to explore and address the unique needs of sub-Saharan African youth and adults struggling with unmet mental health challenges. Future research, however, is necessary to establish the validity of such tools in sub-Saharan African populations.

Second, we only made use of retrospective self-report questionnaires at baseline. Had we included a second time point for these questionnaires (particularly the anxiety questionnaires), we could have tested relations between momentary anxiety and longitudinal changes in anxiety on a validated clinical questionnaire. We also could have assessed whether mobile sensing features relate more strongly to longitudinal changes in clinical self-report questionnaires, or momentary changes in anxiety. Further extending the timeframe participants use the app in (e.g., using it for four weeks instead of only two) could only make these comparisons more powerful.

Third, there are a variety of analyses we were unable to conduct in the scope of the present Master's thesis. We only used simple random forest regressors and classifiers, which were useful for examining feature importances as we sought to but are not particularly useful for investigating longitudinal nor demographic effects of specific mobile sensing features. We could have used more state-of-the-art machine learning models – such as the time series transformer model – to analyze our data in more depth, including at finer time periods than from week-to-week. We chose to make comparisons across each week of the two-week study period because we wanted to maximize power in a structural equation model, but we do not have to stick with this time frame for further analyses. Since completing the analyses presented here for the purpose of thesis defence, we have been working on other machine learning techniques, such as gradient boosting. We are also analyzing different classification problems to explore research questions we were underpowered to assess with the structural equation analyses. For example, we are attempting to classify participants with and without an anxiety and/or depressive disorder diagnosis. We are also investigating whether mobility features (which, as mentioned earlier, may predict worsened anxiety and improved depression) can differentially classify participants with an anxiety diagnosis versus those with a depression diagnosis.

Fourth, our sample was recruited entirely online, rather than through a clinical setting. We initially planned to recruit a sample of youth through a telehealth clinical service but were unable to recruit enough participants through that route in time to include in the present thesis. That study is, however, ongoing independently of this thesis project, as are further analyses of the data presented here. Another limitation of online recruitment is an inability to completely control or analyze which region participants are located in. While the diversity of our sample is a strength, there may be geographical differences in the expression of anxiety which we were unable to capture in moderation analyses due to small sample size. Future researchers should therefore note that greater sample diversity is attainable via online advertising of mobile sensing studies, opening new avenues for generalizing this technology across the world.

Fifth and finally, there are two respects in which the results of this study are difficult to use for clinical interpretations. One is that we did not include any interventions in this thesis. So, while we may observe associations of anxiety and behaviour that clinicians could find notable, we cannot make any strong recommendations based on our study results which could immediately influence clinical practice. Another is that our momentary anxiety item may, in hindsight, have not been the best measure we could have used as an ecological momentary assessment. A brand new preprint shows that Likert scale items (like the one we used) and visual analogue scales (e.g., rating anxiety from one to 100 on a slider) are equally good at measuring within-subjects variance, but visual analogue scales correlate considerably better with clinical measures (Haslbeck et al., 2024). Visual analogue scales, therefore, may be better equipped for comparisons of associations between momentary anxiety and baseline anxiety insofar as they associate with objectively measured behaviours. There are also a large variety of other mobile sensing features which we were unable to engineer from the data collected, and which may have been useful for clinical interpretation. For example, sociability features predicted greater momentary anxiety variance in adults, but less momentary anxiety variance in youth. Without knowing who participants were calling and for what reason(s), we cannot productively speculate about clinical implications of these effects. Perhaps adults were more likely to call people related to their workplace, whereas youth were more likely to call friends or family. Text message features may have also been useful to add to sociability. Past research has shown, for example, that adolescents report more negative mood corresponding with more "self-focused" text messages (e.g., increased use of first-person pronouns; McNeilly et al., 2023). Future clinical research in mobile sensing would be better equipped with a broader range of mobile sensing features addressing specific clinical symptoms and research questions.

4.4. Future Directions

Retrospective self-report questionnaires have a variety of measurement properties which make them useful clinically. They produce little within-subjects variance and large between-subjects variance, making longitudinal changes in the same subject relevant for detecting clinically significant changes in symptomatology (Hedge et al., 2018). They also are useful for factor analyses, which are useful for deriving coherent, explanatory psychological theories of psychopathology which simplify clinical classifications and group comparisons. The SCARED, for example, is validated for detecting anxiety disorders, discriminating a variety of anxiety disorders, and detecting changes of anxiety symptom severity in adolescents (Birmaher et al., 1997). Retrospective self-report questionnaires do not, however, capture within-subjects, real-time fluctuations in behaviour and symptomatology, which may be useful for making clinical predictions, such as in detecting longitudinal changes in symptom severity (Jacobson et al., 2021). Furthermore, they are subject to recall and other self-report biases (Komarahadi et al., 2004; Sato & Kawahara, 2011)

In this study, we demonstrated that mobile sensing features (particularly, four out of five GPS features) produced sufficient within-subjects, real-time variance to make multilevel and longitudinal assessments of anxiety-related behaviour possible. Future studies may find that splitting data into different intervals than across weeks may be more appropriate, and/or that measuring participants for a longer period of time may be useful. Furthermore, this study shows that mobile sensing features do, in fact, relate to anxiety symptoms both at baseline and as assessed via ecological momentary assessment. The features measured here also clustered into three-factor model which was derived in part by reference to clinical literature hypothesizing a set of anxiety-related behaviours. All these facts converge upon a promising future for mobile sensing research, where objectively measured behaviours could be validated for clinical predictions and decision making in ways that could supplement the retrospective self-report tools already available to clinicians.

Another promising part about this study for future research is related to our use of both the classical statistical, structural equation and new machine learning approaches. As with any clinical measurement, the kinds of factors and items that contribute to those factors may vary across populations in which researchers aim to measure the same underlying construct (hence the requirement for measurement invariance testing). The SCARED, for instance, has a considerably different factor structure from the SCAARED, because youth and adults express anxiety differently. Despite our small sample which limited our capacity to make complicated structural equation models, we were still able to perform meaningful group comparisons of youth and adults with anxiety problems. With machine learning analyses, we showed that the kinds of features which best relate to anxiety in youth and adults may differ, and the extent to which any given feature is important for predicting anxiety across the populations may differ as well.

Furthermore, the features which provide the best prediction of anxiety, and the best age-group classification are not necessarily the features which form a factor analytical model which coheres with preconceived theories of anxiety-related behaviour. Both model types likely have their place: factor analyses could help in deriving parsimonious, comprehensible theories of anxiety-related behaviours, while machine learning models could analyze complex relations between features that might not be captured in a factor analysis or structural equation model. Both approaches have been used with success in past studies (e.g., Bai et al., 2021; Zhang et al., 2022), so we hope this study will encourage researchers to pursue directions using and potentially uniting both analytical approaches.

4.5. Conclusion

This thesis contributes to the mobile sensing and anxiety literature in several important respects. To our knowledge, we are either the first or among the first authors who have directly compared youth and adults with clinically significant anxiety problems using the same smartphone mobile sensing app. We also are the first to replicate and extend the findings of MacLeod et al. (2021) with the PROSIT app essentially in their entirety, in a new sample which was collected after the COVID-19 pandemic. We showed that in youth and adults, mobility, sociability, physical activity, screen usage, and sleep features from the PROSIT mobile sensing app can be used to measure their intended behaviours and relate them to anxiety in ways that are consistent with both the available clinical literature and past mobile sensing studies. We showed that mobile sensing features form a coherent, latent factor structure in youth and adults. We also showed that the features which best predict anxiety in youth and adults may differ. Finally, we showed that mobile sensing features can be used to classify youth and adults in machine learning models. In sum, this thesis shows that the PROSIT mobile sensing app can measure anxiety symptoms and related behaviours which may be valuable for clinical observation and group comparisons. The results are promising for the future development and fine-tuning of mobile sensing technologies to various populations experiencing anxiety problems.

REFERENCES

Alamoudi, D., Breeze, E., Crawley, E., & Nabney, I. (2023). The Feasibility of Using Smartphone Sensors to Track Insomnia, Depression, and Anxiety in Adults and Young Adults: Narrative Review. *JMIR mHealth and uHealth*, *11*(1), e44123. https://doi.org/10.2196/44123

Aledavood, T., Torous, J., Triana Hoyos, A. M., Naslund, J. A., Onnela, J.-P., & Keshavan, M. (2019). Smartphone-Based Tracking of Sleep in Depression, Anxiety, and Psychotic Disorders. *Current Psychiatry Reports*, *21*(7), 49. https://doi.org/10.1007/s11920-019-1043-y

Alshoaibi, Y., Bafil, W., & Rahim, M. (2023). The effect of screen use on sleep quality among adolescents in Riyadh, Saudi Arabia. *Journal of Family Medicine and Primary Care*, *12*(7), 1379–1388. https://doi.org/10.4103/jfmpc.jfmpc_159_23

Altamura, A. C., Dell'osso, B., D'Urso, N., Russo, M., Fumagalli, S., & Mundo, E. (2008). Duration of untreated illness as a predictor of treatment response and clinical course in generalized anxiety disorder. *CNS Spectrums*, *13*(5), 415–422. https://doi.org/10.1017/s1092852900016588

Angulo, M., Rooks, B. T., Gill, M., Goldstein, T., Sakolsky, D., Goldstein, B., Monk, K., Hickey, M. B., Diler, R. S., Hafeman, D., Merranko, J., Axelson, D., & Birmaher, B. (2017). Psychometrics of the screen for adult anxiety related disorders (SCAARED)- A new scale for the assessment of DSM-5 anxiety disorders. *Psychiatry Research*, *253*, 84–90. https://doi.org/10.1016/j.psychres.2017.02.034

Asselbergs, J., Ruwaard, J., Ejdys, M., Schrader, N., Sijbrandij, M., & Riper, H. (2016). Mobile Phone-Based Unobtrusive Ecological Momentary Assessment of Day-to-Day Mood: An Explorative Study. *Journal of Medical Internet Research*, *18*(3), e72. https://doi.org/10.2196/jmir.5505

Awick, E. A., Ehlers, D. K., Aguiñaga, S., Daugherty, A. M., Kramer, A. F., & McAuley, E. (2017). Effects of a randomized exercise trial on physical activity, psychological distress and quality of life in older adults. *General Hospital Psychiatry*, *49*, 44–50. https://doi.org/10.1016/j.genhosppsych.2017.06.005

Azizabadi, Z., Aminisani, N., & Emamian, M. H. (2022). Socioeconomic inequality in depression and anxiety and its determinants in Iranian older adults. *BMC Psychiatry*, *22*(1), 761. https://doi.org/10.1186/s12888-022-04433-w

Bai, R., Xiao, L., Guo, Y., Zhu, X., Li, N., Wang, Y., Chen, Q., Feng, L., Wang, Y., Yu, X., Xie, H., & Wang, G. (2021). Tracking and Monitoring Mood Stability of Patients With Major Depressive Disorder by Machine Learning Models Using Passive Digital Data: Prospective Naturalistic Multicenter Study. *JMIR mHealth and uHealth*, *9*(3), e24365. https://doi.org/10.2196/24365

Bandelow, B., & Michaelis, S. (2015). Epidemiology of anxiety disorders in the 21st century. *Dialogues in Clinical Neuroscience*, *17*(3), 327–335.

Barry, T. J., Yeung, S. P., & Lau, J. Y. F. (2018). Meta-analysis of the influence of age on symptom change following cognitive-behavioural treatment for anxiety disorders. *Journal of Adolescence*, *68*, 232–241. https://doi.org/10.1016/j.adolescence.2018.08.008

Baumgartner, H., & Steenkamp, J.-B. E. M. (2001). Response Styles in Marketing Research: A Cross-National Investigation. *Journal of Marketing Research*, *38*(2), 143–156. https://doi.org/10.1509/jmkr.38.2.143.18840

Becker, C. B., Zayfert, C., & Anderson, E. (2004). A survey of psychologists' attitudes towards and utilization of exposure therapy for PTSD. *Behaviour Research and Therapy*, *42*(3), 277–292. https://doi.org/10.1016/S0005-7967(03)00138-4

Behrens, B., Swetlitz, C., Pine, D. S., & Pagliaccio, D. (2019). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Informant discrepancy, measurement invariance, and test-retest reliability. *Child Psychiatry and Human Development*, *50*(3), 473–482. https://doi.org/10.1007/s10578-018-0854-0

Beidel, D. C., Turner, S. M., & Morris, T. L. (1999). Psychopathology of childhood social phobia. *Journal of the American Academy of Child and Adolescent Psychiatry*, *38*(6), 643–650. https://doi.org/10.1097/00004583-199906000-00010

Betancourt, T., Scorza, P., Meyers-Ohki, S., Mushashi, C., Kayiteshonga, Y., Binagwaho, A., Stulac, S., & Beardslee, W. R. (2012). Validating the Center for Epidemiological Studies Depression Scale for Children in Rwanda. *Journal of the American Academy of Child and Adolescent Psychiatry*, *51*(12), 1284–1292. https://doi.org/10.1016/j.jaac.2012.09.003

Birmaher, B., Khetarpal, S., Brent, D., Cully, M., Balach, L., Kaufman, J., & Neer, S. M. (1997). The Screen for Child Anxiety Related Emotional Disorders (SCARED): Scale Construction and Psychometric Characteristics. *Journal of the American Academy of Child & Adolescent Psychiatry*, *36*(4), 545–553. https://doi.org/10.1097/00004583-199704000-00018

Boukhechba, M., Chow, P., Fua, K., Teachman, B. A., & Barnes, L. E. (2018). Predicting Social Anxiety From Global Positioning System Traces of College Students: Feasibility Study. *JMIR Mental Health*, *5*(3), e10101. https://doi.org/10.2196/10101

Boukhechba, M., Huang, Y., Chow, P., Fua, K., Teachman, B. A., & Barnes, L. E. (2017). Monitoring social anxiety from mobility and communication patterns. *Proceedings of the 2017 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2017 ACM International Symposium on Wearable Computers*, 749–753. https://doi.org/10.1145/3123024.3125607

Brailovskaia, J., & Margraf, J. (2024). From fear of missing out (FoMO) to addictive social media use: The role of social media flow and mindfulness. *Computers in Human Behavior*, *150*, 107984. https://doi.org/10.1016/j.chb.2023.107984

Brook, C. A., & Schmidt, L. A. (2008). Social anxiety disorder: A review of environmental risk factors. *Neuropsychiatric Disease and Treatment*, *4*(1), 123–143.

Bruening, M., van Woerden, I., Todd, M., Brennhofer, S., Laska, M. N., & Dunton, G. (2016). A Mobile Ecological Momentary Assessment Tool (devilSPARC) for Nutrition and Physical Activity Behaviors in College Students: A Validation Study. *Journal of Medical Internet Research*, *18*(7), e209. https://doi.org/10.2196/jmir.5969

Bufano, P., Laurino, M., Said, S., Tognetti, A., & Menicucci, D. (2023). Digital Phenotyping for Monitoring Mental Disorders: Systematic Review. *Journal of Medical Internet Research*, *25*, e46778. https://doi.org/10.2196/46778

Canada: Median annual family income by province | Statista. (n.d.). Retrieved May 17, 2024, from https://www.statista.com/statistics/467078/median-annual-family-income-in-canada-by-province/

Canada, P. H. A. of. (2019, April 24). *Infographic: Inequalities in mental illness hospitalization in Canada* [Education and awareness;statistics]. https://www.canada.ca/en/public-health/services/publications/science-research-data/inequalities-mental-illness-hospitalization-infographic.html

Cao, J., Truong, A. L., Banu, S., Shah, A. A., Sabharwal, A., & Moukaddam, N. (2020). Tracking and Predicting Depressive Symptoms of Adolescents Using Smartphone-Based Self-Reports, Parental Evaluations, and Passive Phone Sensor Data: Development and Usability Study. *JMIR Mental Health*, *7*(1), e14045. https://doi.org/10.2196/14045

Caporino, N. E., Sakolsky, D., Brodman, D. M., McGuire, J. F., Piacentini, J., Peris, T. S., Ginsburg, G. S., Walkup, J. T., Iyengar, S., Kendall, P. C., & Birmaher, B. (2017). Establishing Clinical Cutoffs for Response and Remission on the Screen for Child Anxiety Related Emotional Disorders (SCARED). *Journal of the American Academy of Child and Adolescent Psychiatry*, *56*(8), 696–702. https://doi.org/10.1016/j.jaac.2017.05.018

Celano, C. M., Daunis, D. J., Lokko, H. N., Campbell, K. A., & Huffman, J. C. (2016). Anxiety disorders and cardiovascular disease. *Current Psychiatry Reports*, *18*(11), 101. https://doi.org/10.1007/s11920-016-0739-5

Chikersal, P., Doryab, A., Tumminia, M., Villalba, D., Dutcher, J., Liu, X., Cohen, S., Creswell, K., Mankoff, J., Creswell, J., Goel, M., & Dey, A. (2021). Detecting Depression and Predicting its Onset Using Longitudinal Symptoms Captured by Passive Sensing: A Machine Learning Approach With Robust Feature Selection. *ACM Transactions on Computer-Human Interaction*, *28*, 1–41. https://doi.org/10.1145/3422821

Chin, W. Y., Choi, E. P. H., Chan, K. T. Y., & Wong, C. K. H. (2015). The Psychometric Properties of the Center for Epidemiologic Studies Depression Scale in Chinese Primary Care Patients: Factor Structure, Construct Validity, Reliability, Sensitivity and Responsiveness. *PLOS ONE*, *10*(8), e0135131. https://doi.org/10.1371/journal.pone.0135131

Cho, C.-H., Lee, T., Kim, M.-G., In, H. P., Kim, L., & Lee, H.-J. (2019). Mood Prediction of Patients With Mood Disorders by Machine Learning Using Passive Digital Phenotypes Based on the Circadian Rhythm: Prospective Observational Cohort Study. *Journal of Medical Internet Research*, *21*(4), e11029. https://doi.org/10.2196/11029

Choi, A., Ooi, A., & Lottridge, D. (2024). Digital Phenotyping for Stress, Anxiety, and Mild Depression: Systematic Literature Review. *JMIR mHealth and uHealth*, *12*, e40689. https://doi.org/10.2196/40689

Cooper, M. L., Frone, M. R., Russell, M., & Mudar, P. (1995). Drinking to regulate positive and negative emotions: A motivational model of alcohol use. *Journal of Personality and Social Psychology*, *69*(5), 990–1005. https://doi.org/10.1037/0022-3514.69.5.990

Cox, R. C., Sterba, S. K., Cole, D. A., Upender, R. P., & Olatunji, B. O. (2018). Time of day effects on the relationship between daily sleep and anxiety: An ecological momentary assessment approach. *Behaviour Research and Therapy*, *111*, 44–51. https://doi.org/10.1016/j.brat.2018.09.008

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J. F., & Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise*, *35*(8), 1381–1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB de Lijster, J. M., Dierckx, B., Utens, E. M. W. J., Verhulst, F. C., Zieldorff, C., Dieleman, G. C., & Legerstee, J. S. (2017). The Age of Onset of Anxiety Disorders. *Canadian Journal of Psychiatry. Revue Canadienne de Psychiatrie*, *62*(4), 237–246. https://doi.org/10.1177/0706743716640757

DXOMARK. (2023, February 10). GPS on smartphones: Testing the accuracy of location positioning. *DXOMARK*. https://www.dxomark.com/gps-on-your-smartphone-why-youre-not-always-there-when-it-says-youre-there/

Eirich, R., McArthur, B. A., Anhorn, C., McGuinness, C., Christakis, D. A., & Madigan, S. (2022). Association of Screen Time With Internalizing and Externalizing Behavior Problems in Children 12 Years or Younger. *JAMA Psychiatry*, *79*(5), 393–405. https://doi.org/10.1001/jamapsychiatry.2022.0155

Elmore, A. L., & Crouch, E. (2020). The Association of Adverse Childhood Experiences With Anxiety and Depression for Children and Youth, 8 to 17 Years of Age. *Academic Pediatrics*, *20*(5), 600–608. https://doi.org/10.1016/j.acap.2020.02.012

Elmquist, D. L., & McLaughlin, C. L. (2018). Social Media Use among Adolescents Coping with Mental Health. *Contemporary School Psychology*, *22*(4), 503–511. https://doi.org/10.1007/s40688-017-0167-5

Faravelli, C., Sauro, C. L., Godini, L., Lelli, L., Benni, L., Pietrini, F., Lazzeretti, L., Talamba, G. A., Fioravanti, G., & Ricca, V. (2012). Childhood stressful events, HPA axis and anxiety disorders. *World Journal of Psychiatry*, *2*(1), 13. https://doi.org/10.5498/wjp.v2.i1.13

Farhane-Medina, N. Z., Luque, B., Tabernero, C., & Castillo-Mayén, R. (2022). Factors associated with gender and sex differences in anxiety prevalence and comorbidity: A systematic review. *Science Progress*, *105*(4), 00368504221135469. https://doi.org/10.1177/00368504221135469

Faulstich. (1986). Assessment of depression in childhood and adolescence: An evaluation of the Center for Epidemiological Studies Depression Scale for Children (CES-DC). *American Journal of Psychiatry*, *143*(8), 1024–1027. https://doi.org/10.1176/ajp.143.8.1024

Figas, K., Giannouchos, T., & Crouch, E. (2022). *Rural-Urban Differences in Child and Adolescent Mental Health Prior to and During the COVID-19 Pandemic: Results From the National Survey of Children's Health—Rural Health Research Gateway*. Rural Health Research Gateway. https://www.ruralhealthresearch.org/publications/1558

Fisak, B., Penna, A., Mian, N. D., Lamoli, L., Margaris, A., & Cruz, S. A. M. F. D. (2023). The Effectiveness of Anxiety Interventions for Young Children: A Meta-Analytic Review. *Journal of Child and Family Studies*, *32*(8), 2546–2557. https://doi.org/10.1007/s10826-023-02596-y

Foerster, M., Henneke, A., Chetty-Mhlanga, S., & Röösli, M. (2019). Impact of Adolescents' Screen Time and Nocturnal Mobile Phone-Related Awakenings on Sleep and General Health Symptoms: A Prospective Cohort Study. *International Journal of Environmental Research and Public Health*, *16*(3), 518. https://doi.org/10.3390/ijerph16030518

Fukazawa, Y., Ito, T., Okimura, T., Yamashita, Y., Maeda, T., & Ota, J. (2019). Predicting anxiety state using smartphone-based passive sensing. *Journal of Biomedical Informatics*, *93*, 103151. https://doi.org/10.1016/j.jbi.2019.103151

Funkhouser, C. J., Trivedi, E., Li, L. Y., Helgren, F., Zhang, E., Sritharan, A., Cherner, R. A., Pagliaccio, D., Durham, K., Kyler, M., Tse, T. C., Buchanan, S. N., Allen, N. B., Shankman, S. A., & Auerbach, R. P. (2023). Detecting adolescent depression through passive monitoring of linguistic markers in smartphone communication. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*. https://doi.org/10.1111/jcpp.13931

Garber, J., & Weersing, V. R. (2010). Comorbidity of Anxiety and Depression in Youth: Implications for Treatment and Prevention. *Clinical Psychology : A Publication of the Division of Clinical Psychology of the American Psychological Association*, *17*(4), 293–306. https://doi.org/10.1111/j.1468-2850.2010.01221.x

Garcia, K. K. S., & Abrahão, A. A. (2021). Research Development Using REDCap Software. *Healthcare Informatics Research*, *27*(4), 341–349. https://doi.org/10.4258/hir.2021.27.4.341

Garnefski, N., Legerstee, J., Kraaij, V., Van den kommer, T., & Teerds, J. (2002). Cognitive coping strategies and symptoms of depression and anxiety: A comparison between adolescents and adults. *Journal of Adolescence*, *25*(6), 603–611. https://doi.org/10.1006/jado.2002.0507

Gillett, G., McGowan, N. M., Palmius, N., Bilderbeck, A. C., Goodwin, G. M., & Saunders, K. E. A. (2021). Digital Communication Biomarkers of Mood and Diagnosis in Borderline Personality Disorder, Bipolar Disorder, and Healthy Control Populations. *Frontiers in Psychiatry*, *12*, 610457. https://doi.org/10.3389/fpsyt.2021.610457

Goldstein-Piekarski, A. N., Williams, L. M., & Humphreys, K. (2016). A trans-diagnostic review of anxiety disorder comorbidity and the impact of multiple exclusion criteria on studying clinical outcomes in anxiety disorders. *Translational Psychiatry*, *6*(6), e847. https://doi.org/10.1038/tp.2016.108

Goodwin, R. D., Beesdo-Baum, K., Knappe, S., & Stein, D. J. (2013). Life course epidemiology of anxiety disorders. In K. C. Koenen, S. Rudenstine, E. Susser, & S. Galea (Eds.), *A Life Course Approach to Mental Disorders* (pp. 97–110). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199657018.003.0009

Graham, B., Bowes, L., & Ehlers, A. (2022). External Locus of Control but not Self-Esteem Predicts Increasing Social Anxiety Among Bullied Children. *Clinical Psychology in Europe*, *4*(2), e3809. https://doi.org/10.32872/cpe.3809

Grant, B. F., Stinson, F. S., Dawson, D. A., Chou, S. P., Dufour, M. C., Compton, W., Pickering, R. P., & Kaplan, K. (2004). Prevalence and co-occurrence of substance use disorders and independent mood and anxiety disorders: Results from the National Epidemiologic Survey on Alcohol and Related Conditions. *Archives of General Psychiatry*, *61*(8), 807–816. https://doi.org/10.1001/archpsyc.61.8.807

Hall, M., Scherner, P. V., Kreidel, Y., & Rubel, J. A. (2021). A Systematic Review of Momentary Assessment Designs for Mood and Anxiety Symptoms. *Frontiers in Psychology*, *12*. https://www.frontiersin.org/articles/10.3389/fpsyg.2021.642044

Harari, G. M., Müller, S. R., Stachl, C., Wang, R., Wang, W., Bühner, M., Rentfrow, P. J., Campbell, A. T., & Gosling, S. D. (2020). Sensing sociability: Individual differences in young adults' conversation, calling, texting, and app use behaviors in daily life: Journal of Personality and Social Psychology. *Journal of Personality and Social Psychology*, *119*(1), 204–228. https://doi.org/10.1037/pspp0000245

Haslbeck, J. M. B., Jover Martínez, A., Roefs, A., Fried, E. I., Lemmens, L. H. J. M., Groot, E., & Edelsbrunner, P. A. (2024). *Comparing Likert and Visual Analogue Scales in Ecological Momentary Assessment*. https://doi.org/10.31234/osf.io/yt8xw

Hauser, R. M. (1994). Measuring socioeconomic status in studies of child development. *Child Development*, *65*(6), 1541–1545. https://doi.org/10.1111/j.1467-8624.1994.tb00834.x

He, Y., Li, A., Li, K., & Xiao, J. (2021). Neuroticism vulnerability factors of anxiety symptoms in adolescents and early adults: An analysis using the bi-factor model and multi-wave longitudinal model. *PeerJ*, *9*, e11379. https://doi.org/10.7717/peerj.11379

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods*, *50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1

Howes Vallis, E., Zwicker, A., Uher, R., & Pavlova, B. (2020). Cognitive-behavioural interventions for prevention and treatment of anxiety in young children: A systematic review and meta-analysis. *Clinical Psychology Review*, *81*, 101904. https://doi.org/10.1016/j.cpr.2020.101904

Huang, Y., Xiong, H., Leach, K., Zhang, Y., Chow, P., Fua, K., Teachman, B. A., & Barnes, L. E. (2016). Assessing social anxiety using gps trajectories and point-of-interest data. *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 898–903. https://doi.org/10.1145/2971648.2971761

Humeniuk, R., Ali, R., Babor, T. F., Farrell, M., Formigoni, M. L., Jittiwutikarn, J., de Lacerda, R. B., Ling, W., Marsden, J., Monteiro, M., Nhiwatiwa, S., Pal, H., Poznyak, V., & Simon, S. (2008). Validation of the Alcohol, Smoking And Substance Involvement Screening Test (ASSIST). *Addiction (Abingdon, England)*, *103*(6), 1039–1047. https://doi.org/10.1111/j.1360-0443.2007.02114.x

Isaacs, J., MacKinnon, S., Joyce, K., & Stewart, S. (2021). Reactivity to Daily Self-Monitoring of Cannabis Use in Biological Females. *Cannabis*, *4*(2), 17–30. https://doi.org/10.26828/cannabis/2021.02.002

Jacobson, N. C., Lekkas, D., Huang, R., & Thomas, N. (2021). Deep Learning Paired with Wearable Passive Sensing Data Predicts Deterioration in Anxiety Disorder Symptoms across 17–18 Years. *Journal of Affective Disorders*, *282*, 104–111. https://doi.org/10.1016/j.jad.2020.12.086

Jacobson, N. C., Summers, B., & Wilhelm, S. (2020). Digital Biomarkers of Social Anxiety Severity: Digital Phenotyping Using Passive Smartphone Sensors. *Journal of Medical Internet Research*, *22*(5), e16875. https://doi.org/10.2196/16875

Jagesar, R. R., Vorstman, J. A., & Kas, M. J. (2021). Requirements and Operational Guidelines for Secure and Sustainable Digital Phenotyping: Design and Development Study. *Journal of Medical Internet Research*, *23*(4). https://doi.org/10.2196/20996

Kanwar, A., Malik, S., Prokop, L. J., Sim, L. A., Feldstein, D., Wang, Z., & Murad, M. H. (2013). The Association Between Anxiety Disorders and Suicidal Behaviors: A Systematic Review and Meta-Analysis. *Depression and Anxiety*, *30*(10), 917–929. https://doi.org/10.1002/da.22074

Katzman, M. A., Bleau, P., Blier, P., Chokka, P., Kjernisted, K., Van Ameringen, M., & the Canadian Anxiety Guidelines Initiative Group on behalf of the Anxiety Disorders Association of Canada/Association Canadienne des troubles anxieux and McGill University. (2014). Canadian clinical practice guidelines for the management of anxiety, posttraumatic stress and obsessive-compulsive disorders. *BMC Psychiatry*, *14*(Suppl 1), S1. https://doi.org/10.1186/1471-244X-14-S1-S1

Kelly, S. M., Gryczynski, J., Mitchell, S. G., Kirk, A., O'Grady, K. E., & Schwartz, R. P. (2014). Validity of Brief Screening Instrument for Adolescent Tobacco, Alcohol, and Drug Use. *Pediatrics*, *133*(5), 819–826. https://doi.org/10.1542/peds.2013-2346

Kendall, P. C., & Peterman, J. S. (2015). CBT for Adolescents With Anxiety: Mature Yet Still Developing. *American Journal of Psychiatry*, *172*(6), 519–530. https://doi.org/10.1176/appi.ajp.2015.14081061

Kessler, R. C., Petukhova, M., Sampson, N. A., Zaslavsky, A. M., & Wittchen, H.-U. (2012). Twelve-month and lifetime prevalence and lifetime morbid risk of anxiety and mood disorders in the United States. *International Journal of Methods in Psychiatric Research*, *21*(3), 169–184. https://doi.org/10.1002/mpr.1359

Kim, H., Lee, S., Lee, S., Hong, S., Kang, H., & Kim, N. (2019). Depression Prediction by Using Ecological Momentary Assessment, Actiwatch Data, and Machine Learning: Observational Study on Older Adults Living Alone. *JMIR mHealth and uHealth*, *7*(10), e14149. https://doi.org/10.2196/14149

Klaufus, L., Verlinden, E., van der Wal, M., Cuijpers, P., Chinapaw, M., & Smit, F. (2022). Adolescent anxiety and depression: Burden of disease study in 53,894 secondary school pupils in the Netherlands. *BMC Psychiatry*, *22*(1), 225. https://doi.org/10.1186/s12888-022-03868-5

Klein, Z., Shner, G., Ginat-Frolich, R., Vervliet, B., & Shechner, T. (2020). The effects of age and trait anxiety on avoidance learning and its generalization. *Behaviour Research and Therapy*, *129*, 103611. https://doi.org/10.1016/j.brat.2020.103611

Klemmer, B., Kinnafick, F. E., Spray, C., & Chater, A. M. (2023). The effectiveness of structured sport and exercise interventions in enhancing the mental health of adolescents with mild to moderate mental health problems: A systematic review. *International Review of Sport and Exercise Psychology*, *0*(0), 1–24. https://doi.org/10.1080/1750984X.2023.2266823

Koenig, H. G., & Al Zaben, F. (2021). Moral Injury: An Increasingly Recognized and Widespread Syndrome. *Journal of Religion and Health*, *60*(5), 2989–3011. https://doi.org/10.1007/s10943-021-01328-0

Kolenikov, S., & Bollen, K. A. (2012). Testing Negative Error Variances: Is a Heywood Case a Symptom of Misspecification? *Sociological Methods & Research*, *41*(1), 124–167. https://doi.org/10.1177/0049124112442138

Komarahadi, F. L., Maurischat, C., Härter, M., & Bengel, J. (2004). [Relationship of depression and anxiety with social desirability in chronic pain patients]. *Schmerz (Berlin, Germany)*, *18*(1), 38–44. https://doi.org/10.1007/s00482-003-0282-2

Lan, Y., Roberts, H., Kwan, M.-P., & Helbich, M. (2022). Daily space-time activities, multiple environmental exposures, and anxiety symptoms: A cross-sectional mobile phone-based sensing study. *Science of The Total Environment*, *834*, 155276. https://doi.org/10.1016/j.scitotenv.2022.155276

Laugharne, J., Lillee, A., & Janca, A. (2010). Role of psychological trauma in the cause and treatment of anxiety and depressive disorders. *Current Opinion in Psychiatry*, *23*(1), 25. https://doi.org/10.1097/YCO.0b013e3283345dc5

Lebowitz, E. R., Shic, F., Campbell, D., Basile, K., & Silverman, W. K. (2015). Anxiety sensitivity moderates behavioral avoidance in anxious youth. *Behaviour Research and Therapy*, *74*, 11–17. https://doi.org/10.1016/j.brat.2015.08.009

Lijster, J. M. de, Dierckx, B., Utens, E. M. W. J., Verhulst, F. C., Zieldorff, C., Dieleman, G. C., & Legerstee, J. S. (2017). The Age of Onset of Anxiety Disorders. *Canadian Journal of Psychiatry. Revue Canadienne De Psychiatrie*, *62*(4), 237–246. https://doi.org/10.1177/0706743716640757

Lind, M. N., Byrne, M. L., Wicks, G., Smidt, A. M., & Allen, N. B. (2018). The Effortless Assessment of Risk States (EARS) Tool: An Interpersonal Approach to Mobile Sensing. *JMIR Mental Health*, *5*(3), e10334. https://doi.org/10.2196/10334

Lind, M. N., Kahn, L. E., Crowley, R., Reed, W., Wicks, G., & Allen, N. B. (2023). Reintroducing the Effortless Assessment Research System (EARS). *JMIR Mental Health*, *10*, e38920. https://doi.org/10.2196/38920

Loerinc, A. G., Meuret, A. E., Twohig, M. P., Rosenfield, D., Bluett, E. J., & Craske, M. G. (2015). Response rates for CBT for anxiety disorders: Need for standardized criteria. *Clinical Psychology Review*, *42*, 72–82. https://doi.org/10.1016/j.cpr.2015.08.004

Louise, P., Siobhan, O., Louise, M., & Jean, G. (2017). The burden of generalized anxiety disorder in Canada. *Health Promotion and Chronic Disease Prevention in Canada : Research, Policy and Practice*, *37*(2), 54–62.

Mabrouk, A., Mbithi, G., Chongwo, E., Too, E., Sarki, A., Namuguzi, M., Atukwatse, J., Ssewanyana, D., & Abubakar, A. (2022). Mental health interventions for adolescents in sub-Saharan Africa: A scoping review. *Frontiers in Psychiatry*, *13*, 937723. https://doi.org/10.3389/fpsyt.2022.937723

MacIntyre, M. M., Zare, M., & Williams, M. T. (2023). Anxiety-Related Disorders in the Context of Racism. *Current Psychiatry Reports*, *25*(2), 31–43. https://doi.org/10.1007/s11920-022-01408-2

Mack, D. L., DaSilva, A. W., Rogers, C., Hedlund, E., Murphy, E. I., Vojdanovski, V., Plomp, J., Wang, W., Nepal, S. K., Holtzheimer, P. E., Wagner, D. D., Jacobson, N. C., Meyer, M. L., Campbell, A. T., & Huckins, J. F. (2021). Mental Health and Behavior of College Students During the COVID-19 Pandemic: Longitudinal Mobile Smartphone and Ecological Momentary Assessment Study, Part II. *Journal of Medical Internet Research*, *23*(6), e28892. https://doi.org/10.2196/28892

Mackinnon, S. P., Curtis, R., & O'Connor, R. (2020). *A Tutorial in Longitudinal Measurement Invariance and Cross-lagged Panel Models Using Lavaan*. https://doi.org/10.31234/osf.io/tkzrb

MacLeod, L., Suruliraj, B., Gall, D., Bessenyei, K., Hamm, S., Romkey, I., Bagnell, A., Mattheisen, M., Muthukumaraswamy, V., Orji, R., & Meier, S. (2021). A Mobile Sensing App to Monitor Youth Mental Health: Observational Pilot Study. *JMIR mHealth and uHealth*, *9*(10), e20638. https://doi.org/10.2196/20638

Maeng, L. Y., & Milad, M. R. (2015). Sex Differences in Anxiety Disorders: Interactions between Fear, Stress, and Gonadal Hormones. *Hormones and Behavior*, *76*, 106–117. https://doi.org/10.1016/j.yhbeh.2015.04.002

Marin-Dragu, S., Forbes, A., Sheikh, S., Iyer, R. S., Pereira Dos Santos, D., Alda, M., Hajek, T., Uher, R., Wozney, L., Paulovich, F. V., Campbell, L. A., Yakovenko, I., Stewart, S. H., Corkum, P., Bagnell, A., Orji, R., & Meier, S. (2023). Associations of active and passive smartphone use with measures of youth mental health during the COVID-19 pandemic. *Psychiatry Research*, *326*, 115298. https://doi.org/10.1016/j.psychres.2023.115298

Matteo, D. D., Fotinos, K., Lokuge, S., Mason, G., Sternat, T., Katzman, M. A., & Rose, J. (2021). Automated Screening for Social Anxiety, Generalized Anxiety, and Depression From Objective Smartphone-Collected Data: Cross-sectional Study. *Journal of Medical Internet Research*, *23*(8), e28918. https://doi.org/10.2196/28918

Maurya, C., Muhammad, T., Maurya, P., & Dhillon, P. (2022). The association of smartphone screen time with sleep problems among adolescents and young adults: Cross-sectional findings from India. *BMC Public Health*, *22*, 1686. https://doi.org/10.1186/s12889-022-14076-x

McNeilly, E. A., Mills, K. L., Kahn, L. E., Crowley, R., Pfeifer, J. H., & Allen, N. B. (2023). Adolescent Social Communication Through Smartphones: Linguistic Features of Internalizing Symptoms and Daily Mood. *Clinical Psychological Science : A Journal of the Association for Psychological Science*, *11*(6), 1090–1107. https://doi.org/10.1177/21677026221125180

Meier, S. M., Manuel, M., Ole, M., Mortensen, P. B., Laursen, T. M., & Penninx, B. W. (2016). Increased mortality among people with anxiety disorders: Total population study. *The British Journal of Psychiatry*, *209*(3), 216–221. https://doi.org/10.1192/bjp.bp.115.171975

Meier, S. M., Pavlova, B., Dalsgaard, S., Nordentoft, M., Mors, O., Mortensen, P. B., & Uher, R. (2018). Attention-deficit hyperactivity disorder and anxiety disorders as precursors of bipolar disorder onset in adulthood. *The British Journal of Psychiatry*, *213*(3), 555–560. https://doi.org/10.1192/bjp.2018.111

Meier, S. M., Uher, R., Mors, O., Dalsgaard, S., Munk-Olsen, T., Laursen, T. M., Mattheisen, M., Nordentoft, M., Mortensen, P. B., & Pavlova, B. (2016). Specific anxiety disorders and subsequent risk for bipolar disorder: A nationwide study. *World Psychiatry*, *15*(2), 187–188. https://doi.org/10.1002/wps.20314

Melcher, J., Lavoie, J., Hays, R., D'Mello, R., Rauseo-Ricupero, N., Camacho, E., Rodriguez-Villa, E., Wisniewski, H., Lagan, S., Vaidyam, A., & Torous, J. (2023). Digital phenotyping of student mental health during COVID-19: An observational study of 100 college students. *Journal of American College Health*, *71*(3), 736–748. https://doi.org/10.1080/07448481.2021.1905650

Mendes, J. P. M., Moura, I. R., Ven, P. V. de, Viana, D., Silva, F. J. S., Coutinho, L. R., Teixeira, S., Rodrigues, J. J. P. C., & Teles, A. S. (2022). Sensing Apps and Public Data Sets for Digital Phenotyping of Mental Health: Systematic Review. *Journal of Medical Internet Research*, *24*(2), e28735. https://doi.org/10.2196/28735

Mendu, S., Baglione, A., Baee, S., Wu, C., Ng, B., Shaked, A., Clore, G., Boukhechba, M., & Barnes, L. (2020). A Framework for Understanding the Relationship between Social Media Discourse and Mental Health. *Proceedings of the ACM on Human-Computer Interaction*, *4*(CSCW2), 1–23. https://doi.org/10.1145/3415215

Mengelkoch, S., Moriarity, D. P., Novak, A. M., Snyder, M. P., Slavich, G. M., & Lev-Ari, S. (2024). Using Ecological Momentary Assessments to Study How Daily Fluctuations in Psychological States Impact Stress, Well-Being, and Health. *Journal of Clinical Medicine*, *13*(1), Article 1. https://doi.org/10.3390/jcm13010024

Messner, E.-M., Sariyska, R., Mayer, B., Montag, C., Kannen, C., Schwerdtfeger, A., & Baumeister, H. (2019). Insights – Future Implications of Passive Smartphone Sensing in the Therapeutic Context. *Verhaltenstherapie*, *32*(Suppl. 1), 86–95. https://doi.org/10.1159/000501951

Meyerhoff, J., Liu, T., Kording, K. P., Ungar, L. H., Kaiser, S. M., Karr, C. J., & Mohr, D. C. (2021). Evaluation of Changes in Depression, Anxiety, and Social Anxiety Using Smartphone Sensor Features: Longitudinal Cohort Study. *Journal of Medical Internet Research*, *23*(9), e22844. https://doi.org/10.2196/22844

Montag, C., & Baumeister, H. (Eds.). (2023). *Digital Phenotyping and Mobile Sensing: New Developments in Psychoinformatics*. Springer International Publishing. https://doi.org/10.1007/978-3-030-98546-2

Moshe, I., Terhorst, Y., Opoku Asare, K., Sander, L. B., Ferreira, D., Baumeister, H., Mohr, D. C., & Pulkki-Råback, L. (2021). Predicting Symptoms of Depression and Anxiety Using Smartphone and Wearable Data. *Frontiers in Psychiatry*, *12*, 625247. https://doi.org/10.3389/fpsyt.2021.625247

Narchal, R. (2017). Loneliness and Anxiety Sensitivity: Understanding Behavioural Avoidance in the Lonely. *Acta Psychopathologica*, *03*(05). https://doi.org/10.4172/2469-6676.100130

Nicholas, J., Shilton, K., Schueller, S. M., Gray, E. L., Kwasny, M. J., & Mohr, D. C. (2019). The Role of Data Type and Recipient in Individuals' Perspectives on Sharing Passively Collected Smartphone Data for Mental Health: Cross-Sectional Questionnaire Study. *JMIR mHealth and uHealth*, *7*(4), e12578. https://doi.org/10.2196/12578

Niemeijer, K., Mestdagh, M., & Kuppens, P. (2022). Tracking Subjective Sleep Quality and Mood With Mobile Sensing: Multiverse Study. *Journal of Medical Internet Research*, *24*(3), e25643. https://doi.org/10.2196/25643

Nishiyama, Y., Ferreira, D., Sasaki, W., Okoshi, T., Nakazawa, J., Dey, A. K., & Sezaki, K. (2020). Using iOS for inconspicuous data collection: A real-world assessment. *Adjunct Proceedings of the 2020 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2020 ACM International Symposium on Wearable Computers*, 261–266. https://doi.org/10.1145/3410530.3414369

Orr, M., MacLeod, L., Bagnell, A., McGrath, P., Wozney, L., & Meier, S. (2023). The comfort of adolescent patients and their parents with mobile sensing and digital phenotyping. *Computers in Human Behavior*, *140*, 107603. https://doi.org/10.1016/j.chb.2022.107603

Pittig, A., Kotter, R., & Hoyer, J. (2019). The Struggle of Behavioral Therapists With Exposure: Self-Reported Practicability, Negative Beliefs, and Therapist Distress About Exposure-Based Interventions. *Behavior Therapy*, *50*(2), 353–366. https://doi.org/10.1016/j.beth.2018.07.003

Putnick, D. L., & Bornstein, M. H. (2016). Measurement Invariance Conventions and Reporting: The State of the Art and Future Directions for Psychological Research. *Developmental Review : DR*, *41*, 71–90. https://doi.org/10.1016/j.dr.2016.06.004

Racine, N., McArthur, B. A., Cooke, J. E., Eirich, R., Zhu, J., & Madigan, S. (2021). Global Prevalence of Depressive and Anxiety Symptoms in Children and Adolescents During COVID-19: A Meta-analysis. *JAMA Pediatrics*, *175*(11), 1142–1150. https://doi.org/10.1001/jamapediatrics.2021.2482

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Raina, P. S., Wolfson, C., Kirkland, S. A., Griffith, L. E., Oremus, M., Patterson, C., Tuokko, H., Penning, M., Balion, C. M., Hogan, D., Wister, A., Payette, H., Shannon, H., & Brazil, K. (2009). The Canadian longitudinal study on aging (CLSA). *Canadian Journal on Aging = La Revue Canadienne Du Vieillissement*, *28*(3), 221–229. https://doi.org/10.1017/S0714980809990055

Ramsawh, H. J., Raffa, S. D., Edelen, M. O., Rende, R., & Keller, M. B. (2009). Anxiety in middle adulthood: Effects of age and time on the 14-year course of panic disorder, social phobia and generalized anxiety disorder. *Psychological Medicine*, *39*(4), 615–624. https://doi.org/10.1017/S0033291708003954

Ratan, Z. A., Parrish, A.-M., Alotaibi, M. S., & Hosseinzadeh, H. (2022). Prevalence of Smartphone Addiction and Its Association with Sociodemographic, Physical and Mental Well-Being: A Cross-Sectional Study among the Young Adults of Bangladesh. *International Journal of Environmental Research and Public Health*, *19*(24), 16583. https://doi.org/10.3390/ijerph192416583

Renn, B. N., Pratap, A., Atkins, D. C., Mooney, S. D., & Areán, P. A. (2018). Smartphone-based passive assessment of mobility in depression: Challenges and opportunities. *Mental Health and Physical Activity*, *14*, 136–139. https://doi.org/10.1016/j.mhpa.2018.04.003

Richard Dawkins. (1999). *The Extended Phenotype: The Long Reach of the Gene: Vol. Revised edition*. OUP Oxford. http://ezproxy.library.dal.ca/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=e000xna&AN=665066&site=ehost-live

Romans, S., Cohen, M., & Forte, T. (2011). Rates of depression and anxiety in urban and rural Canada. *Social Psychiatry and Psychiatric Epidemiology*, *46*(7), 567–575. https://doi.org/10.1007/s00127-010-0222-2

Ruscio, A. M., Brown, T. A., Chiu, W. T., Sareen, J., Stein, M. B., & Kessler, R. C. (2008). Social Fears and Social Phobia in the United States: Results from the National Comorbidity Survey Replication. *Psychological Medicine*, *38*(1), 15–28. https://doi.org/10.1017/S0033291707001699

Saeb, S., Lattie, E. G., Kording, K. P., & Mohr, D. C. (2017). Mobile Phone Detection of Semantic Location and Its Relationship to Depression and Anxiety. *JMIR mHealth and uHealth*, *5*(8), e7297. https://doi.org/10.2196/mhealth.7297

Saeb, S., Zhang, M., Karr, C. J., Schueller, S. M., Corden, M. E., Kording, K. P., & Mohr, D. C. (2015). Mobile Phone Sensor Correlates of Depressive Symptom Severity in Daily-Life Behavior: An Exploratory Study. *Journal of Medical Internet Research*, *17*(7), e4273. https://doi.org/10.2196/jmir.4273

Santiago, F. L., Oliveira Da Silva, A., Souza Silva, R. I. D., Castro Melo, W. V. D., Rodrigues Filho, E. D. A., Torres Pirauá, A. L., Farah, B. Q., & Machado-Oliveira, L. (2022). Association between screen time exposure, anxiety, and sleep quality in adolescents. *Salud Mental*, *45*(3), 125–133. https://doi.org/10.17711/SM.0185-3325.2022.017

Santos, R. M. S., Ventura, S. de A., Nogueira, Y. J. de A., Mendes, C. G., Paula, J. J. de, Miranda, D. M., & Romano-Silva, M. A. (2024). The Associations Between Screen Time and Mental Health in Adults: A Systematic Review. *Journal of Technology in Behavioral Science*. https://doi.org/10.1007/s41347-024-00398-7

Sardar, A. W., Ullah, F., Bacha, J., Khan, J., Ali, F., & Lee, S. (2022). Mobile sensors based platform of Human Physical Activities Recognition for COVID-19 spread minimization. *Computers in Biology and Medicine*, *146*, 105662. https://doi.org/10.1016/j.compbiomed.2022.105662

Sato, H., & Kawahara, J. (2011). Selective bias in retrospective self-reports of negative mood states. *Anxiety, Stress, & Coping*, *24*(4), 359–367. https://doi.org/10.1080/10615806.2010.543132

Sequeira, M., Singh, S., Fernandes, L., Gaikwad, L., Gupta, D., Chibanda, D., & Nadkarni, A. (2022). Adolescent Health Series: The status of adolescent mental health research, practice and policy in sub-Saharan Africa: A narrative review. *Tropical Medicine & International Health*, *27*(9), 758–766. https://doi.org/10.1111/tmi.13802

Sharifian, N., Kraal, A. Z., Zaheed, A. B., Sol, K., Morris, E. P., & Zahodne, L. B. (2021). Measurement Invariance of Social Media Use in Younger and Older Adults and Links to Socioemotional Health. *Innovation in Aging*, *5*(2), igab009. https://doi.org/10.1093/geroni/igab009

Shiffman, S. (2009). Ecological Momentary Assessment (EMA) in Studies of Substance Use. *Psychological Assessment*, *21*(4), 486–497. https://doi.org/10.1037/a0017074

Shimada-Sugimoto, M., Otowa, T., & Hettema, J. M. (2015). Genetics of anxiety disorders: Genetic epidemiological and molecular studies in humans. *Psychiatry and Clinical Neurosciences*, *69*(7), 388–401. https://doi.org/10.1111/pcn.12291

Sohn, S. Y., Krasnoff, L., Rees, P., Kalk, N. J., & Carter, B. (2021). The Association Between Smartphone Addiction and Sleep: A UK Cross-Sectional Study of Young Adults. *Frontiers in Psychiatry*, *12*. https://doi.org/10.3389/fpsyt.2021.629407

Stamatis, C. A., Meyerhoff, J., Meng, Y., Lin, Z. C. C., Cho, Y. M., Liu, T., Karr, C. J., Liu, T., Curtis, B. L., Ungar, L. H., & Mohr, D. C. (2024). Differential temporal utility of passively sensed smartphone features for depression and anxiety symptom prediction: A longitudinal cohort study. *Npj Mental Health Research*, *3*(1), 1–8. https://doi.org/10.1038/s44184-023-00041-y

Starcevic, V. (2022). Representation of Benzodiazepines in Treatment Guidelines: The Paradox of Undesirable Objectivity. *Psychotherapy and Psychosomatics*, *91*(5), 295–299. https://doi.org/10.1159/000524772

Stewart, S. H., & Conrod, P. J. (Eds.). (2008). *Anxiety and substance use disorders: The vicious cycle of comorbidity*. Springer.

Stoyanova, M., & Hope, D. A. (2012). Gender, gender roles, and anxiety: Perceived confirmability of self report, behavioral avoidance, and physiological reactivity. *Journal of Anxiety Disorders*, *26*(1), 206–214. https://doi.org/10.1016/j.janxdis.2011.11.006

Su, L., Wang, K., Fan, F., Su, Y., & Gao, X. (2008). Reliability and validity of the screen for child anxiety related emotional disorders (SCARED) in Chinese children. *Journal of Anxiety Disorders*, *22*(4), 612–621. https://doi.org/10.1016/j.janxdis.2007.05.011

Szuhany, K. L., & Simon, N. M. (2022). Anxiety Disorders: A Review. *JAMA*, *328*(24), 2431. https://doi.org/10.1001/jama.2022.22744

Tlachac, M., Melican, V., Reisch, M., & Rundensteiner, E. (2021). Mobile Depression Screening with Time Series of Text Logs and Call Logs. *2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, 1–4. https://doi.org/10.1109/BHI50953.2021.9508582

Vaidyam, A., Halamka, J., & Torous, J. (2022). Enabling Research and Clinical Use of Patient-Generated Health Data (the mindLAMP Platform): Digital Phenotyping Study. *JMIR mHealth and uHealth*, *10*(1), e30557. https://doi.org/10.2196/30557

van Santvoort, F., Hosman, C. M. H., Janssens, J. M. A. M., van Doesum, K. T. M., Reupert, A., & van Loon, L. M. A. (2015). The Impact of Various Parental Mental Disorders on Children's Diagnoses: A Systematic Review. *Clinical Child and Family Psychology Review*, *18*(4), 281–299. https://doi.org/10.1007/s10567-015-0191-9

Vasey, M. W., Bosmans, G., & Ollendick, T. H. (2014). The Developmental Psychopathology of Anxiety. In M. Lewis & K. D. Rudolph (Eds.), *Handbook of Developmental Psychopathology* (pp. 543–560). Springer US. https://doi.org/10.1007/978-1-4614-9608-3_27

Vega, J., Li, M., Aguillera, K., Goel, N., Joshi, E., Khandekar, K., Durica, K. C., Kunta, A. R., & Low, C. A. (2021). Reproducible Analysis Pipeline for Data Streams: Open-Source Software to Process Data Collected With Mobile Devices. *Frontiers in Digital Health*, *3*. https://doi.org/10.3389/fdgth.2021.769823

Vermani, M., Marcus, M., & Katzman, M. A. (2011). Rates of detection of mood and anxiety disorders in primary care: A descriptive, cross-sectional study. *The Primary Care Companion for CNS Disorders*, *13*(2), PCC.10m01013. https://doi.org/10.4088/PCC.10m01013

Waller, C. R., Leal, A. S. M., & Silvers, J. A. (2023). Disparities in Depression and Anxiety That Impact Self-Identified Sexual Minority People Affect a Broader Group of Same-Gender Attracted Young Adults. *Journal of Adolescent Health*, *73*(4), 739–745. https://doi.org/10.1016/j.jadohealth.2023.05.024

Wang, R., Wang, W., daSilva, A., Huckins, J. F., Kelley, W. M., Heatherton, T. F., & Campbell, A. T. (2018). Tracking Depression Dynamics in College Students Using Mobile Phone and Wearable Sensing. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *2*(1), 43:1-43:26. https://doi.org/10.1145/3191775

Wanjau, M. N., Möller, H., Haigh, F., Milat, A., Hayek, R., Lucas, P., & Veerman, J. L. (2023). Physical Activity and Depression and Anxiety Disorders: A Systematic Review of Reviews and Assessment of Causality. *AJPM Focus*, *2*(2), 100074. https://doi.org/10.1016/j.focus.2023.100074

Weisberg, R. B., Dyck, I., Culpepper, L., & Keller, M. B. (2007). Psychiatric treatment in primary care patients with anxiety disorders: A comparison of care received from primary care providers and psychiatrists. *The American Journal of Psychiatry*, *164*(2), 276–282. https://doi.org/10.1176/ajp.2007.164.2.276

WHO ASSIST Working Group. (2002). The Alcohol, Smoking and Substance Involvement Screening Test (ASSIST): Development, reliability and feasibility. *Addiction (Abingdon, England)*, *97*(9), 1183–1194. https://doi.org/10.1046/j.1360-0443.2002.00185.x

Wolfson, A. R., Carskadon, M. A., Acebo, C., Seifer, R., Fallone, G., Labyak, S. E., & Martin, J. L. (2003). Evidence for the validity of a sleep habits survey for adolescents. *Sleep*, *26*(2), 213–216. https://doi.org/10.1093/sleep/26.2.213

Woodward, L. J., & Fergusson, D. M. (2001). Life course outcomes of young people with anxiety disorders in adolescence. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(9), 1086–1093. https://doi.org/10.1097/00004583-200109000-00018

Xiao, S., Shi, L., Zhang, J., Li, X., Lin, H., Xue, Y., Xue, B., Chen, Y., Zhou, G., & Zhang, C. (2023). The role of anxiety and depressive symptoms in mediating the relationship between subjective sleep quality and cognitive function among older adults in China. *Journal of Affective Disorders*, *325*, 640–646. https://doi.org/10.1016/j.jad.2023.01.048

Xu, X., Liu, X., Zhang, H., Wang, W., Nepal, S., Sefidgar, Y., Seo, W., Kuehn, K. S., Huckins, J. F., Morris, M. E., Nurius, P. S., Riskin, E. A., Patel, S., Althoff, T., Campbell, A., Dey, A. K., & Mankoff, J. (2022). GLOBEM: Cross-Dataset Generalization of Longitudinal Human Behavior Modeling. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *6*(4), 1–34. https://doi.org/10.1145/3569485

Zhang, Y., Folarin, A. A., Sun, S., Cummins, N., Vairavan, S., Bendayan, R., Ranjan, Y., Rashid, Z., Conde, P., Stewart, C., Laiou, P., Sankesara, H., Matcham, F., White, K. M., Oetzmann, C., Ivan, A., Lamers, F., Siddi, S., Vilella, E., … RADAR-CNS consortium. (2022). Longitudinal Relationships Between Depressive Symptom Severity and Phone-Measured Mobility: Dynamic Structural Equation Modeling Study. *JMIR Mental Health*, *9*(3), e34898. https://doi.org/10.2196/34898

APPENDIX A: Tables

Table 1

Comparison of three other available mobile sensing apps.

App Name	Platforms Available	Sensors Available	Citation
AWARE Framework	iOS; Android	GPS; accelerometer; call logs; text logs; screen usage; light sensor; physical activity recognition; battery status; app usage; Bluetooth connections; WiFi status	Nishiyama et al. (2020)
EARS	iOS; Android	GPS; accelerometer; call logs (iOS only); text logs; music and media; screen usage; app usage; light sensor (Android only); battery status (iOS only)	Lind et al. (2018; 2023)
mindLAMP	iOS; Android	GPS; acclerometer; voice; screen usage; metadata; call logs; text logs; Apple Health and Apple SensorKit; Google Fit	Vaidyam et al. (2022)
PROSIT	iOS; Android	GPS; accelerometer*; call logs; voice*; screen usage*; app usage; light/noise; physical activity recognition; magnetometer*; Bluetooth status*; power state*; gyroscope*; Apple Health and Apple SensorKit*; Google Fit*	MacLeod et al. (2021)

Note. *Stars (*) for the PROSIT app indicate sensors which were not directly used to engineer features in the present analyses.*

Table 2

Descriptive sample demographic statistics.

	Youth ($N = 65$)	Adults ($N = 96$)
Mean Age in Years (Range)	19.5 (16-21)	32.8 (27-40)
Phone Type		
iOS	27 (41.5%)	42 (43.8%)
Android	38 (58.5%)	54 (56.3%)
Biological Sex		
Female	36 (55.4%)	58 (60.4%)
Male	28 (41.31%)	38 (39.6%)
Gender Identity		
Woman/Girl	33 (50.8%)	56 (58.3%)
Man/Boy	28 (43.1%)	38 (39.6%)
Non-Binary/ENBY	2 (3.1%)	2 (2.1%)
Agender	1 (1.5%)	0 (0%)
Sexual Orientation		
Asexual	2 (3.1%)	3 (3.1%)
Bisexual	11 (16.9%)	8 (8.3%)
Gay	1 (1.5%)	0 (0%)
Lesbian	1 (1.5%)	1 (1%)
Pansexual	2 (3.1%)	6 (6.3%)
Queer	2 (3.1%)	0 (0%)
Questioning	2 (3.1%)	1 (1%)
Straight	41 (63.1%)	77 (80.2%)
Ethnicity		
Asian	18 (27.7%)	9 (9.4%)
Black/African	16 (24.6%)	18 (18.8%)
White	30 (46.2%)	68 (70.8%)
Hispanic	3 (4.6%)	3 (3.1%)
Indigenous	1 (1.5%)	3 (3.1%)
Pacific Islander	1 (1.5%)	0 (0%)
Middle Eastern	1 (1.5%)	0 (0%)
Country		
Canada	22 (33.85%)	48 (50%)
United States	8 (12.31%)	20 (20.83%)
Kenya	12 (18.46%)	19 (19.79%)
Nigeria	1 (1.54%)	0 (0%)
United Kingdom	1 (1.54%)	2 (2.08%)
The Bahamas	1 (1.54%)	0 (0%)
Mental Health Diagnoses		
No Diagnoses	48 (73.85%)	64 (66.67%)

Depression	8 (12.3%)	19 (19.8%)
Bipolar Disorder	1 (1.5%)	3 (3.1%)
Premenstrual Dysphoric Mood Disorder	0 (0%)	1 (1.0%)
Anroexia Nervosa	0 (0%)	2 (2.1%)
Bulimia	0 (0%)	1 (1.0%)
Attention Deficit / Hyperactivity Disorder	1 (1.5%)	8 (8.3%)
Autism Spectrum Disorder	2 (3.1%)	3 (3.1%)
Tourette's Disorder		
Anxiety Disorder	14 (21.5%)	28 (29.2%)
Panic Disorder	2 (3.1%)	1 (1.0%)
Obsessive Compulsive Disorder	1 (1.5%)	4 (4.2%)
Posttraumatic Stress Disorder		7 (7.3%)
Specific Phobia		1 (1.0%)
Seasonal Affective Disorder		2 (2.1%)
Social Anxiety Disorder	3 (4.6%)	5 (5.2%)
Location		
Urban	47 (72.3%)	57 (59.4%)
Suburban	16 (24.6%)	26 (27.1%)
Rural	2 (3.1%)	13 (13.5%)
Parental Education		
Did not finish high school	2 (3.1%)	NA
Finished high school	6 (9.2%)	NA
Further education	3 (4.6%)	NA
Student at university	2 (3.1%)	NA
Finished university	51 (78.5%)	NA
Annual Income		
Less than $20k	NA	5 (5.2%)
$20k-$50k	NA	21 (21.9%)
$50k-$100k	NA	26 (27.1%)
$100k-$150k	NA	25 (26.0%)
$150k+	NA	17 (17.7%)
SCARED/SCAARED Mean Score (*SD*)		
Overall Score	47.48 (10.91)	44.56 (13.87)
Panic Symptoms	9.58 (3.08)	14.91 (7.26)
Generalized Anxiety Symptoms	11.89 (3.82)	17.55 (4.86)
Separation Anxiety Symptoms	7.28 (3.34)	5.96 (3.61)
Social Anxiety Symptoms	9.40 (3.11)	9.41 (3.44)
School Avoidance Symptoms	3.43 (1.74)	NA

Note. Percentage values that do not add to 100 indicate missingness.

Table 3

Mobile sensing feature descriptives.

Feature	Youth		Adults	
Mobility	Mean	SD	Mean	SD
Location Variance	-6.723064	3.704153	-5.302661	4.840735
Haversine Distance	116632	376003.8	100757800	1189310000
Number of Stay Points	19.13793	14.47787	24.31176	35.83612
Location Entropy	1.289588	0.7419626	1.099307	0.7822579
Normalized Entropy	0.4698186	0.2419966	0.3811832	0.2120649
Sociability				
Total Call Duration (secs)	1794.145	3048.262	2380.371	2903.884
Total Incoming Calls	12.76613	23.84847	9.666667	13.87304
Total Outgoing Calls	23.40323	47.0471	19.21505	27.59528
Screen Usage				
Total Screen Time (secs)	3221.291	2723.368	4044.299	3456.234
Total Unlocks	192.6239	218.8812	279.1739	294.287
Sleep				
Total Evening Screen Time (secs)	937.735	1239.955	1055.636	1077.191
Total Evening Unlocks	68.17094	80.24223	82.40761	79.70733
Total Nightly Screen Time (secs)	796.1368	1060.348	890.0435	1020.888
Total Nightly Unlocks	26.65812	29.88748	39.34239	53.35536
Ambient Light Intensity	7001.924	35199.77	1744.548	3233.362
Proportion Evening Stationary Time	0.6735199	0.2761628	0.6578664	0.2400534
Proportion Night Stationary Time	0.6902991	0.3366885	0.7157568	0.2683065
Physical Activity				
Activity Entropy	0.4965365	0.2344401	0.5946178	0.2366051
Normalized Activity Entropy	0.1458957	0.07489532	0.1577014	0.06521026
Proportion Stationary Time	0.7067591	0.2435814	0.6844249	0.211789

Table 4

Spearman correlations of baseline anxiety items and momentary anxiety scores in youth and adults.

Questionnaire Item	Momentary Anxiety Correlation Coefficient (*r*, *p*-value)
Youth (SCARED)	
"I get stomachaches at school"	0.19, .03
"I get shaky"	0.24, .006
"I have nightmares about something bad happening to me"	0.28, .001
"I get really frightened for no reason at all"	0.19, .03
"When I get frightened, I feel like I am choking"	0.30, .0006
"I don't like to be away from my family"	-0.22, .01
"I feel shy with people I don't know well"	0.18, .04
"When I get frightened, I feel dizzy"	0.28, .001
Adults (SCAARED)	
"I don't like to be with people I don't know well"	0.15, .03
"I worry about sleeping alone"	-0.21, .04
"I am afraid to be alone in the house"	-0.21, .003
"When I get anxious, I feel like I'm choking"	-0.16, .03
"I don't like to be away from my family"	-0.22, .002
"When I worry a lot, I feel restless"	-0.21, .004

Table 5

Covariate predictor effects of successive models with baseline anxiety as the outcome, covariates as predictors (Model 1), and covariates and mobile sensing features as predictors (Model 2).

	All Participants Included (*b*, *p*-value)		Youth Only (*b*, *p*-value)		Adults Only (*b*, *p*-value)	
Predictor	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
(Intercept)	49.96, < .001	35.25, < .001	51.44, < .001	58.82, < .001	48.97, . < .001	23.69, .079
Sex (Male)	2.64, .271	3.15, .243	-5.84, .162	-9.44, .036	6.43, .039	5.35, .145
Ethnicity (Asian)	-0.81, .768	-1.53, .583	0.51, .889	2.73, .477	-4.47, .332	-2.25, .640
Ethnicity (Black/African)	-3.50, .201	-0.13, .967	3.36, .343	-2.73, .480	-4.76, .203	2.70, .554
Diagnosis (Some Diagnosis(es))	1.68, .443	2.31, .300	3.46, .311	2.39, .518	4.11, .166	3.55, .257
SES (Relatively Lower SES)	2.33, .224	0.81, .680	-0.98, .764	-0.06, .987	3.29, .196	1.33, .618
Sexual Orientation (Non-Straight)	2.67, .249	2.61, .285	7.23, .032	4.19, .260	-2.31, .495	-2.73, .449
Location (Non-Urban)	-2.40, .223	-2.21, .285	1.43, .649	1.17, .731	-2.88, .250	-2.80, .309
Past 90 Day Alcohol Use (Yes)	-3.73, .048	-3.37, .089	-3.92, .180	-2.46, .456	-5.08, .053	-5.71, .174
Phone Type (iOS)	-1.45, .452	-2.98, .365	-6.49, .026	5.02, .422	-0.41, .873	23.69, .079

Table 6

Covariate predictor effects of successive models with momentary anxiety as the outcome, covariates as predictors (Model 1), and covariates and mobile sensing features as predictors (Model 2).

	All Participants Included (*b*, *p*-value)		Youth Only (*b*, *p*-value)		Adults Only (*b*, *p*-value)	
Predictor	Model 1	Model 2	Model 1	Model 2	Model 1	Model 2
(Intercept)	3.96, < .001	4.32, < .001	4.168, < .001	5.337, .006	3.861, < 0.001	3.573, 0.002
Sex (Male)	-0.10, .638	0.26, .266	-0.153, .731	0.232, .663	-0.144, 0.575	0.238, 0.445
Ethnicity (Asian)	-0.47, .042	-0.48, .045	-0.406, .296	-0.566, .227	-0.370, 0.333	-0.125, 0.757
Ethnicity (Black/African)	-0.27, .235	-0.13, .613	-0.353, .350	0.136, .772	-0.166, 0.591	0.194, 0.614
Diagnosis (Some Diagnosis(es))	0.05, .805	0.12, .531	0.423, .248	0.663, .142	-0.059, 0.811	-0.121, 0.646
SES (Relatively Lower SES)	0.02, .896	-0.08, .654	-0.280, .424	-0.837, .064	0.113, 0.592	0.089, 0.693
Sexual Orientation (Non-Straight)	0.29, .134	0.30, .150	0.363, .306	0.560, .214	0.241, 0.391	0.428, 0.159
Location (Non-Urban)	0.01, .955	0.04, .806	0.030, .929	-0.004, .992	0.062, 0.766	-0.054, 0.815
Past 90 Day Alcohol Use (Yes)	-0.29, .068	-0.38, .027	-0.317, .308	-0.208, .602	-0.283, 0.192	-0.330, 0.152
Phone Type (iOS)	-0.06, .712	-0.29, .302	-0.467, .129	-1.777, .022	0.047, 0.823	-0.308, 0.381

Table 7

Between-group measurement invariance testing of a configural, three-factor model of mobile sensing features including mobility, sociability, and physical activity latent factors.

Invariance Testing Step	Model Comparison	*ΔCFI*	*ΔRMSEA*	*ΔSRMR* < .030	$\chi^2(df_{diff})$, p-value	Decision
Metric Invariance	Configural (baseline) versus Metric (factor loadings constrained)	-0.002	-0.002	-0.006	3.63 (6), .73	Accept
Scalar Invariance	Metric versus Scalar (item intercepts constrained)	-0.03	0.03	0.009	72.85 (6), < .001	Reject
Decision Criteria	—	< -.01	< .015	< .030 (metric) < .015 (scalar)	$p < .05$	—

Table 8

Longitudinal invariance testing of a configural, three-factor model of mobile sensing features including mobility, sociability, and physical activity latent factors across two weekly timepoints.

Invariance Testing Step	Model Comparison	*ΔCFI*	*ΔRMSEA*	*ΔSRMR* < .030	$\chi^2(df_{diff})$, p-value	Decision
Metric Invariance	Configural (baseline) versus Metric (factor loadings constrained)	-0.01	0.005	0.02	13.19 (9), .15	Accept
Scalar Invariance	Metric versus Scalar (item intercepts constrained)	0.0008	-0.004	0.001	6.84 (9), .65	Accept
Decision Criteria	—	< -.01	< .015	< .030 (metric) < .015 (scalar)	$p < .05$	—

Table 9

Standardized regression statistics from a partially invariant and an assumed fully invariant structural models of momentary anxiety scores.

Invariance	Fit Indices	Fit Measure	Predictor	Youth *b** [95% CI]	Adults *b** [95% CI]
Partial	*CFI*	0.95	Mobility	0.16 [-0.06, 0.38]	0.08 [-0.03, 0.18]
	RMSEA	0.08	Sociability	0.03 [-0.19, 0.25]	-0.17 [-0.34, 0.01]
	SRMR	0.08	Physical Activity	-0.06 [-0.26, 0.15]	0.00 [-0.16, 0.16]
			Baseline Anxiety	0.02 [-0.19, 0.24]	-0.16 [-0.33, 0.00]
			Baseline Depression	0.02 [-0.26, 0.30]	0.22 [0.03, 0.40]
Full	*CFI*	0.93	Mobility	0.16 [-0.06, 0.38]	0.08 [-0.03, 0.18]
	RMSEA	0.09	Sociability	0.03 [-0.19, 0.25]	-0.17 [-0.34, 0.01]
	SRMR	0.09	Physical Activity	-0.05 [-0.25, 0.15]	0.00 [-0.16, 0.15]
			Baseline Anxiety	0.02 [-0.19, 0.24]	-0.16 [-0.33, 0.00]
			Baseline Depression	0.02 [-0.26, 0.30]	0.22 [0.03, 0.40]

Note. b indicates standardized regression coefficients.*

Table 10

Standardized regression statistics from a partially invariant and an assumed fully invariant structural models of momentary anxiety variances.

Invariance	Fit Indices	Fit Measure	Predictor	Youth *b** [95% CI]	Adults *b** [95% CI]
Partial	*CFI*	0.95	Mobility	0.18 [-0.08, 0.44]	-0.20 [-0.38, -0.02]
	RMSEA	0.08	Sociability	-0.18 [-0.33, -0.04]	0.11 [-0.10, 0.31]
	SRMR	0.09	Physical Activity	0.02 [-0.16, 0.21]	0.10 [-0.05, 0.24]
			Baseline Anxiety	-0.19 [-0.33, -0.04]	-0.03 [-0.23, 0.16]
			Baseline Depression	0.23 [0.05, 0.41]	0.09 [-0.13, 0.31]
Full	*CFI*	0.93	Mobility	0.18 [-0.08, 0.44]	-0.20 [-0.38, -0.02]
	RMSEA	0.09	Sociability	-0.19 [-0.33, -0.04]	0.11 [-0.10, 0.32]
	SRMR	0.09	Physical Activity	0.03 [-0.15, 0.22]	0.10 [-0.05, 0.24]
			Baseline Anxiety	-0.19 [-0.33, -0.04]	-0.03 [-0.23, 0.16]
			Baseline Depression	0.23 [0.05, 0.41]	0.09 [-0.13, 0.31]

Note. b indicates standardized regression coefficients.*

Table 11

Standardized regression statistics from a partially invariant and an assumed fully invariant structural models of baseline anxiety scores.

Invariance	Fit Indices	Fit Measure	Predictor	Youth *b** [95% CI]	Adults *b** [95% CI]
Partial	CFI	0.96	Mobility	0.18 [-0.04, 0.40]	0.05 [-0.05, 0.16]
	RMSEA	0.08	Sociability	-0.04 [-0.20, 0.12]	-0.11 [-0.31, 0.10]
	SRMR	0.08	Physical Activity	0.06 [-0.17, 0.28]	0.07 [-0.10, 0.23]
			Baseline Depression	0.47 [0.33, 0.61]	0.43 [0.30, 0.57]
Full	CFI	0.94	Mobility	0.18 [-0.04, 0.40]	0.05 [-0.05, 0.16]
	RMSEA	0.10	Sociability	-0.04 [-0.20, 0.12]	-0.11 [-0.32, 0.09]
	SRMR	0.09	Physical Activity	0.05 [-0.17, 0.27]	0.07 [-0.10, 0.23]
			Baseline Depression	0.48 [0.34, 0.61]	0.43 [0.30, 0.57]

Note. b indicates standardized regression coefficients.*

Table 12

Standardized regression statistics from a partially invariant and an assumed fully invariant structural models of baseline depression scores.

Invariance	Fit Indices	Fit Measure	Predictor	Youth *b** [95% CI]	Adults *b** [95% CI]
Partial	CFI	0.95	Mobility	0.06 [-0.21, 0.34]	-0.03 [-0.13, 0.08]
	RMSEA	0.09	Sociability	0.10 [-0.06, 0.25]	0.08 [-0.10, 0.26]
	SRMR	0.09	Physical Activity	0.06 [-0.14, 0.26]	0.03 [-0.15, 0.20]
			Baseline Anxiety	0.44 [0.32, 0.57]	0.47 [0.32, 0.61]
Full	CFI	0.94	Mobility	0.06 [-0.21, 0.33]	-0.03 [-0.13, 0.08]
	RMSEA	0.10	Sociability	0.10 [-0.06, 0.25]	0.08 [-0.10, 0.26]
	SRMR	0.09	Physical Activity	0.06 [-0.14, 0.26]	0.03 [-0.15, 0.20]
			Baseline Anxiety	0.45 [0.32, 0.57]	0.47 [0.32, 0.61]

Note. B indicates standardized regression coefficients.*

APPENDIX B: Figures

Figure 1

Spearman correlations of internalizing symptoms and mobility features.

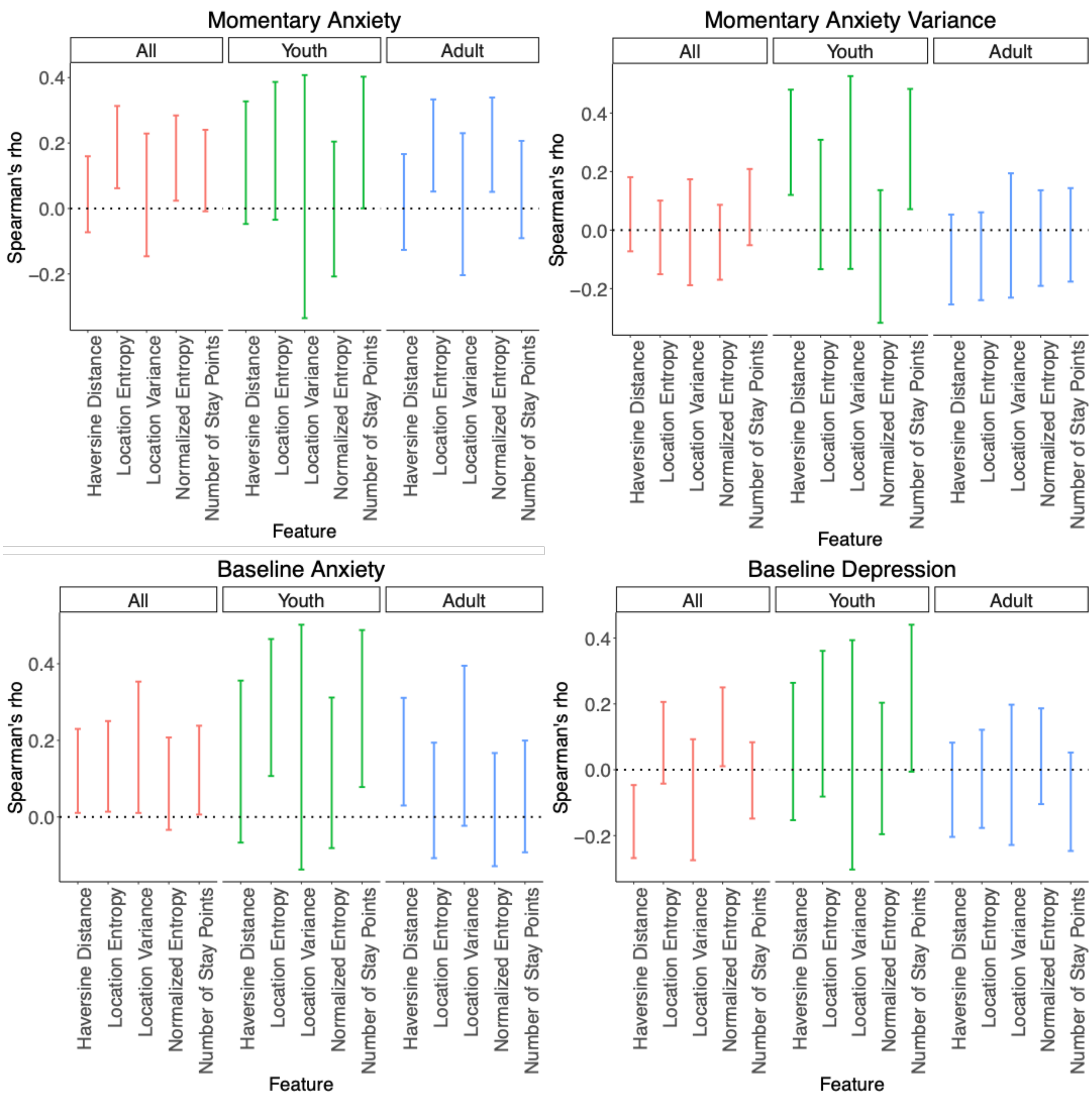

Figure 2

Spearman correlations of internalizing symptoms and sociability features.

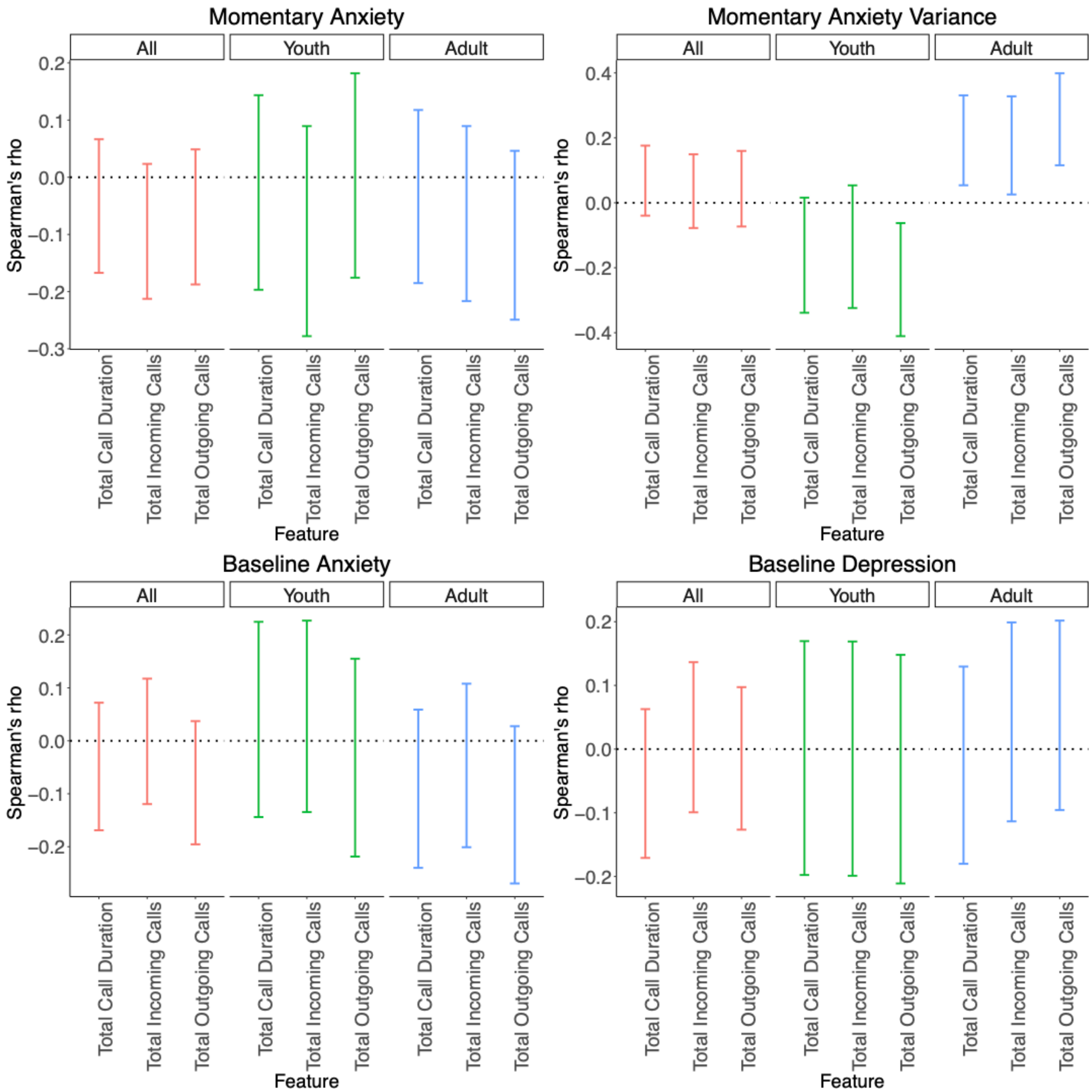

Figure 3

Spearman correlations of internalizing symptoms and screen usage features.

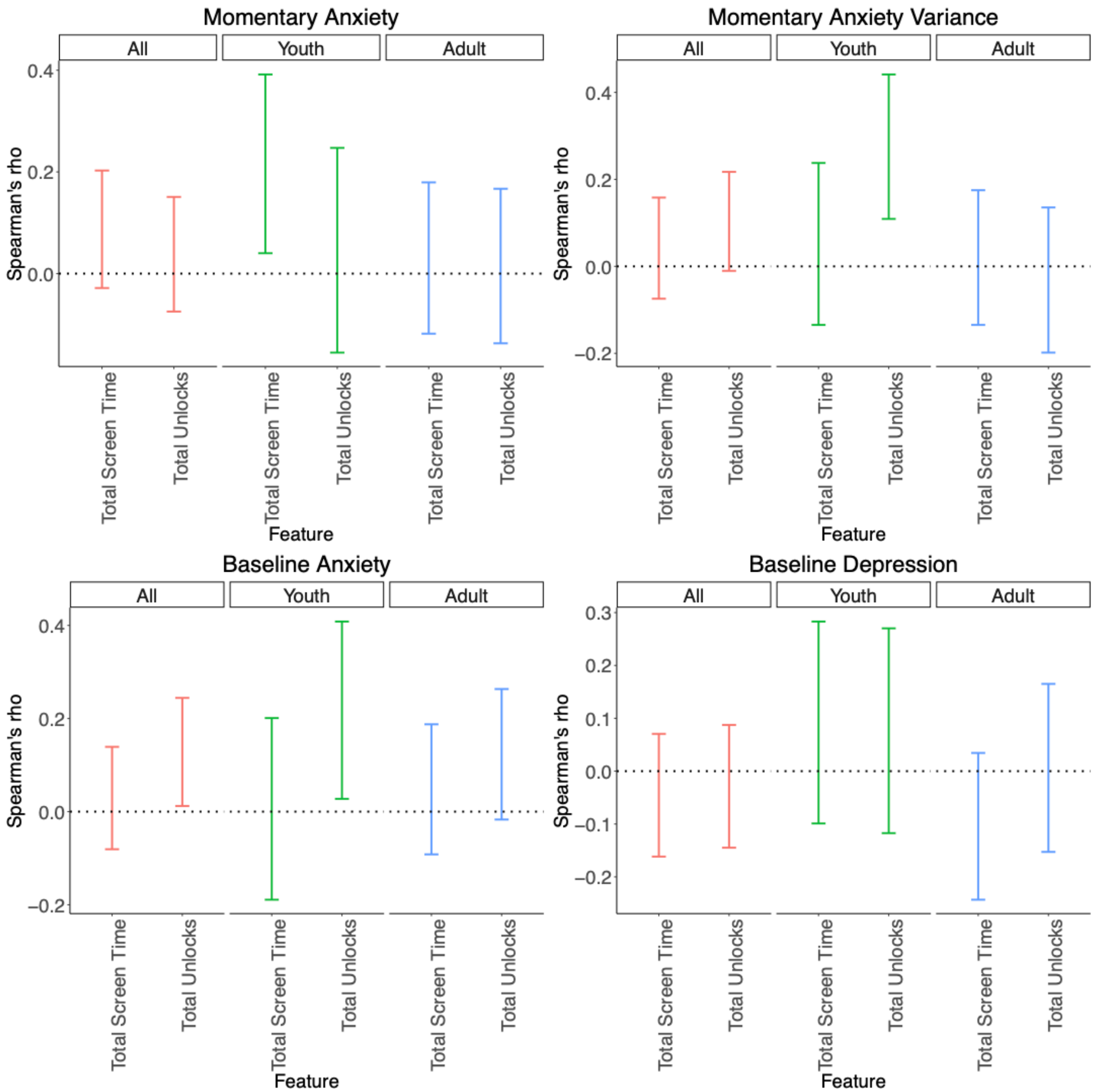

Figure 4

Spearman correlations of internalizing symptoms and sleep features.

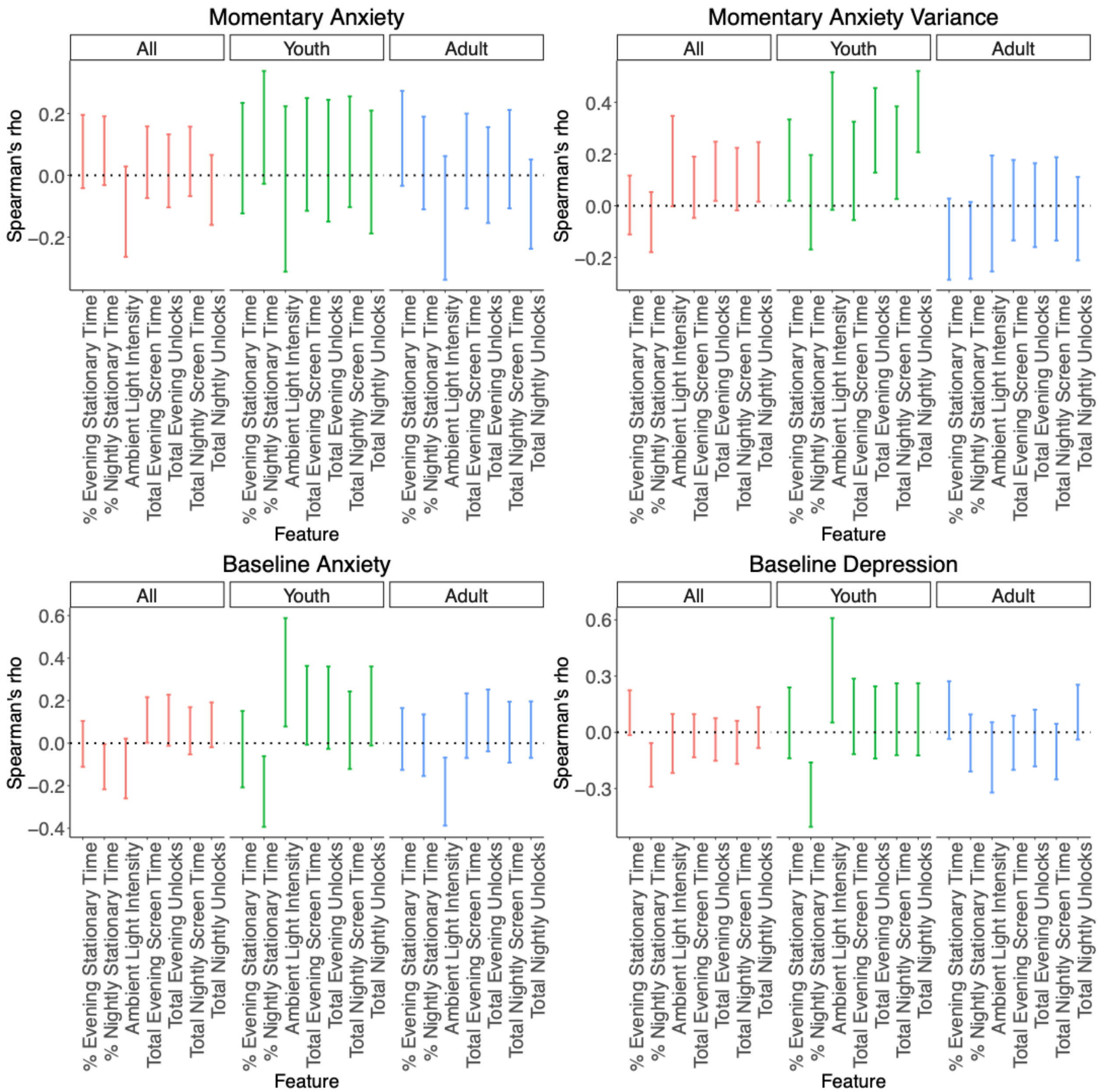

Figure 5

Spearman correlations of internalizing symptoms and physical activity features.

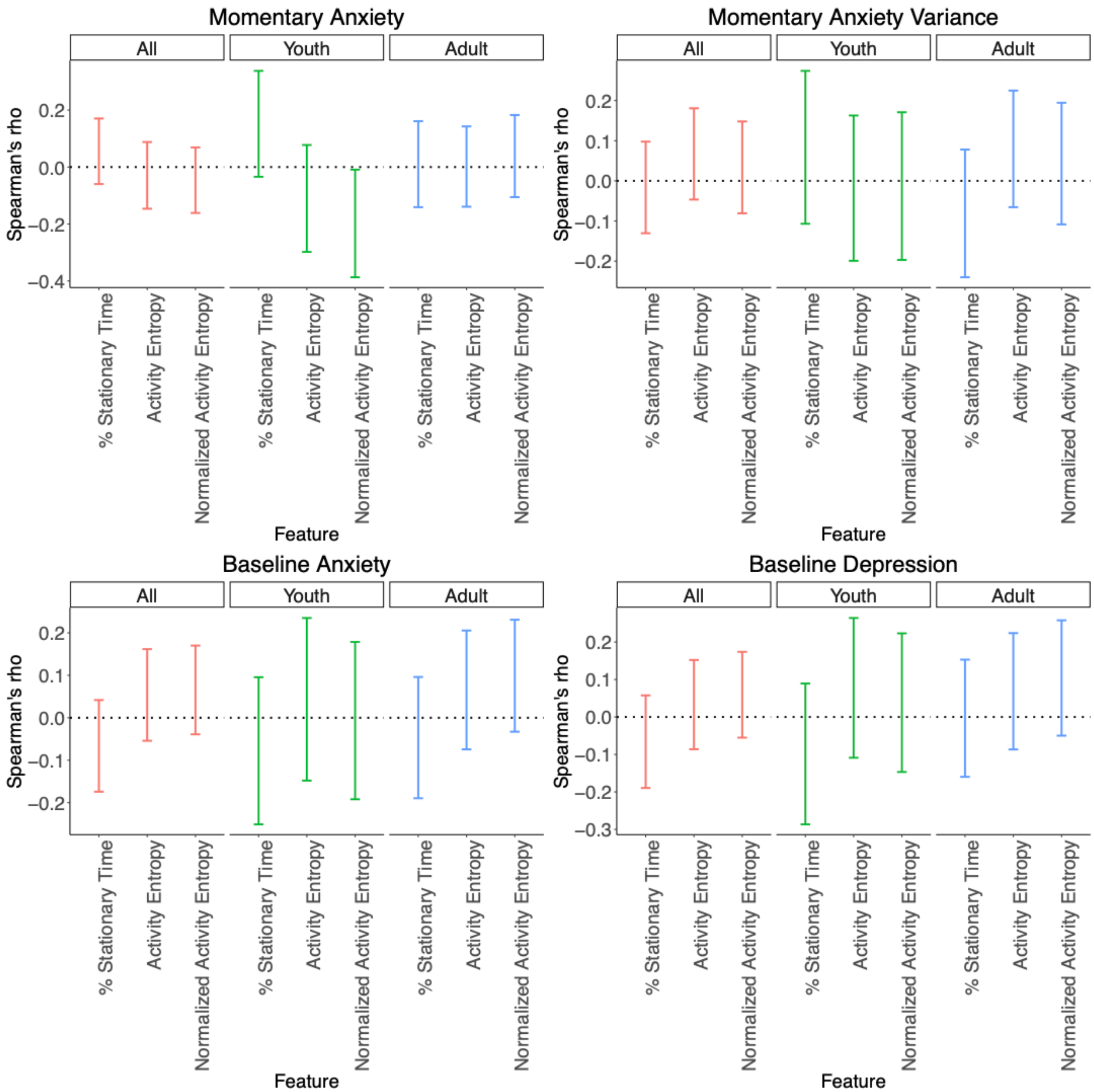

Figure 6

Structural path plot of a confirmatory factor analysis containing PROSIT mobile sensing features, with latent coefficients for youth (lower-most paths) and adults (upper-most paths). Stars denote statistical significance.

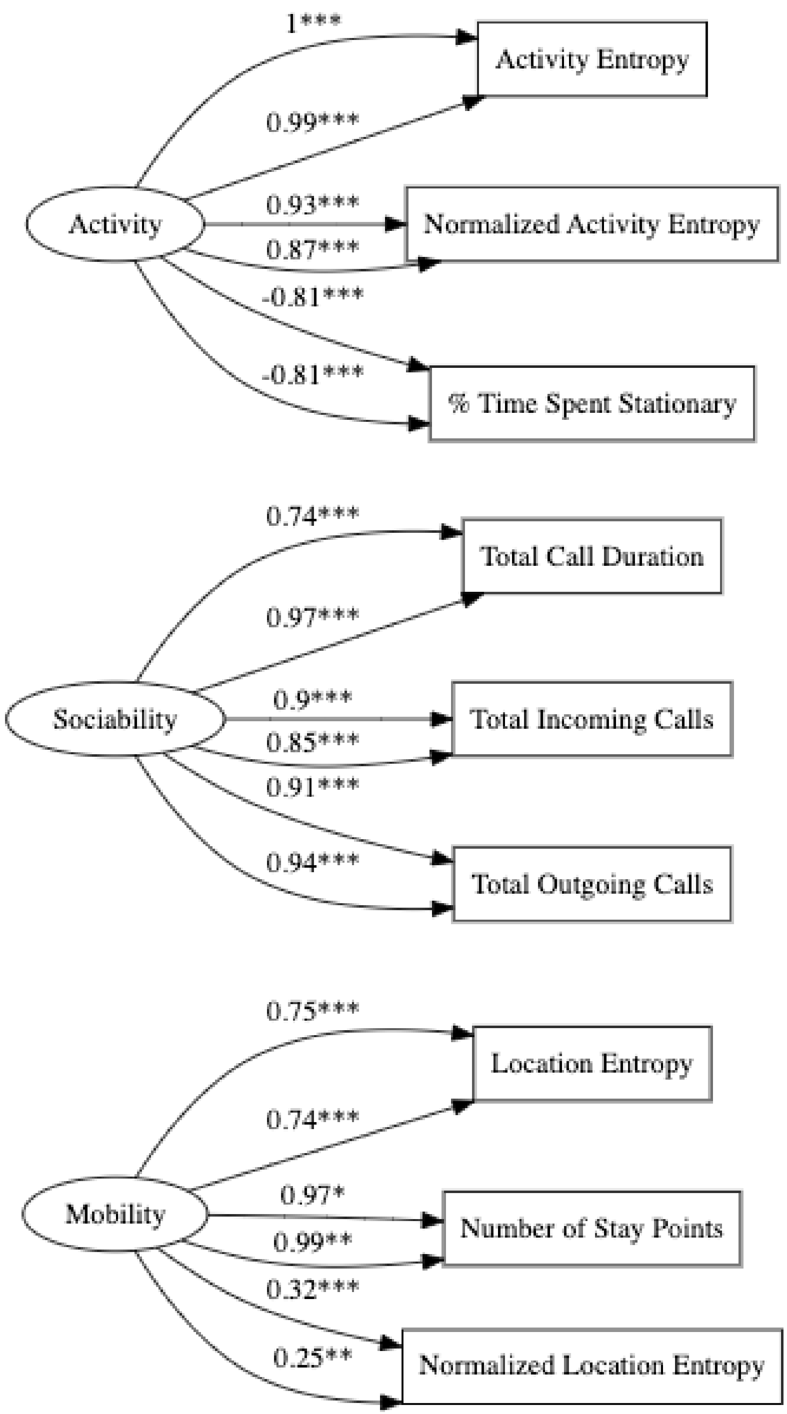

Figure 7

Cross-lagged and lagged effects of mobile sensing features, where the cross-lagged effects run with the first week of feature scores regressed upon the second week of momentary anxiety scores. Missing lines (e.g., for total outgoing calls) indicate model non-convergence.

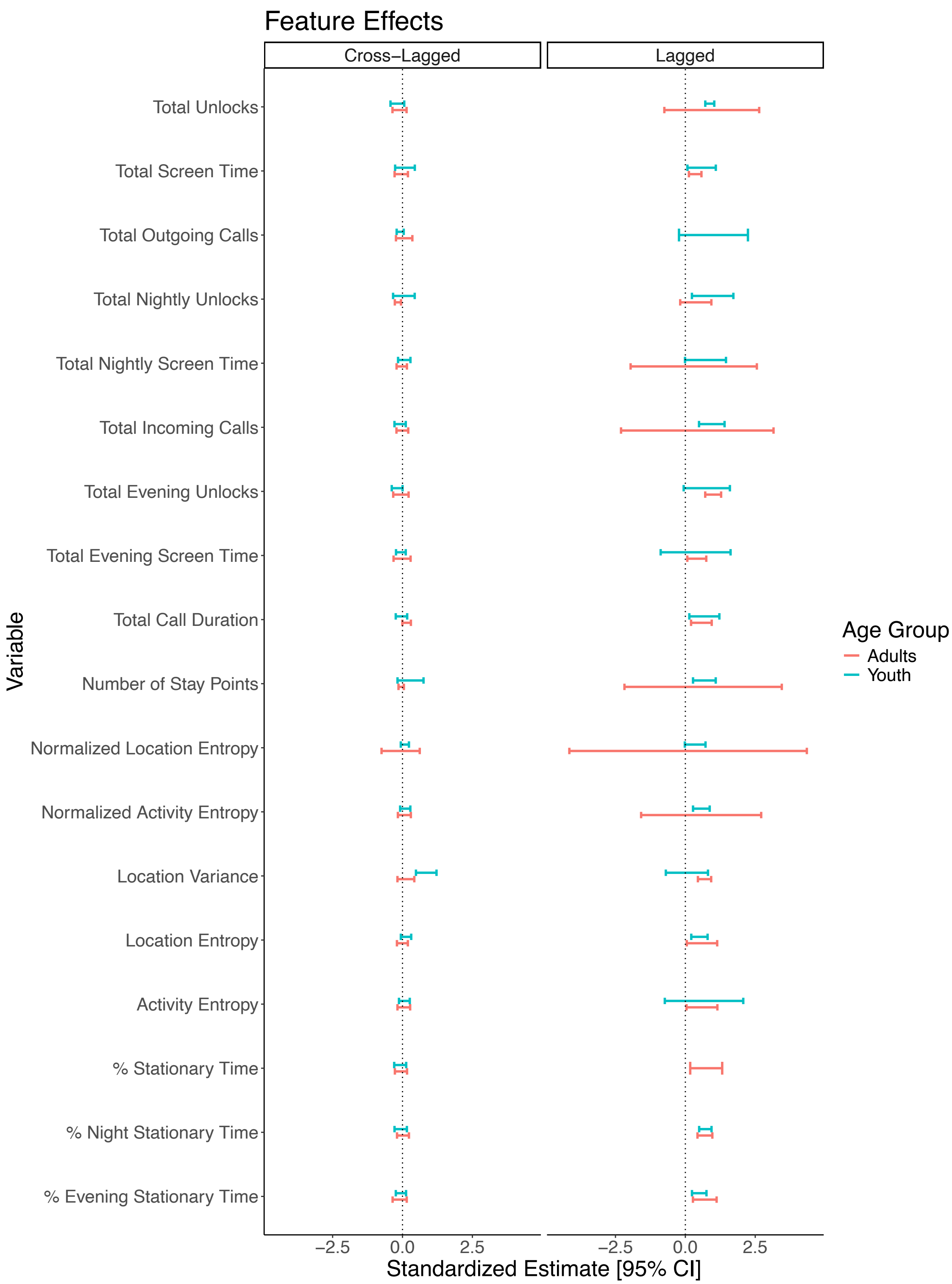

Figure 8

Cross-lagged and lagged effects of mobile sensing features, where the cross-lagged effects run with the first week of momentary anxiety scores regressed upon the second week of feature scores. Missing lines (e.g., for total call duration) indicate model non-convergence.

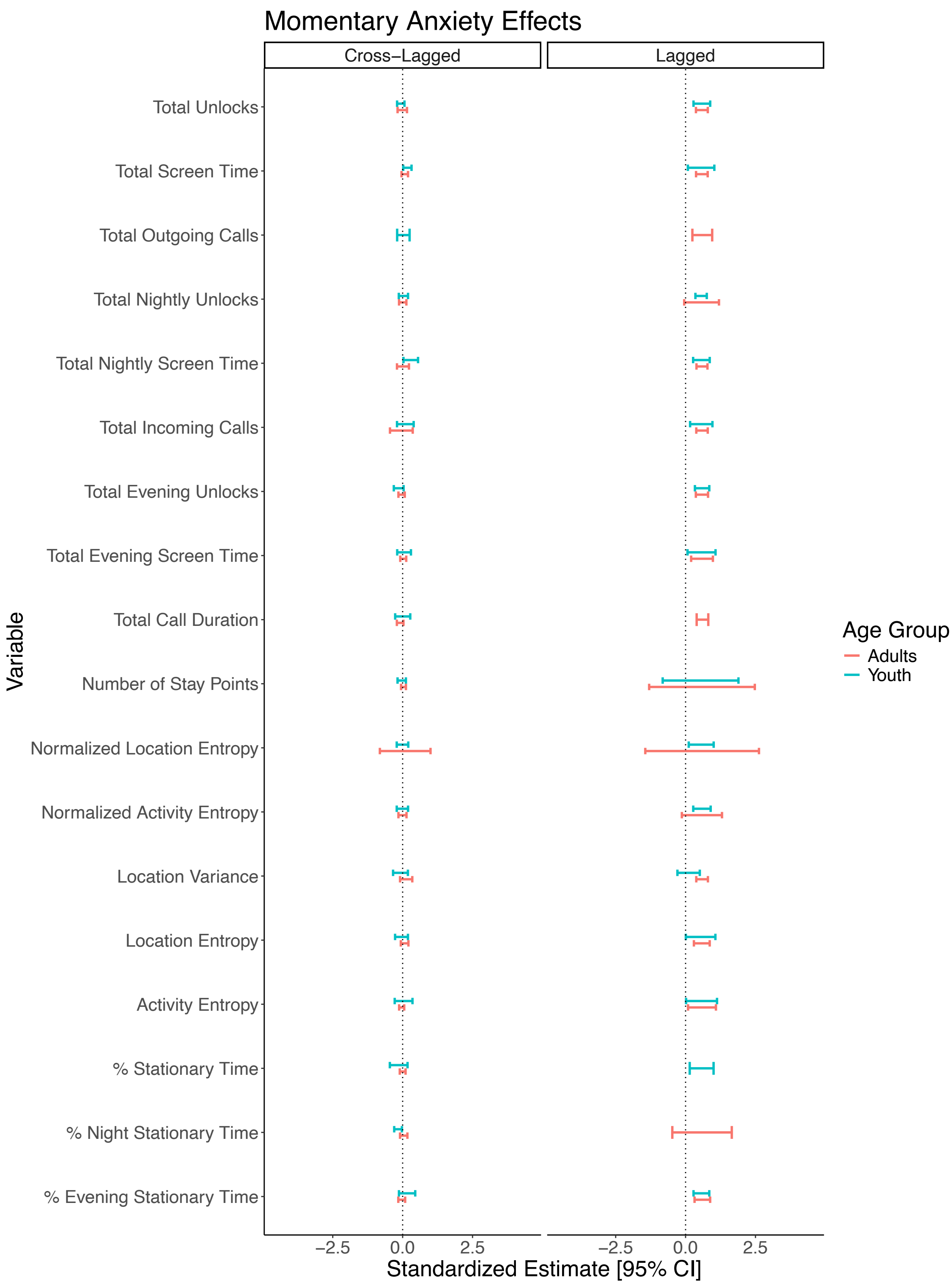

Figure 9

Cross-lagged and lagged effects of mobile sensing features, where the cross-lagged effects run with the first week of feature scores regressed upon the second week of momentary anxiety variances. Missing lines (e.g., for total outgoing calls) indicate model non-convergence.

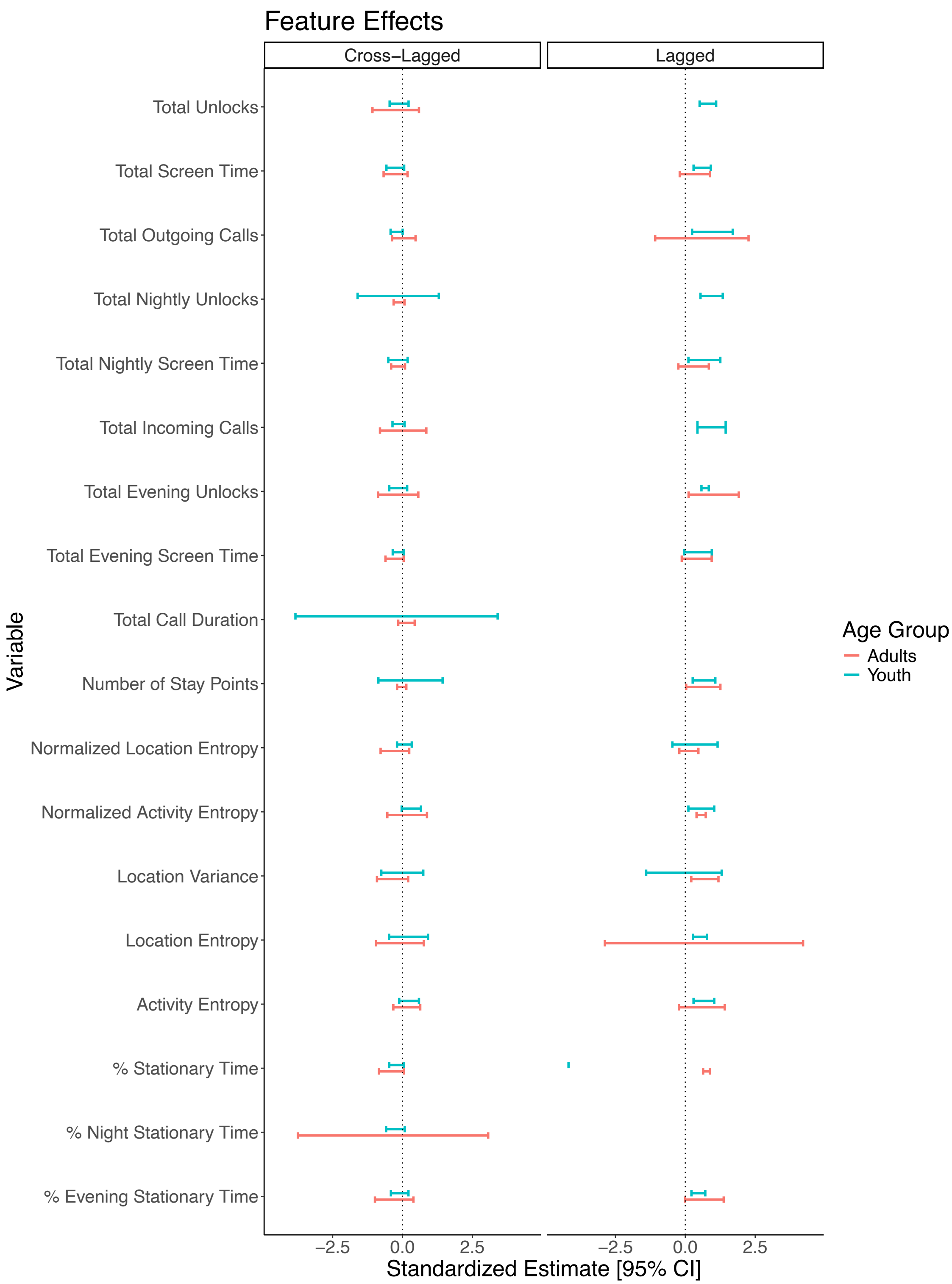

Figure 10

Cross-lagged and lagged effects of mobile sensing features, where the cross-lagged effects run with the first week of momentary anxiety variances regressed upon the second week of feature scores. Missing lines (e.g., for total call duration) indicate model non-convergence.

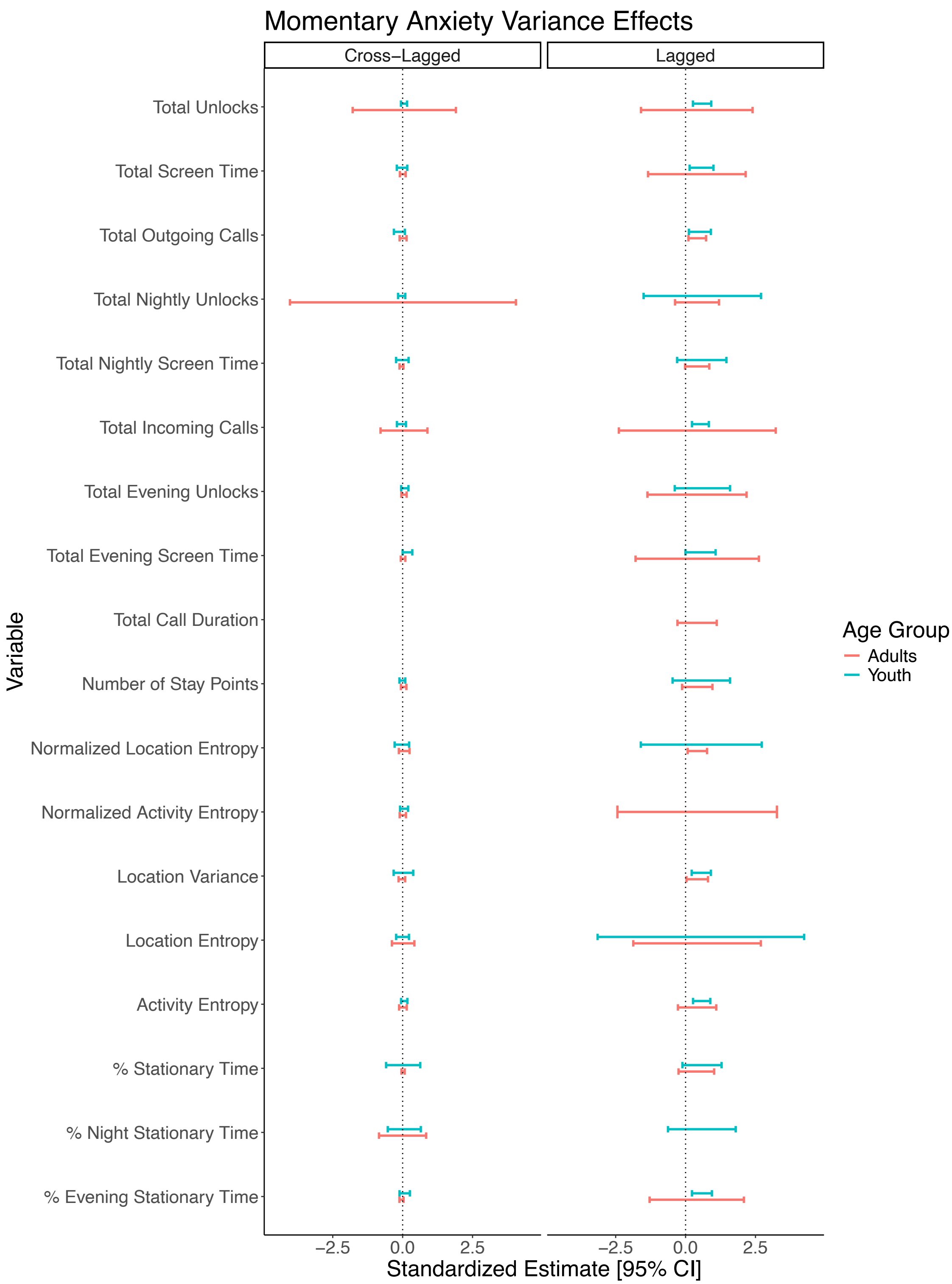

Figure 11

Multilevel effects of mobile sensing features on momentary anxiety scores.

Figure 12

Multilevel effects of mobile sensing features on momentary anxiety variances.

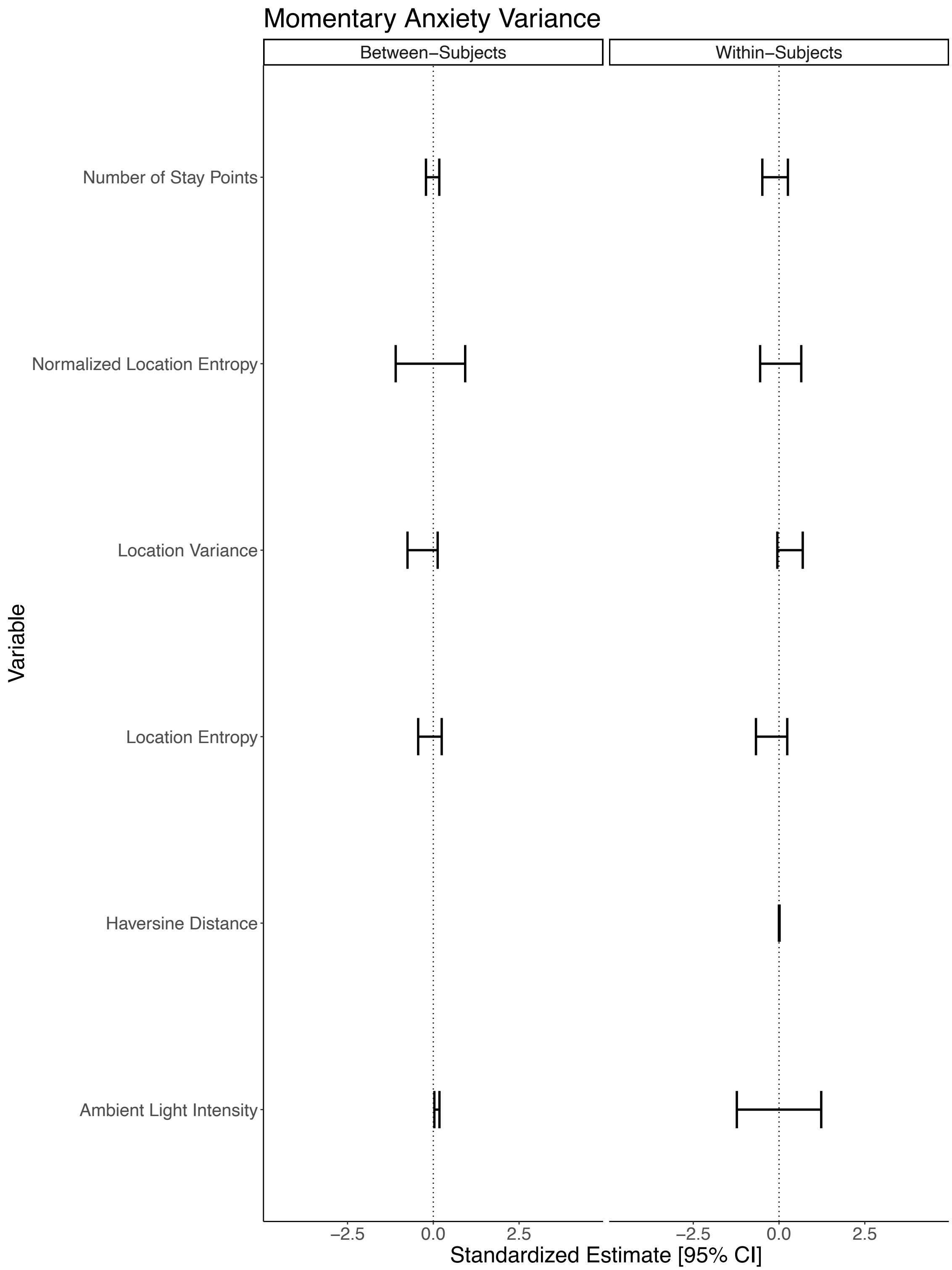

Figure 13

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest regression predicting momentary anxiety scores (across the entire sample of participants).

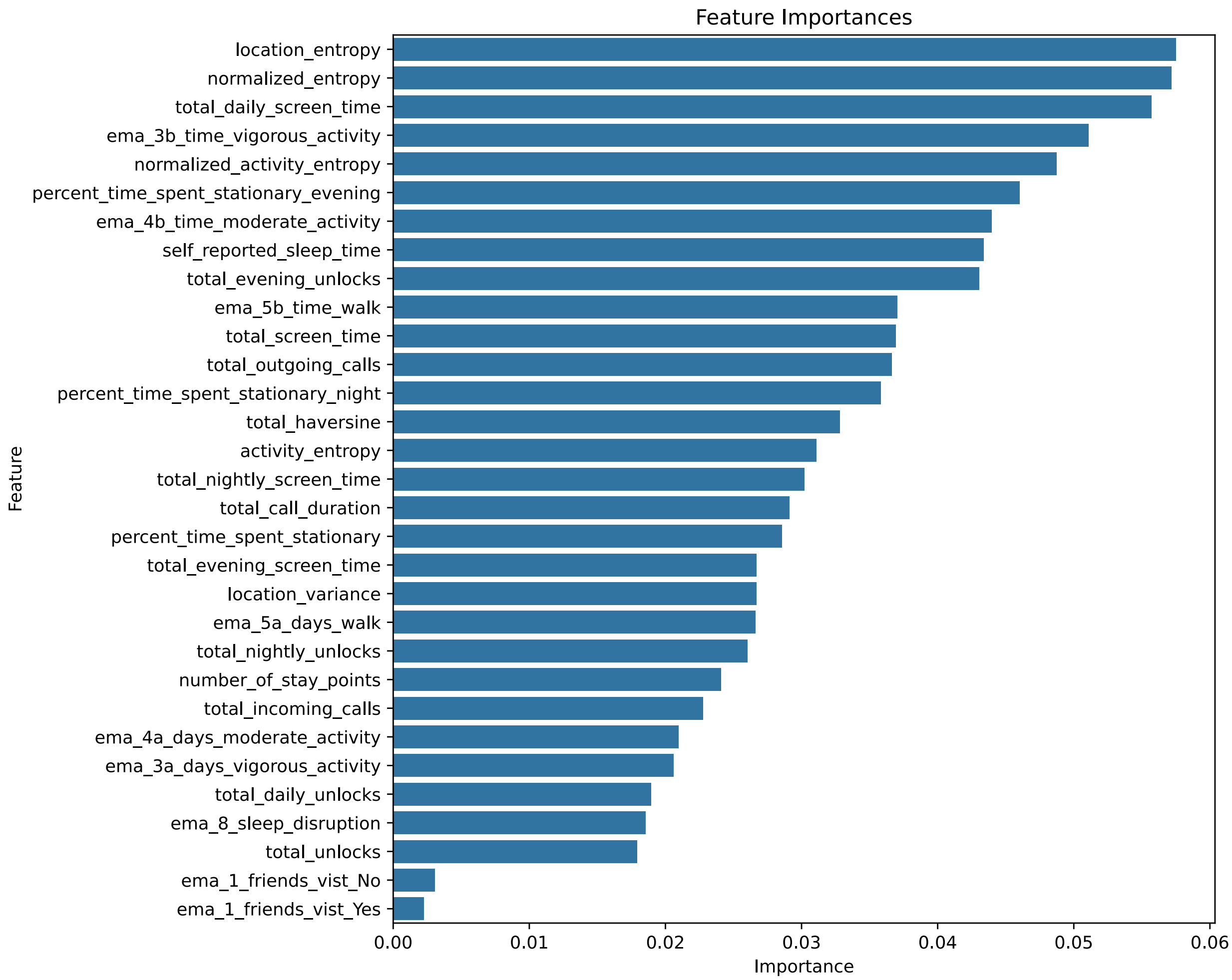

Figure 14

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest regression predicting momentary anxiety variance (across the entire sample of participants).

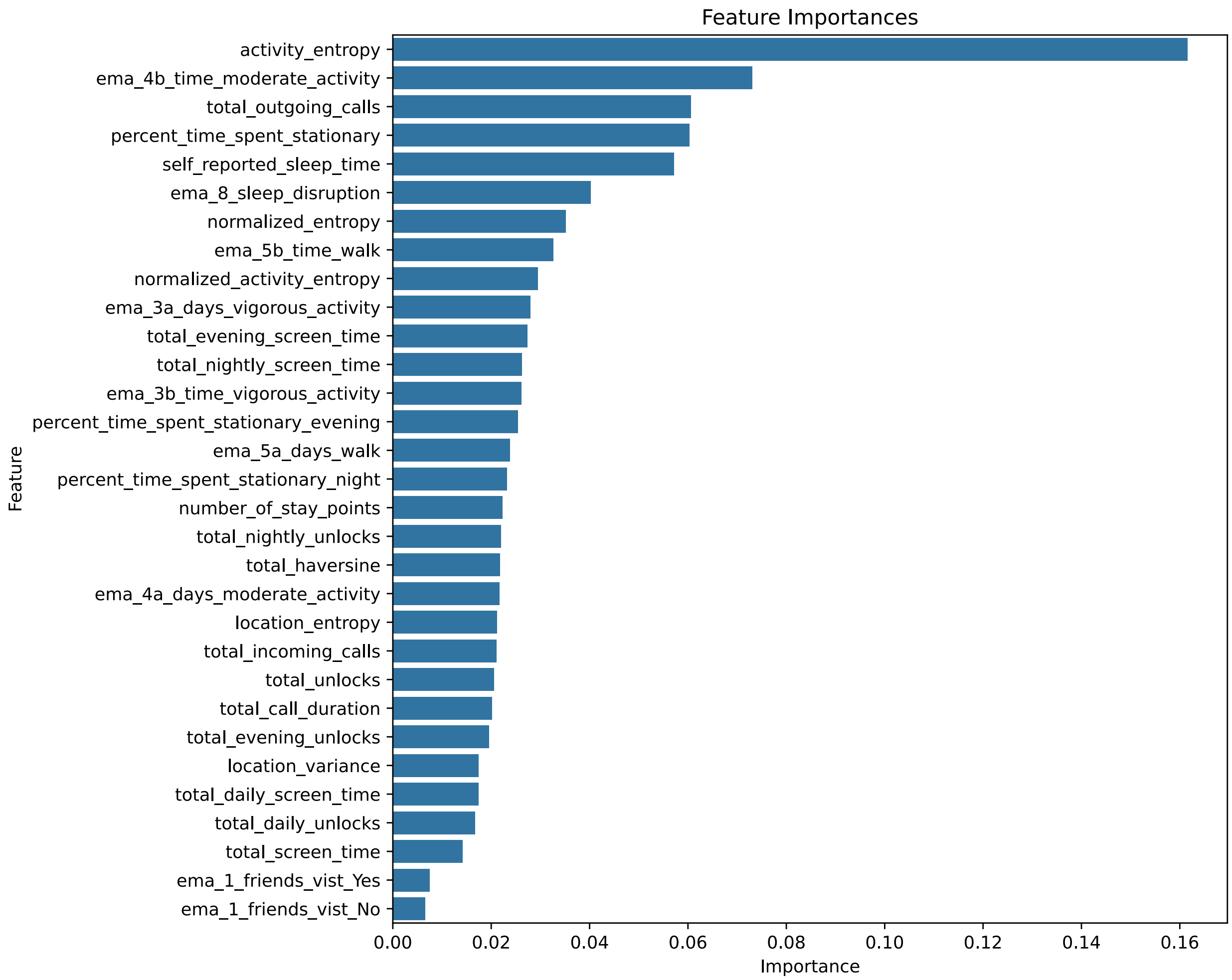

Figure 15

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest regression predicting momentary anxiety scores including only youth participants).

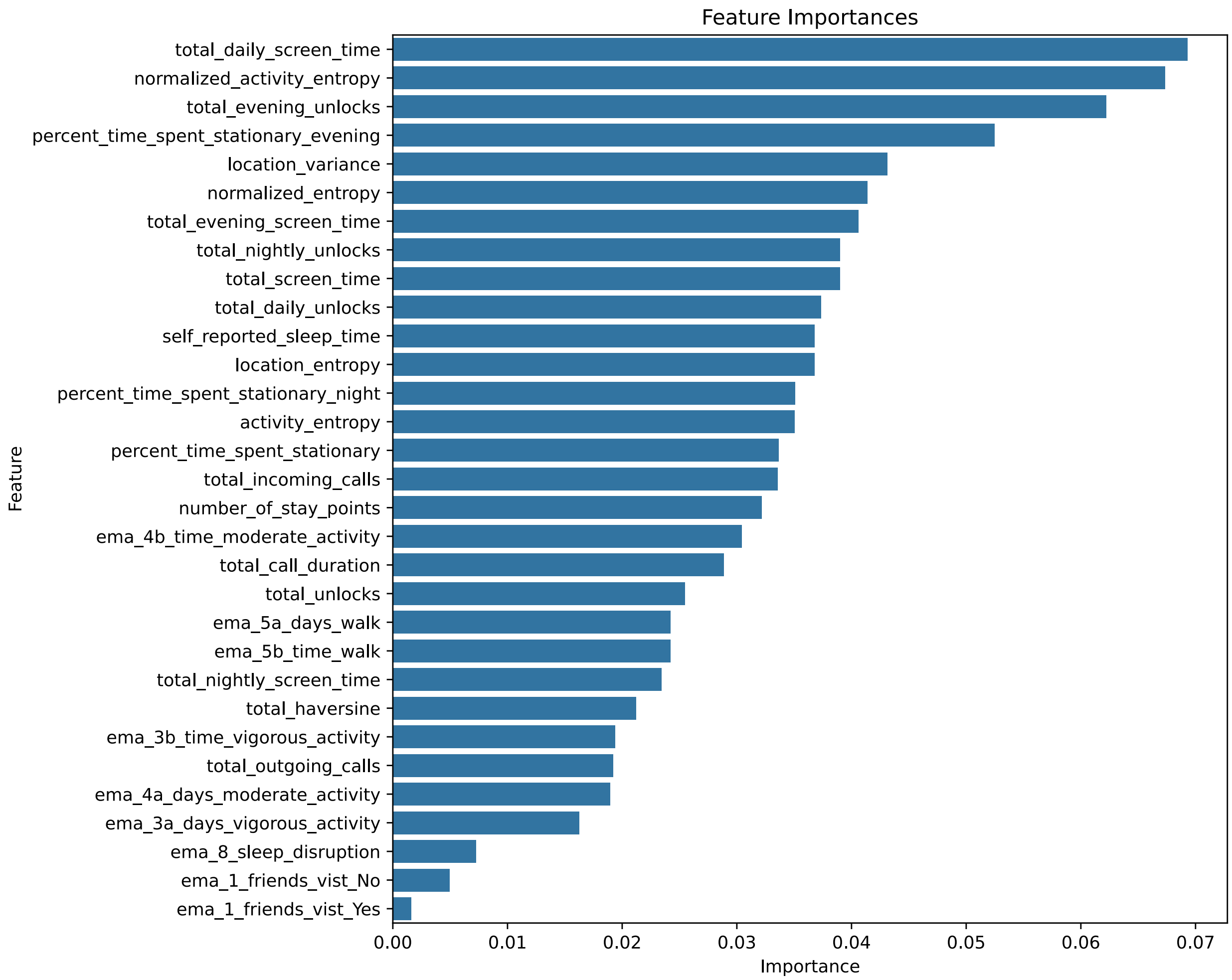

Figure 16

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest regression predicting momentary anxiety scores including only adult participants).

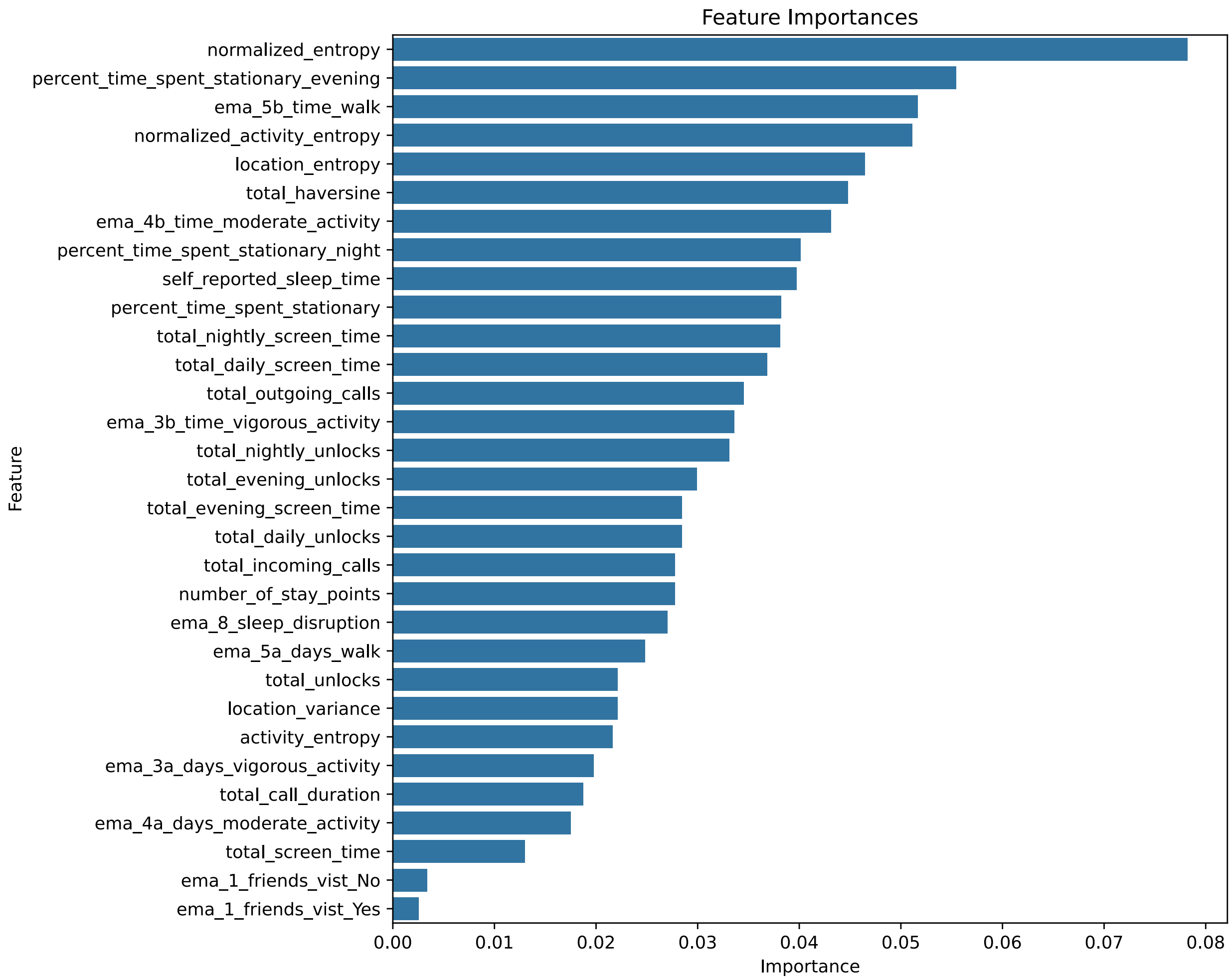

Figure 17

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest regression predicting momentary anxiety variance (including only youth participants).

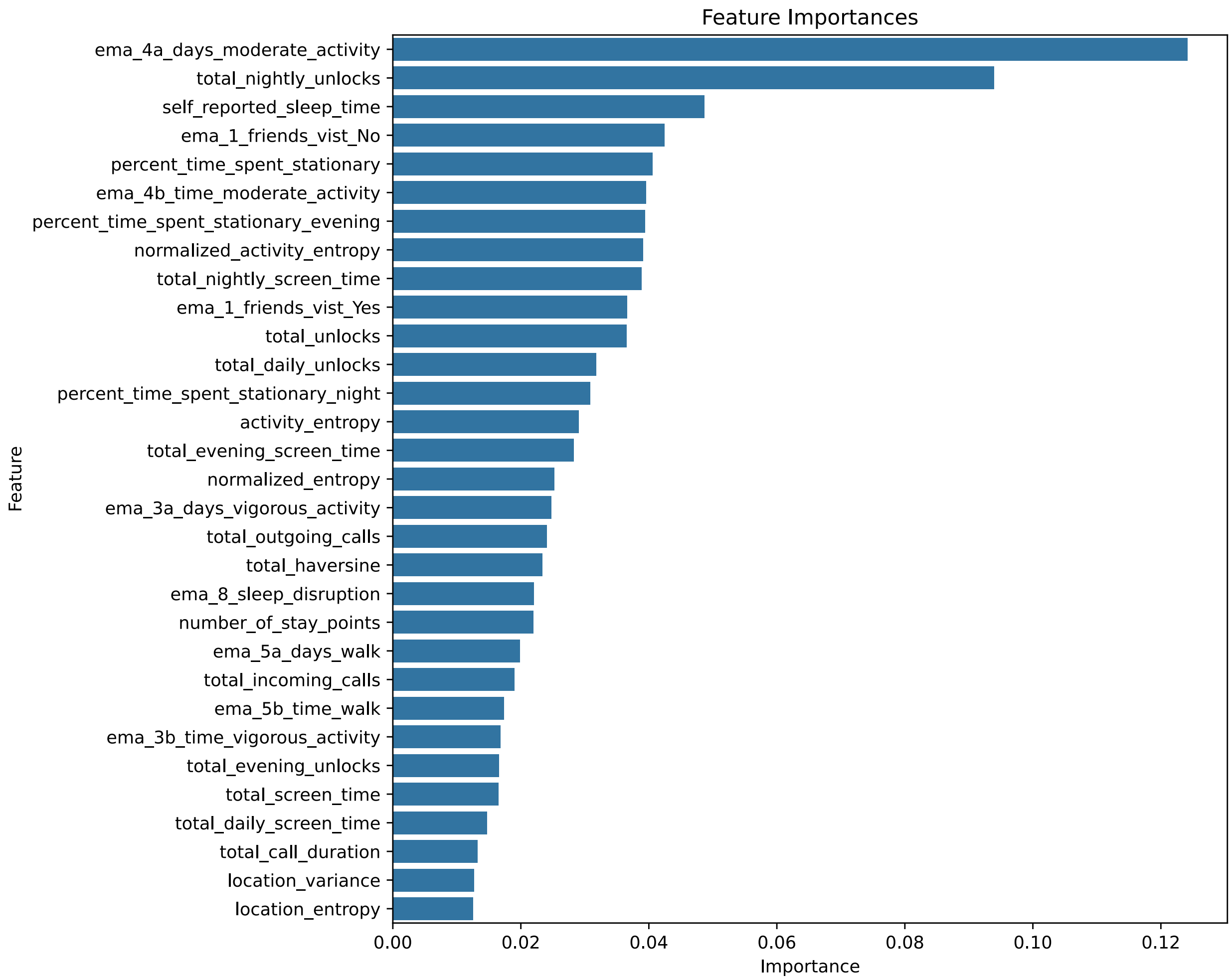

Figure 18

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest regression predicting momentary anxiety variance (including only adult participants).

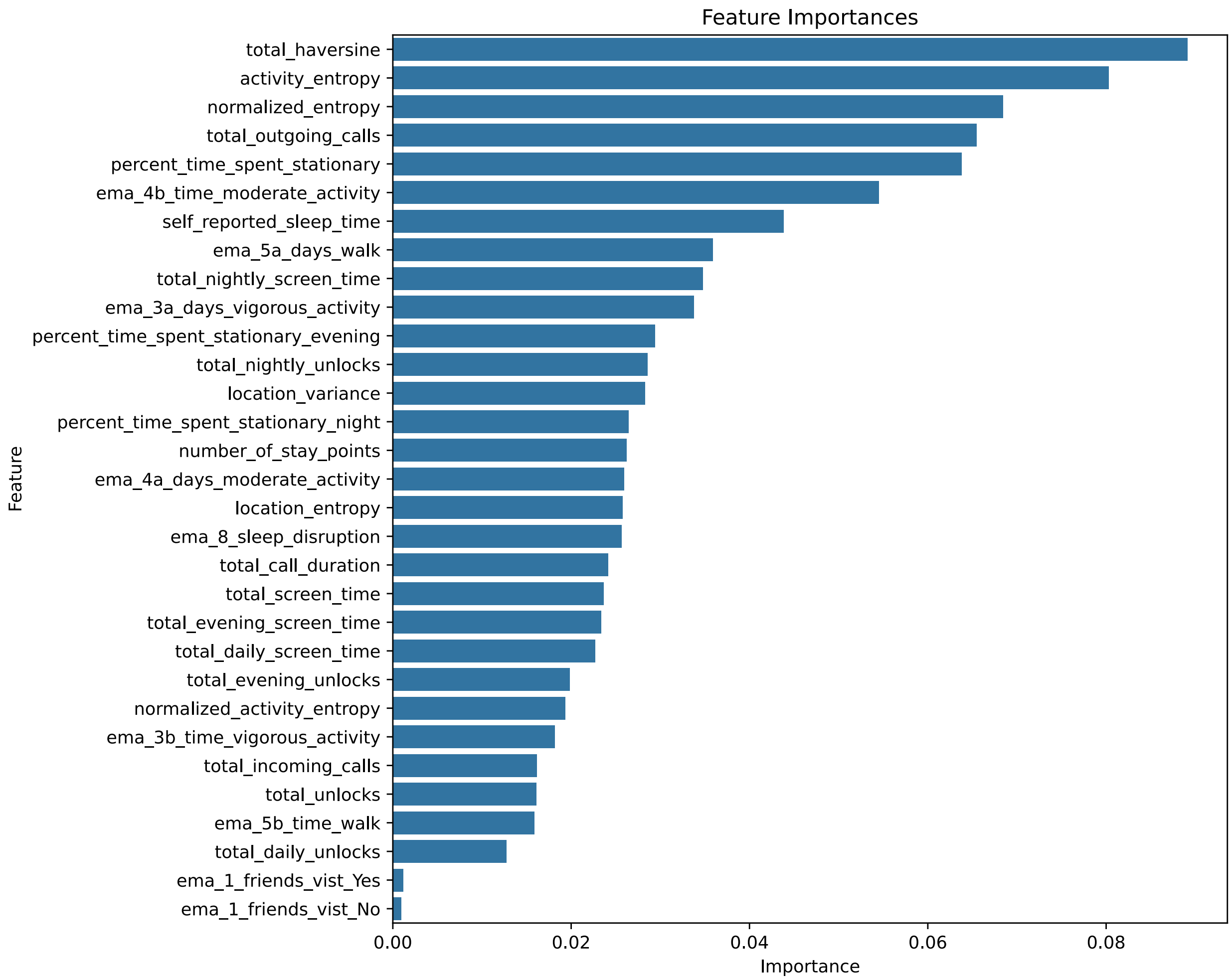

Figure 19

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest recursive feature elimination regression predicting momentary anxiety scores (including all participants).

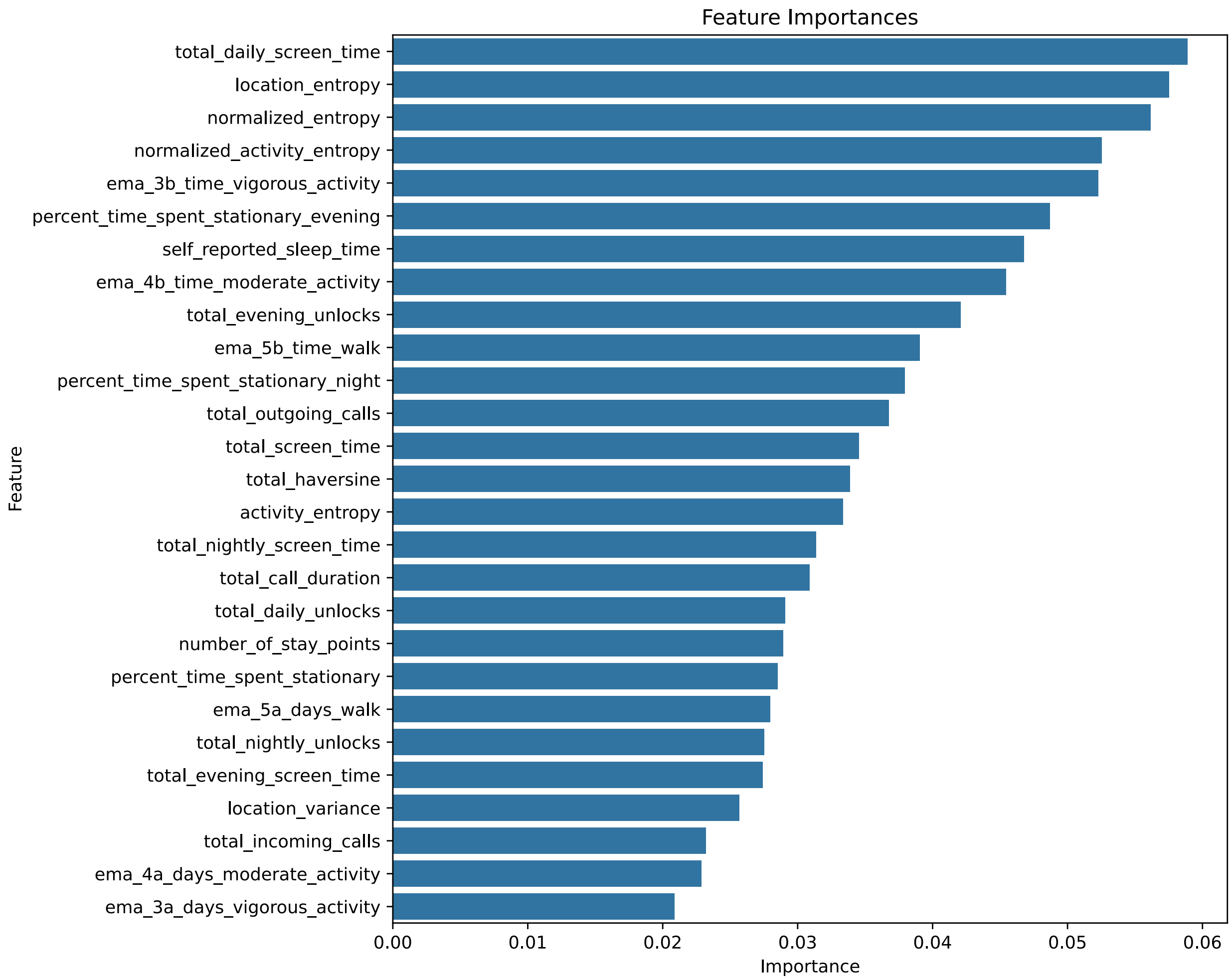

Figure 20

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest recursive feature elimination regression predicting momentary anxiety scores (including only youth participants).

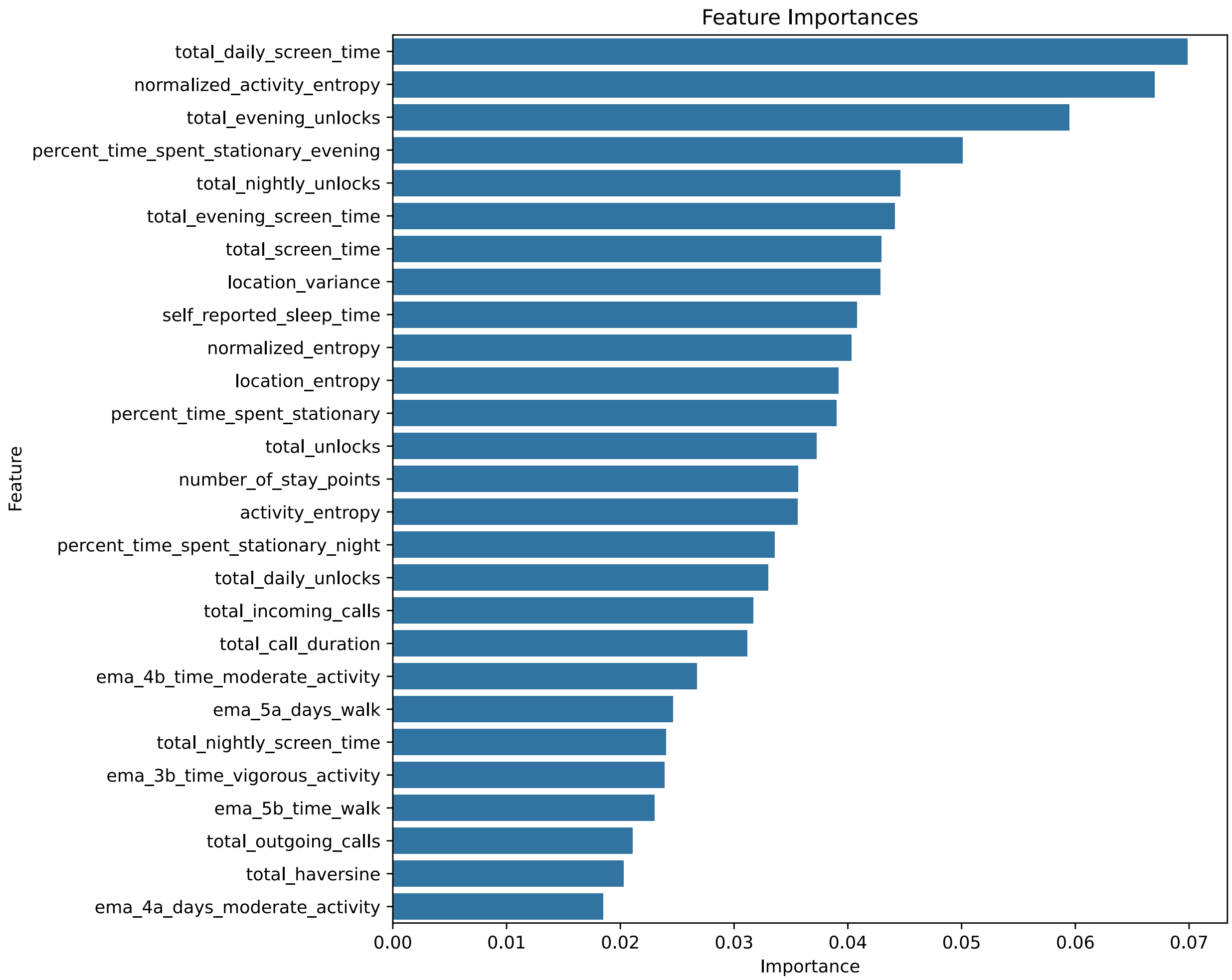

Figure 21

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest recursive feature elimination regression predicting momentary anxiety scores (including only adult participants).

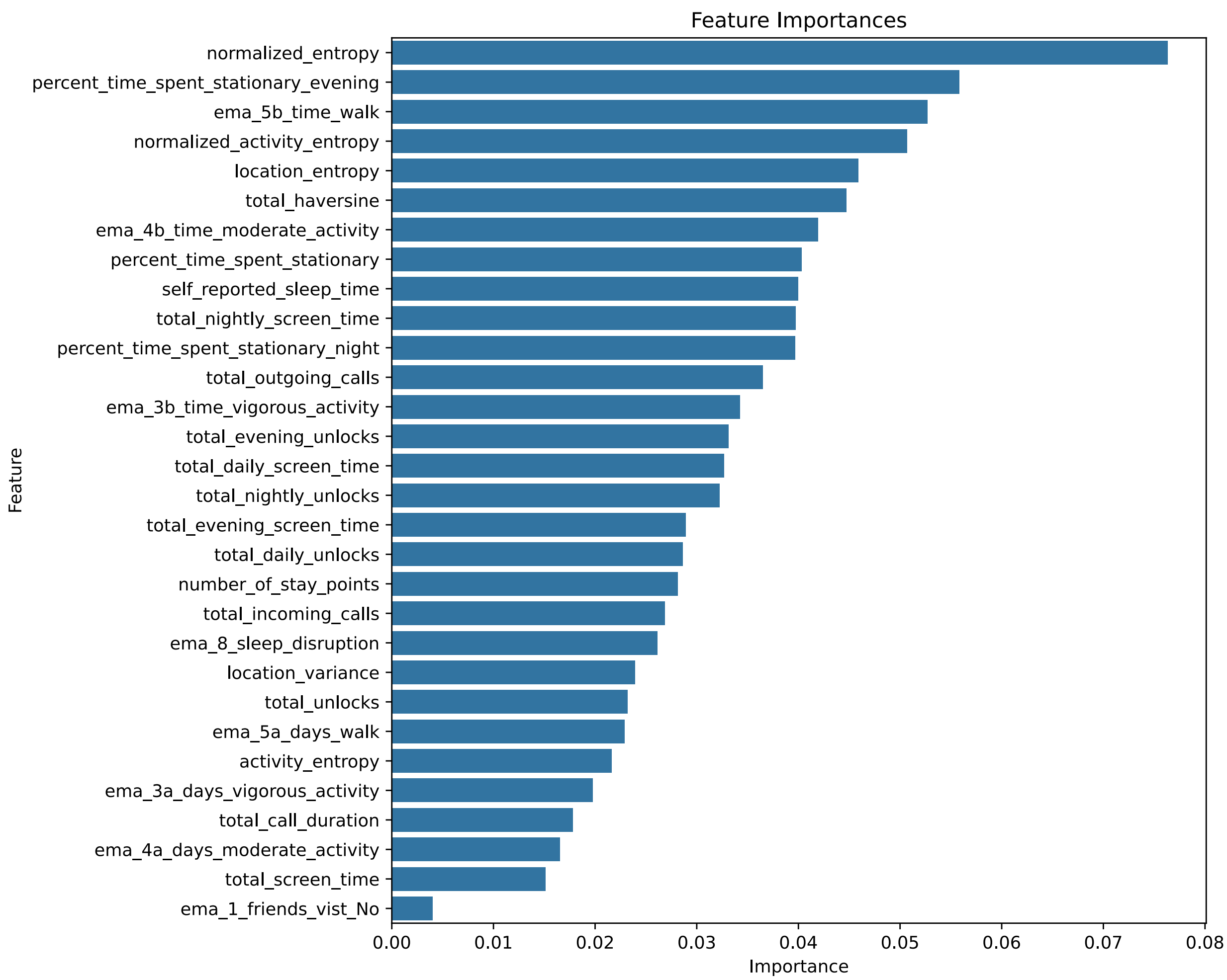

Figure 22

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest recursive feature elimination regression predicting momentary anxiety variance (including only youth participants).

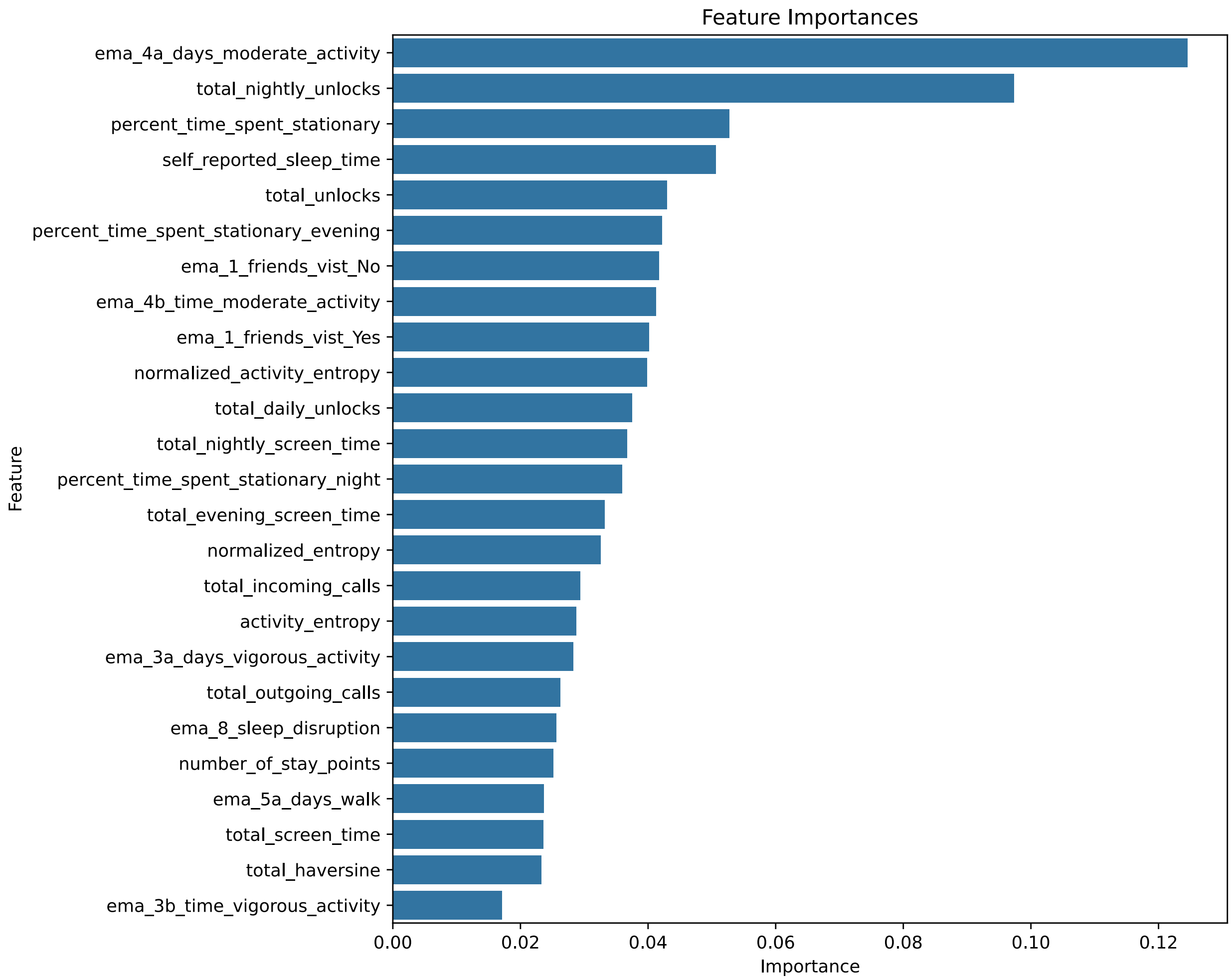

Figure 23

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest recursive feature elimination regression predicting momentary anxiety variance (including only adult participants).

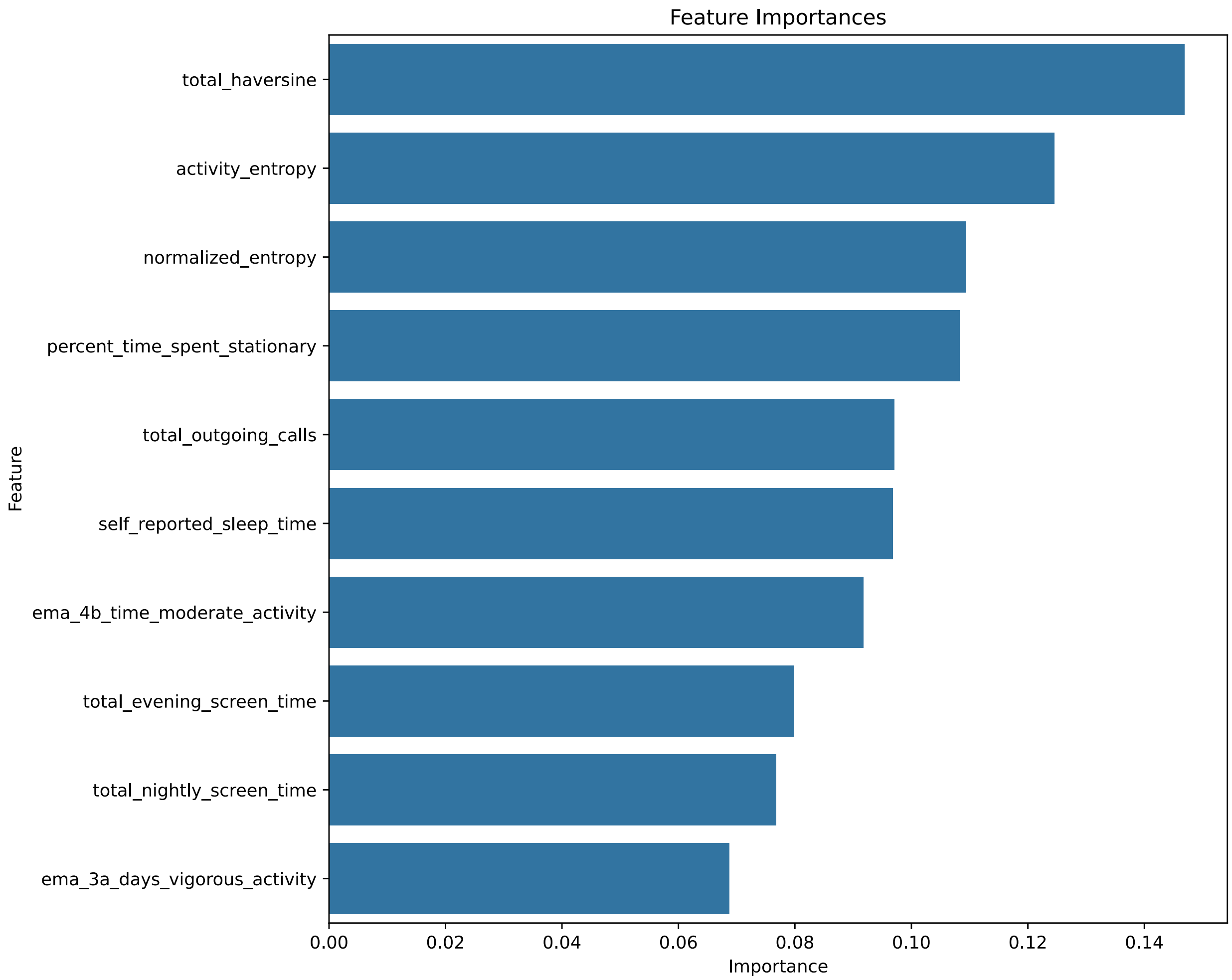

Figure 24

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest recursive feature elimination classifier, classifying participants as either youth or adults while including momentary anxiety variances as a feature.

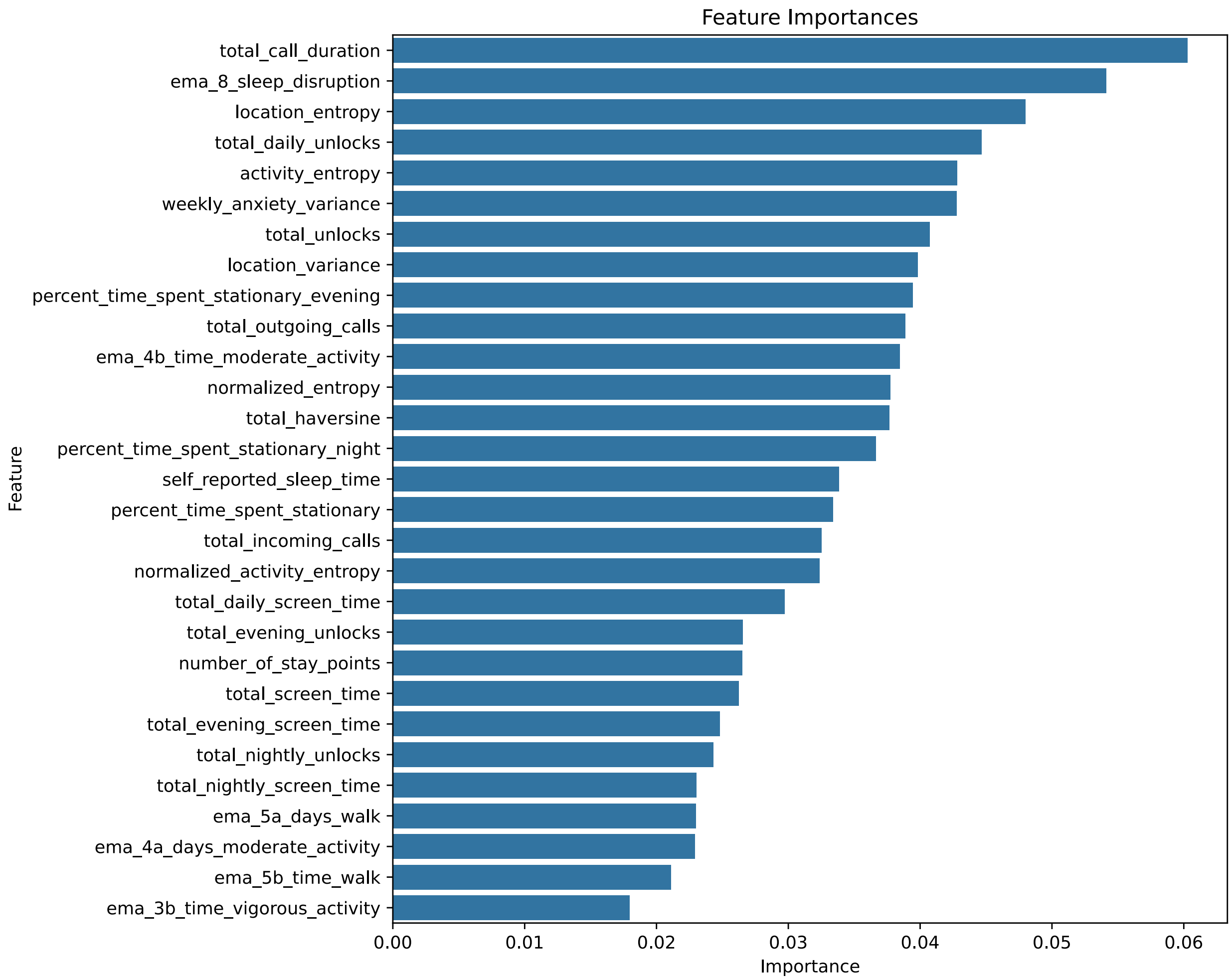

Figure 25

Feature importance plot of mobile sensing features and ecological momentary assessments used in a random forest recursive feature elimination classifier, classifying participants as either youth or adults without including momentary anxiety scores or variances as features.