Square eyes or all lies? Investigating the methodology used in screen use research on children

Booker, Bridget

Booker, Bridget (2024). Square eyes or all lies? Investigating the methodology used in screen use research on children [PhD Thesis]. Australian Catholic University. DOI: https://doi.org/10.26199/acu.91183

SQUARE EYES OR ALL LIES? INVESTIGATING THE METHODOLOGY USED IN SCREEN USE RESEARCH ON CHILDREN

Submitted by

Bridget Booker

Bachelor of Public Health and Exercise Science

Bachelor Applied Public Health (Honors, First Class)

A thesis submitted in total fulfilment of the requirements of the degree of

Doctor of Philosophy

Institute for Positive Psychology and Education

Faculty of Health Sciences

Australian Catholic University

Date Submitted: 12th February 2024

Principal supervisor: Dr Taren Sanders

Co-supervisor: Professor Chris Lonsdale

Associate supervisor: Christine Innes-Hughes

Statement of Authorship

This thesis contains no material that has been extracted in whole or in part from a thesis that I have submitted towards the award of any other degree or diploma in any other tertiary institution. No other person's work has been used without due acknowledgment in the main text of the thesis. All research procedures reported in the thesis received the approval of the relevant Ethics/Safety Committees (where required).

I was involved in all aspects of the presented research and responsible for coordinating the projects under the supervision of Dr Taren Sanders and Professor Chris Lonsdale.

Bridget Booker

Statement of Appreciation

Doing a PhD has been simultaneously the most challenging and enriching experience I have undertaken. Without the encouragement, support, and mentorship from the following wonderful people, the writing of this thesis would not have been possible.

To my mentors and supervisors, your patience and knowledge are admirable. I feel incredibly fortunate to have learned from you over the past four years. Your encouragement and support have been helpful for more than just completing this thesis.

Taren, your passion and excitement for all things research is infectious. Your knowledge and guidance have been invaluable to completing this PhD. I will always be grateful for this opportunity and the potential you saw in me as an undergraduate student completing work experience.

Chris, I know you are a very busy man, so I thank you for your dedication and time. Your feedback has made me a better writer, researcher, and critical thinker. I am grateful for the time you have taken to help me build and hone my skills.

Christine, thank you for your insightful feedback. I know you have held some important roles, especially through the COVID-19 pandemic, so I appreciate you dedicating your time to providing me with valuable insights.

To M&B Lab, thank you for all the laughs, distractions and advice. I could not have made it without you, or at least it would have been much less fun. From now on, I will always reserve rainy days for peppermint hot chocolates.

And last but not least, to my mum, family and friends. Thank you for showing me unwavering support and belief in me over the past four years. I would not be where I am today without you.

I would also like to give a special acknowledgement to my Uncle, Gregory, and my Pop, Robert, who passed away during my candidature. Your unwavering love and support will always be cherished.

The research has been supported by an Australia Government Research Training Program Scholarship.

Table of Contents

List of Tables

List of Figures

Abstract

Screen-based media devices, such as televisions, smartphones, and tablets, are integral to modern living. Children today are spending more time on screen-based media devices than ever before. However, the influence of these devices on children's health outcomes remains controversial, with research showing inconsistent findings. Much of the inconsistency between studies may be due to the way screen use has been measured, as the field has largely relied on unvalidated self- or parent-reported measures. Automated wearable cameras present an opportunity to assess children's screen use more accurately. The primary purpose of this thesis was to examine the use of automated wearable cameras to establish a more accurate measure of children's screen use. The secondary purpose was to examine whether estimates of screen use from device-based measurements are similar to self-report measurements. Study 1 (Chapter 2) was a systematic review of the literature providing an overview of the evidence on the use of automated wearable cameras to measure health behaviours in youth. This study found that automated wearable cameras may provide a reliable method for measuring specific health behaviours; however, there was limited evidence on the validity of automated wearable camera measurements. Study 2 (Chapter 4) investigated the convergent validity of automated wearable camera measurements for measuring children's screen use compared to direct observation. The findings from this study demonstrated that automated wearable camera measurements of screen use duration, type of device, social environment, content, associated behaviours, such as multitasking and eating, and location of the screen use show excellent agreement and strong convergent validity with direct observation measurements but poor agreement and weak convergent validity for food-related behaviours. Study 3 (Chapter 5) investigated whether estimates of screen use from the automated wearable camera measurements were similar to a self-report measure of screen use. The study found that parent- and child-reported screen use measurements were inaccurate compared to automated

wearable camera measurements, with children and parents being more likely to overestimate their children's screen use. The thesis's implications, strengths, and limitations are discussed in Chapter 6. Overall, the findings of this thesis raise concerns about the accuracy of child- and parent-reported measurements of screen use. Given the challenge of accurately measuring children's screen use behaviours; automated wearable cameras have the potential for accurately measuring complex screen use behaviours such as the content and context of the behaviour.

Chapter 1: Introduction and Overview

Over the past decade, screen-based media devices, such as televisions, tablets, and smartphones, have profoundly changed childhood. The current generation of children and young people have been described as ‘digital natives’, who have grown up surrounded by screen-based media devices and screen media (Bittman et al., 2011). These devices have become an integral part of children’s daily lives and have transformed how children learn, socialise, express, entertain themselves and interact with the world around them (Brown & Bobkowski, 2011). However, the influence of these devices on children’s health and behaviours remains a controversial topic, with the current research showing inconsistent findings (Cain & Gradisar, 2010; Hale & Guan, 2015). Much of the inconsistency between studies may be due to the way screen use has been measured, as the field relies largely on unvalidated self- or parent-reported measurements (Cain & Gradisar, 2010; Hale & Guan, 2015). Automated wearable cameras present an opportunity to assess children’s screen use more accurately.

My primary objective in this thesis was to examine the use of automated wearable cameras to establish a more accurate measurement of children’s screen use. Furthermore, I examined whether estimates of screen use from device-based measurements are similar to self-report measurements.

Screen Use

What is Screen Use and Screen Time?

Within the literature, screen use and screen time definitions vary and are often used interchangeably. The lack of consensus within the literature presents numerous challenges relating to the measurement and comparison of screen use research. The differentiation of definitions and conceptualisations has led to many different approaches when measuring screen use. In this thesis, I followed recommendations from Kaye et al. (2020), where “screen use" is conceptualised as goal-directed behaviour and "screen time" is conceptualised as a numerical measurement. Screen use refers to the goal-directed behaviours that screen-based media devices facilitate (Kaye et al., 2020). For example, screen use consists of using a screen-based media device to fulfil different motivations, such as using a computer for homework (Kaye et al., 2020). Screen time refers to time spent on screen-based media devices (Tremblay et al., 2017). For example, screen time is the duration spent watching a television programme. Screen time is further split into four subtypes: recreational screen time, sedentary screen time, stationary screen time and active screen time.

Recreational Screen Time

Recreational screen time refers to time spent using screen-based media devices that does not involve school or work (Tremblay et al., 2017). For example, this may include watching cartoons and playing stationary video games.

Sedentary Screen Time

Sedentary screen time refers to time spent using screen-based media devices while in a sitting, lying down or reclining posture (Tremblay et al., 2017). Sedentary screen time is characterised by having a low energy expenditure (≤1.5 metabolic equivalents; Tremblay et al., 2017). For example, this may include sitting or lying down and watching television.

Stationary Screen Time

Stationary screen time refers to time spent using a screen-based media device while standing lying, or sitting, with no ambulation, regardless of energy expenditure (Tremblay et al., 2017). For example, this may include using a tablet while sitting down or using a mobile phone while standing in a line. Stationary and sedentary screen time definitions overlap in some respects but differ according to energy expenditure. Sedentary screen time is defined as an individual's energy expenditure while using screen-based media devices. The definition of stationary screen time does not consider an individual's energy expenditure while using screen-based media devices.

Active Screen Time

Active screen time refers to time spent using screen-based media devices with movement (Tremblay et al., 2017). For example, this includes time spent playing active video games such as exergaming.

When considering the content of the screen-based media device, screen time can be further divided into two additional subtypes: passive screen time and interactive screen time.

Passive Screen Time

Passive screen time includes passively receiving screen-based media information (Sweetser et al., 2012). This type of screen time requires no input or interaction during the exposure, with the viewer only receiving information. For example, this may include watching a television.

Interactive Screen Time

Interactive screen time includes cognitively engaging in screen-based media activities (Sweetser et al., 2012). This type of screen time requires real-time input from the user during the screen exposure. For example, this may include playing video games or completing homework.

In summary, the definitions of screen use and screen time vary. Screen use is conceptualised as goal-directed behaviour, while screen time refers to a numerical measurement (Kaye et al., 2020). Furthermore, screen time can be split into four subtypes: recreational screen time, sedentary screen time, stationary screen time and active screen time. Screen time can also be further split into two additional subtypes when considering the content of the screen-based media device: passive screen time and interactive screen time. Most research in this thesis refers to recreational and sedentary screen time. In some cases, the research presented refers to the broader "screen time" category, not the subgroup classification. I have referred to the appropriate classifications of screen time wherever possible. Table 1.1 presents a summary of screen use and screen time terms and definitions.

Table 1.1

Summary of Screen Use and Screen Time Definitions

Term	Definition
Screen Use	Goal-directed behaviours that screen-based media devices are facilitating (Kaye et al., 2020).
Recreational screen time	Time spent using screen-based media devices that does not involve school or work (Tremblay et al., 2017).
Sedentary screen time	Time spent using screen-based media devices with low energy expenditure (≤1.5 metabolic equivalents; Tremblay et al., 2017).
Stationary screen time	Time spent using a screen-based media device while standing without movement, lying, or sitting, regardless of energy expenditure (Tremblay et al., 2017).
Active screen time	Time spent using screen-based media devices while moving (Tremblay et al., 2017).
Passive screen time	Passively receiving screen-based media information (Sweetser et al., 2012).
Interactive screen time	Cognitively engaging in screen-based media activities (Sweetser et al., 2012).

Developmental Differences in Screen Use and Time

Screen-based behaviours and their impacts may be shaped by the development stages, each with its own set of challenges and implications. Understanding these differences is essential for creating appropriate screen time guidelines and interventions tailored to each age group.

Childhood is a critical developmental stage where screen use habits may have lasting effects on cognitive, social and emotional development. Children primarily engage with screen-based media devices for entertainment and learning (Rosen et al., 2014) While educational screen media have the potential to enhance education and learning, especially when co-viewed with a parent/caregiver (Liu et al., 2022), excessive screen time and media multitasking can negatively affect executive function, sensorimotor development, and academic outcomes (Muppala et al., 2023). For example, high levels of screen use during early childhood (i.e., more than 2 hours of daily screen time) have been associated with impaired language development, reduced executive functioning, and poorer academic performance (Muppala et al., 2023). Thus, the National guidelines for screen time in Australia recommend that children under 2 years of age have no screen time, children aged 2-5 years have no more than one hour per day, and youth between 5-17 years have less than 2 hours a day of sedentary recreational screen time (The Department of Health, 2019; Tremblay, Leblanc et al., 2011). Moreover, the impact of early screen exposure can be long-lasting. A longitudinal study of 2411 adolescents from the Raine Study (a prospective birth cohort study in Australia) found that screen use habits formed in childhood, such as prolonged TV viewing, may be predictive of obesity and other health issues that persist into adolescence and adulthood (McVeigh et al., 2016). Excessive screen use may also lead to problems in social-emotional development, including sleep disturbances, depression, and anxiety, and impair social and emotional competence (Muppala et al., 2023). Moreover, longitudinal data suggest that this developmental period shows the highest rates of change in self-reported TV viewing (McVeigh et al., 2016). The findings emphasise the importance of addressing screen use in early childhood, as the habits formed during this period can have lasting implications for health outcomes into adulthood.

Adolescents is a developmental stage characterised by the exploration and redefinition of personal identities and relationships (Zhu et al., 2023). Screen-based media plays a crucial role in these developmental tasks, providing a platform for social interaction and self-expression (Borca et al., 2015) Evidence suggests that during adolescence, screen time patterns shift to fulfil individuals changing needs. Coyne et al. (2018) found that while time spent on traditional media devices like television and video games remains relatively stable, time spent on social media and communication platforms increases, peaking in late adolescence. This shift reflects the growing importance of digital communication in peer relationships and identity exploration. Moreover, during adolescence researchers need to consider the variations in adolescent development trajectories of screen time. For example, a cohort study of 2411 adolescents in Australia estimated the developmental trajectories of time spent on TV from childhood (5 years old) to early adulthood (20 years old) found three distinct screen time patterns: consistently high television use, consistently low television use, and a sharp increase of television use during the adolescent years (McVeigh et al., 2016). Another study of 425 adolescents in United States of America examined the growth trajectories of texting over a 6-year period and found four distinct groups: perpetuals (i.e., showed high levels of texting during early adolescence, with slightly decreasing levels over the course of adolescence), decreasers (i.e., showed high levels of texting during early adolescence but decreased over the course of adolescence), moderates (i.e., levels of texting slightly increased over the course of adolescence), and increasers (i.e., showed a steep increase in texting levels followed by a moderate decrease over the course of adolescence). This evidence suggests there is considerable variance in adolescent screen time developmental patterns. In comparison to children’s screen use, adolescent screen use may demonstrate more complex media use patterns. One possible reason for this is that adolescents are more likely to have their own smartphone, which allows for more unsupervised use of screen-based media devices (Thomas et al., 2019).

In adulthood, screen use becomes an integral part of daily routines, including work, communications, and leisure. Unlike children and adolescents, whose screen time is often driven by education, entertainment, and socialisations, in addition to that, adults frequently engage with screen-based media devices for professional purposes (e.g., emails or online meetings). However, for adults, there are currently no screen time guidelines for total daily screen time that is adequate to maintain heath outcomes. Instead, screen time guidelines for adults are tied to sedentary behaviour guidelines (i.e., minimising long bouts of sitting; The Department of Health, 2019). Long bouts of sitting time (i.e., watching TV or sitting at a desk working on a computer) have been associated with increased risk of obesity and heart disease (Biddle, 2017). However, this does not take into account unique impacts of screen time, which may go beyond other sedentary behaviours. For instance, a systematic review of 32 studies (22 cross-sectional, 7 longitudinal, and 3 randomised control trial studies.) found an association between high levels of screen time with depression, burnout, anxiety, and lower well-being (Santos et. al., 2024). These findings suggest that screen use in adults may have broader implications for mental health, highlighting the need for targeted guidelines that address the use of screen use in this population.

In summary, while screen-based behaviours vary across different developmental stages, the habits formed in childhood may have lasting implications for health outcomes into adulthood. This evidence highlights the importance of focusing on children's screen use when developing recommendations and guidelines. Addressing screen time effectively from early childhood through to adulthood may help mitigate potential harmful health outcomes across the lifespan. Through targeting intervention during childhood, policy makers, parents and researchers may be able to promote healthy screen habits as children grow into adolescence and adulthood.

Prevalence of Screen Use in Children

The rapid development of digital technology in recent years has given children unprecedented access to screen-based media devices. Almost all children in developed countries have daily access to screen-based media devices (Rideout, 2015). By the age of 10, most children have access to as many as five different screen-based media devices at home and regularly consume multiple types of screen media simultaneously (i.e., media multitasking; Sigman, 2012).

It is no surprise that children are now spending more time on screen-based media devices than ever before. Between 2000 and 2015, children’s screen time increased from 2.9 to 3.5 hours per day (Mullan, 2018). Supporting this, a recently published scoping review examining screen-based media device use among 5–18-year-olds in 19 studies from across the world found that, on average, children and adolescents spend 3.6 hours per day on screen-based media devices (Thomas et al., 2019). Studies on Australian children’s screen time has shown similar findings. A study conducted by Yu and Baxter (2016) showed that, on average, Australian children aged 12-13 years old spend three hours per weekday on a screen device, with that time increasing to almost four hours per weekend day. Comparable findings were also found among slightly younger children, with 11–12-year-olds spending, on average, 4.75 hours per weekday and 5.8 hours per weekend day (Granich et al., 2011). However, other studies have reported much higher screen time rates. Results from the Ontario Student Drug Use and Health Survey conducted in 2017 reported that while almost two-thirds of students spend at least three hours a day on screen-based media devices, some students reported spending more than 7.5 hours per day on screen-based media devices (Boak et al., 2020). In summary, children spend a large proportion of their time on screen-based media devices.

Impact of Children's Screen Use and Time on Health

Screen use research has expanded as children's screen time has increased. Specifically, there has been an ongoing discussion in the literature about the impact of screen-based media devices on children's health and behaviours (Kaye et al., 2020). Health behaviours are actions performed by individuals that shape the health and wellbeing of individuals and populations (Short & Mollborn, 2015). These actions can be intentional or unintentional and can promote or detract from health (Short & Mollborn, 2015). In the context of children, these behaviours often include physical activity, dietary habits, screen use, and sleep patterns.

In this thesis, I follow the theoretical concept of 'health lifestyles'. Interventions and policies targeting health behaviours tend to focus on a single behaviour, rather than multiple behaviours, and often find difficulty in changing the behaviour (Short & Mollborn, 2015). The 'health lifestyle' approach instead views behaviours as influencing each other, with an understanding in the interplay between health behaviours seen as fundamental to change health behaviours (Short & Mollborn, 2015). Screen use is particularly significant as it may contribute to other health behaviours such as sedentary behaviours. Sedentary behaviours is any waking behaviour that is characterised by having low energy expenditure (≤1.5 metabolic equivalents; Tremblay et al., 2017). Sedentary behaviours can have a wide range of adverse effects, including cardiovascular disease, all-cause mortality, increased risk of cancer, and an increased risk of metabolic disorders (i.e., diabetes mellitus and hypertension; Park et al. 2020). Sedentary behaviours may also contribute to musculoskeletal issues and mental health indicators such as depression and cognitive impairment (Park et al. 2020). Thus, reducing sedentary behaviours is crucial for promoting public health.

In this section I will discuss the current evidence on the impact of children's screen use on physical health indicators, well-being and mental health indicators, sleep, and other health outcomes. In general, past studies have concluded that high levels of screen time harms children's mental health, physical health, and academic achievement (Tremblay, LeBlanc et al., 2011). As a result, national and organisational screen time guidelines advise that lower levels of screen time (i.e., < 2 hours per day) are associated with benefits for children (The Department of Health, 2019; Tremblay, Leblanc et al., 2011). However, the evidence to support these concerns is lacking.

Impact of Screen Use and Time on Physical Health Indicators

Much of the research on the impact of screen use on children's physical health shows inconsistent findings. The evidence on the impact of children's screen use on adiposity is mixed. Some systematic reviews have found an association between children's screen time and increased adiposity (Stiglic & Viner, 2019; Tremblay, LeBlanc et al., 2011). For example, a recently published systematic review of 13 systematic reviews examining the harms and benefits of screen use on children and adolescents found moderately strong evidence that higher television viewing times were associated with greater adiposity (Stiglic & Viner, 2019). However, other studies have reported small and inconsistent associations with little evidence of causality. For example, a synthesis of 29 systematic reviews and meta-analyses examined the association between sedentary behaviour and adiposity in children and adolescents and found small associations between screen time and adiposity from cross-sectional studies but inconsistent associations from longitudinal studies (Biddle et al., 2017). Supporting this, recent reviews and meta-analyses have consistently concluded that the evidence for the impact of screen time on children's health outcomes is difficult to interpret (Biddle et al., 2017; Stiglic & Viner, 2019). These inconsistent findings highlight the complexity of interpretations within the field. In the literature, there is agreement that associations of screen use on adiposity are small; however, the interpretations and clinical significance remain a debated issue.

The relationship between screen time and other physical health indicators shows inconsistent findings. Some systematic reviews have found evidence for an inverse association between screen time and cardiorespiratory fitness and musculoskeletal fitness (Carson et al., 2016; Tremblay, LeBlanc et al., 2011). For example, Carson et al. (2016) examined the relationship between objectively (i.e., device-based) and subjectively (i.e., questionnaire) measured sedentary behaviours and health indicators in children and adolescents from 235 studies (two experiment studies and 233 observation studies). They found that studies using a two-hour cut point found that higher screen time was significantly associated with lower fitness. However, this has only been examined in a limited number of studies. Other systematic reviews have shown that the evidence for an association between screen use and fitness was weak and inconsistent (Costigan et al., 2013; Tremblay, LeBlanc et al., 2011). Moreover, a recent synthesis of reviews reported a narrative synthesis across three systematic reviews and concluded that evidence for an association between screen time and cardiovascular risk factors, such as blood pressure, metabolic syndrome, insulin insensitivity and cholesterol, was weak (Stiglic & Viner, 2019). Stiglic and Viner (2019) report that the weak evidence may be related to the lack of literature rather than weak associations as there was no clear evidence for an association between individual cardiovascular risk factors and screen time.

Impact of Screen Use and Time on Well-being and Mental Health Indicators

The evidence on the impact of screen use on children’s well-being and mental health indicators is also mixed. For example, a systematic review of 91 studies, including cross-sectional, longitudinal and randomised control trial studies, examined screen-based sedentary behaviour and mental health indicators in children and adolescents found strong evidence for an association between high levels of screen time and depression and mental health outcomes, including hyperactivity/inattention problems, lower psychological well-being, internalising problems and perceived quality of life (Suchert et al., 2015). In contrast, a recent systematic review of 13 systematic reviews found weak evidence for an association between screen use and mental health outcomes, including hyper/inattention problems, poorer well-being, and poorer psychosocial health, but found moderately strong evidence for an association between screen use and depressive symptoms (Stiglic & Viner, 2019). Others have stated that the effect of screen use on children and adolescents' mental health outcomes may be too small to be meaningful (Orben & Przybylski, 2019). While, in some cases, researchers have shown that the context in which young people use specific media, such as the internet or video games, can affect mental health outcomes differently. For example, Selfhout et al. (2009) examined the longitudinal associations of time spent on internet activities for communication purposes and found that adolescents who perceived that their friendship qualities were low, using the internet to communicate predicted less depression, whereas internet use for non-communication (i.e., browsing) predicted more depression and more social anxiety. Gentile et al. (2009) investigated the effects of prosocial gaming in children and adolescents in a correlation study and two longitudinal studies and found that young people who play more prosocial games behaved more prosocially (correlation study), and prosocial game play predicted increases in prosocial behaviours 4-months later.

Thus, there is evidence of a small association between high levels of screen time and depression but inconsistent evidence for other mental health indicators. While the evidence for associations between screen use and mental health indicators remain inconsistent, the potential public health implications of these small effects should not be underestimated. The widespread prevalence of screen use means that small associations may translate into significant impacts on population-level mental health outcomes. For example, a small increase in depressive symptoms due to high screen time could lead to a burden on population-level mental health outcomes when applied across the large population of children in Australia.

Impact of Screen Use and Time on Sleep

Research has consistently shown that children's screen use and sleep are inversely related (Hale & Guan, 2015). Some reviews have found consistent evidence that screen use was associated with short sleep duration (Hale & Guan, 2015; Lund et al., 2021). However, there are mixed results regarding whether the type of screen (e.g., smartphone, computer), content (e.g., interactivity level), or context (e.g., media multitasking) of children's screen use affects sleep duration. For example, a study of 2,048 children in grades four and seven found that having a television or a small screen device (e.g., a smartphone) near where you sleep was associated with shorter sleep time (Falbe et al., 2015). In contrast, a meta-analysis of 85,561 adolescents found no association between television viewing and sleep duration but found that computer use was associated with shorter sleep time (Bartel et al., 2015). Other studies have reported that interactive screen use, such as video games, has a greater impact on sleep than passive screen use (Falbe et al., 2015; Hale & Guan, 2015). Thus, consistent evidence shows screen use is associated with short sleep duration; however, further nuanced evidence is needed for researchers to have a better sense of the magnitude and clinical significance of the observed associations.

Other Impacts on Other Health Outcomes

New findings suggest that moderate screen time levels (i.e., three to six hours of screen use each day) could have benefits over high use or abstinence (Ferguson, 2017). A cohort study with over 35,000 children and their caregivers found that moderate amounts of television-based and device-based screen time were associated with higher levels of psychosocial functioning than non-users (Przybylski et al., 2020). Supporting this, other studies have found a curvilinear relationship between screen time and children and adolescent health and well-being outcomes (Bélanger et al., 2011; Przybylski & Weinstein, 2017). However, a recently published study examining linear and curvilinear relationships between screen time and 4013 children's physical health, psychological outcomes, and educational outcomes found little evidence of curvilinear relationships. Instead, they found small linear associations (i.e., standardised effects < 0.07) in outcomes moderated by the type of screen use (Sanders et al., 2019). In particular, passive screen use was associated with worse outcomes, educational screen use was associated with slight benefits in school achievement, and interactive screen use was associated with positive educational outcomes but had negative associations with other outcomes (Sanders et al., 2019). While these effect sizes may be small, their public health significance should not be underestimated. In the context of a behaviour that is as ubiquitous as screen use, small linear relationships can have implications at the population level. For example, the negative association between passive screen use and health outcomes, although small, could translate into a significant impact on public health given the high prevalence of passive screen activities such as television viewing.

Other studies suggest that while some types of screen use, such as television viewing, may be negatively associated with children's health and behaviour (Hale & Guan, 2015; Stiglic & Viner, 2019; Sweetser et al., 2012), evidence for other types, such as video games and social media, remain less certain, and in some instances may be beneficial (Boot et al., 2008). Further, studies suggest that the context of children's and adolescent's screen use may affect health and behaviour outcomes differently. For example, another study found that total screen time, particularly television viewing time, was detrimental to almost all aspects of health; however, computer and video game use was not consistently associated with any physical health indicators (Carson et al., 2016). These findings suggest that the type of device and context of children's screen use may have a greater impact on screen use estimates than the screen time itself. Supporting this idea, others have argued that the measurement of screen time is meaningless unless the type and context of the screen time are considered (Kaye et al., 2020; Odgers & Jensen, 2020).

In summary, much of the research on the impact of screen use on children's health and behaviours shows inconsistent findings. While most of the findings suggest that high levels of screen time may be detrimental to children's health outcomes, most observed associations between screen time and health outcomes are small, and their interpretations and clinical significance remain contentious. Research evidence also highlights potential benefits from different types and contexts of children's screen use. Moreover, the inconsistent findings may stem from several sources of bias and methodological limitations in the literature.

First, there is variability in study designs, including differences in how screen time is measured and reported. For example, some studies rely on self-reported screen time, which is prone to social desirability or inaccurate recall, leading to potential over- or underestimation of screen time (Kaye et al., 2020). Additionally, the diversity in types of screen-based media devices (i.e.., television, computers and smartphones) and content (i.e., interactive screen time and passive screen time) presents challenges when directly comparing the effects of screen time. The evidence suggests that passive screen use, such as televising viewing, is more consistently associated with negative health outcomes compared to interactive or educational screen use (Sanders et al., 2019). However, the differentiation between the different types of screen-based media devices and content is not always made in the literature, which may lead to misleading conclusions on the overall impact of screen time.

Further, many studies rely on cross-sectional designs, which only show associations rather than casual relationships between screen time and children's health outcomes. This limitation is highlighted by the inconsistency in longitudinal findings, where some studies report small effects of no clear patterns over time (Biddle et al. 2017). The lack of robust longitudinal data diminishes the ability for researchers to draw definitive conclusions about the causality and long-term impacts of screen time on children's health outcomes.

In light of these considerations, it is evident that while high levels of screen time are associated with some negative health outcomes, the evidence is nuanced and may be influenced by several biases and methodological limitations. Additionally, when interpreting the small effect sizes in the literature, it is important to consider the potential public health implications. Given the widespread prevalence of screen use among children, small negative associations with health outcomes can translate into significant public health challenges. In line with recent discussions in the literature, further evidence is needed on how the different types and contexts of screen use, rather than solely screen time, are associated with children's health outcomes (Kaye et al., 2020). Understanding the how different types and contexts of screen use impact children's health outcomes will guide policy makers, educators, and parents in making informed decisions on children's screen use.

Screen Use Measurement

In the literature, debate continues regarding the size and clinical significance of screen use associations with health outcomes; however, researchers are now highlighting the methodological and philosophical issues, such as poor conceptualisation of screen use and the use of non-standardised measurements, surrounding screen use research (Byrne et al., 2021; Kaye et al., 2020). As I alluded to above, there are several complex and nuanced issues in screen use measurement that may be contributing to the inconsistent findings. In this section, I will discuss the importance of valid and reliable measurement in research. I will provide an overview of the current methodology used in screen use research on children. I will then discuss the current methodological issues in screen use research that are central to this thesis.

What is a 'Good' Measurement?

Given the importance of improving the quality of research on children's screen use, it is important to understand the concepts of measurement. Measurement is the assigning of scores to observations so that we can quantify a phenomena (Kimberlin & Winterstein, 2008). Measurement involves the operationalisation of constructs to variables and the development and application of instruments to quantify these variables (Kimberlin & Winterstein, 2008). If an instrument has the ability to measure the accurate value of a constuct, the scientific quality of such research will increase (Kimberlin & Winterstein, 2008). However, during the operationalisation of constructs, development of an instrument, and application of an instrument, it is common for errors to be introduced, which may impact the quality and integrity of the measurement (Kimberlin & Winterstein, 2008). In general, a key indicator of the quality of a measurement instrument is the validity and reliability of the measurements (Kimberlin & Winterstein, 2008).

In the literature, some definitions and terminology of validity and reliability are used interchangeably or with different categorisations (Mokkink et al., 2009). The lack of consensus has led to confusion about measurement properties and the concepts they represent (Mokkink et al., 2009). In this thesis, I will use the concepts and definitions of validity and reliability defined by the COSMIN (Consensus-based Standards for the Selection of Health Measurement Instruments) taxonomy of measurement properties (Mokkink, Terwee, Patrick, et al., 2010). The COSMIN definitions of measurement properties are based on an international consensus on terminology and definitions for measurement properties related to health-related patient-reported outcomes (Mokkink, Terwee, Patrick, et al., 2010).

Validity

Validity is defined as the degree to which an instrument measures the construct it aims to measure (Mokkink, Terwee, Patrick, et al., 2010). Validity is not the property of the instrument itself. Instead it is the property of the instrument's score and interpretations (Kimberlin & Winterstein, 2008). In the context of screen use measurement, validity describes how well researchers can trust the scores of an instrument when measuring screen time, content and context. In general, validity is distinguished by three types of validity: content validity, criterion validity and construct validity.

Content Validity. Content validity is defined as the extent to which the content of a measurement instrument adequately reflects the construct it aims to measure (Mokkink, Terwee, Patrick, et al., 2010). An instrument with high content validity should measure all relevant parts of the theoretical concept, theme, or idea it aims to measure. For example, if the construct we want to measure is screen use, we need to measure the screen time, content, and context of the child's screen use.

An aspect of content validity is face validity (Mokkink, Terwee, Patrick, et al., 2010). Face validity is defined as the extent to which a measurement instrument appears to adequately reflect the construct it aims to measure (Mokkink, Terwee, Patrick, et al., 2010). Face validity concerns whether an instrument measurement appears to be relevant and appropriate for measuring screen use. For example, if an instrument is designed to assess violent content in screen use, face validity would be concerned with whether the instrument seems to appropriately capture violent content based on its design and item content. Face validity is usually assessed without empirical testing and is often considered the weakest form of validity (de Vet et al., 2011).

Criterion Validity. Criterion validity is defined as the extent to which the scores of a measurement instrument adequately reflect those of a "gold standard" measurement instrument (Mokkink, Terwee, Patrick, et al., 2010). Criterion validity can only be evaluated by an instrument that is considered as 'gold standard' (Scholtes et al., 2011). If the comparison instrument is not considered as 'gold standard', it may lead to misleading results.

Direct observation (either by video or a researcher) is considered a gold standard method for measuring screen use (Anderson et al., 1985; Perez et al., 2023). Thus, a new screen use measurement tool should be compared with a well-established direct observation tool to demonstrate how well the new tool aligns with the direct observation measurements.

Criterion validity can be further subdivided into concurrent validity and predictive validity. Concurrent validity assesses the measurement of the instrument and the measurement of the gold standard at the same time. For example, if a phone application tracking screen time scores adequately reflect the amount of screen time record by direct observation, the phone application demonstrates good concurrent validity. In contrast, predictive validity assesses whether the instrument measurement predicts the gold standard measurement in the future (de Vet et al., 2011). For example, if a questionnaire estimates future screen time and the actual screen time measured at a future time reflects the questionnaire estimates, this indicates good predictive validity.

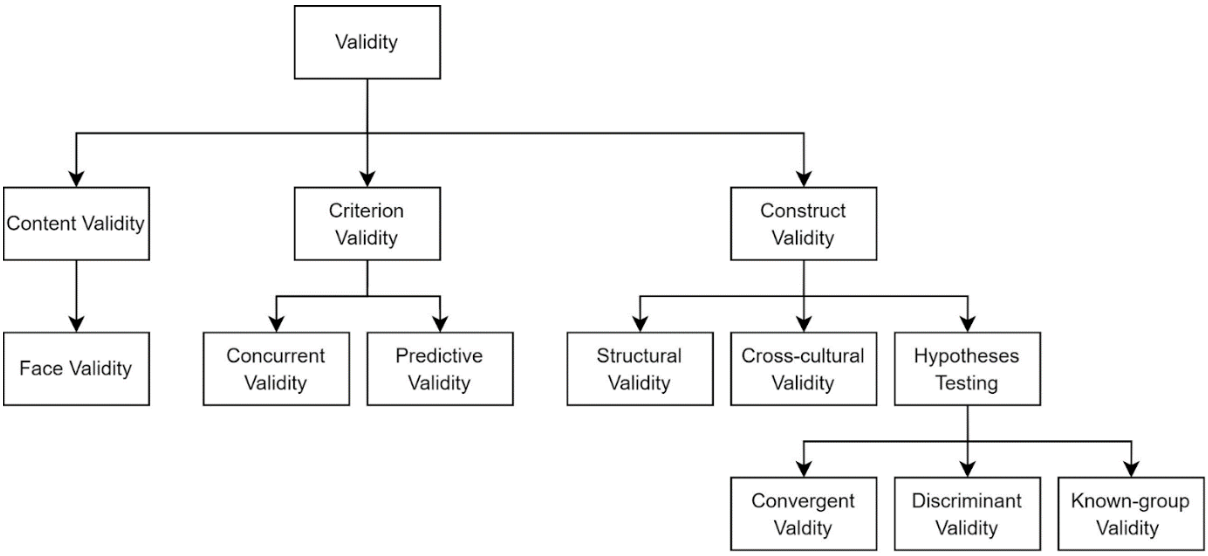
Construct Validity. Construct validity is defined as the extent to which the scores of an instrument measurement are consistent with the hypotheses (Mokkink, Terwee, Patrick, et al., 2010). Hypotheses testing refers to the investigation of the magnitude and direction of a correlation or difference to what is expected of the construct being measured (Mokkink, Terwee, Knol, et al., 2010). Therefore, hypotheses testing for construct validity includes convergent validity and discriminant validity. Convergent validity refers to whether the scores of the instrument measurement correlate with another instrument it should theoretically be related to (de Vet et al., 2011). For example, if two different measurement instruments measure screen time, both measurement instrument screen time results should be correlated. An instrument with high convergent validity would have a high agreement or concordance with another instrument that measures the same construct. Conversely, discriminant validity refers to whether the scores of the instrument measurements that are theoretically unrelated to each other are, in fact, unrelated (de Vet et al., 2011). For instance, if you are measuring screen time, the screen time tool should not be correlated strongly with an instrument measuring physical activity. If the screen time measurement instrument does not show a strong relationship with physical activity instrument measurement, it indicates that the measurement instruments is specifically measuring screen time and not physical activity.

Other Types of Validity. It is important to note that there are other types of validity in the literature that have not been discussed here. For example, external validity is the extent to which results can be generalised to the larger population (Steckler & McLeroy, 2008), whereas internal validity refers to the extent to which a study establishes the cause-and-effect relationship and thus is not due to methodological errors (Willis, 2007).

Accuracy. In the literature the term 'accuracy' is often used interchangeably with validity and reliability (Menditto et al., 2007). In this thesis, accuracy refers to a qualitative performance characteristic, which expresses the closeness of agreement between a measurement value and the true value of the quantity being measured (Menditto et al., 2007).

Figure 1.1

Validity Measurement Properties based on COSMIN Taxonomy of Measurement Properties



Reliability

Reliability is the "degree to which the measurement is free from measurement error" (Mokkink, Terwee, Patrick, et al., 2010, p. 743). Reliability is a broad concept and contains the measurement properties: internal consistency, reliability (including test-retest, inter-rater and intra-rater) and measurement error (including test-retest, inter-rater and intra-rater). The extended definition of reliability is the extent to which participant scores are the same for repeated measurements under different conditions (Mokkink, Terwee, Patrick, et al., 2010). For example, some conditions may include administrating the same questionnaire at different time points (i.e., test-retest; Mokkink, Terwee, Patrick, et al., 2010). Other conditions are related to the person who is administrating the instrument. For example, administrating the same questionnaire on the same occasion by different persons (i.e., inter-rater) or on different occasions by the same person (i.e., intra-rater; Mokkink, Terwee, Patrick, et al., 2010).

Reliability. The measurement property of reliability is defined as the proportion of the overall variance in the instrument measurements attributed to "true" differences between participants (Mokkink, Terwee, Patrick, et al., 2010). In this context, reliability refers to the extent to which variations in screen time measurements are due to 'true' differences in screen use among participants, rather than inconsistencies in the measurement instrument itself. Reliability is commonly assessed by the Intra-Class Correlation Coefficient (ICC), a Generalisability Coefficient or Kappa (Mokkink et al., 2020).

Measurement Error. Measurement error is defined as the random and systematic error of a participant's score that cannot be attributed to the true changes in the construct it aims to measure (Mokkink, Terwee, Patrick, et al., 2010). In the context of screen time, measurement error refers to how close the scores of repeated measurements of screen use in the same participant are (Mokkink et al., 2020). For example, if you use a questionnaire to measure a child's daily screen time, measurement error would be reflected in how similar the questionnaire measurements of screen time are for the same day, under the same conditions. Measurement error is commonly assessed by percentage total agreement or percentage specific agreement for categorical outcomes (Mokkink et al., 2020). For continuous outcomes, measurement error is assessed by the Standard Error of Measurement Limits of Agreement (Mokkink et al., 2020).

Both reliability and measurement error consist of three aspects: test-retest reliability, inter-rater reliability, and intra-rater reliability. It is important to note that reliability and measurement error are assessed using the same study design and data collection but have different statistical methods (Mokkink et al., 2020). These measurement properties are related but have unique statistical purposes. For example, if I calculated the inter-rater reliability of screen time measurements given to a participant by two raters using Kappa, I would be assessing the reliability of the raters. If I calculated the inter-rater reliability of the same screen time measurements given to a participant by two raters using percentage agreement, I would be assessing the measurement error of the two raters. Therefore, test-rest reliability, inter-rater reliability and intra-reliability can be considered as a reliability measurement or a measurement error measurement depending on the statistical method used (Mokkink et al., 2020).

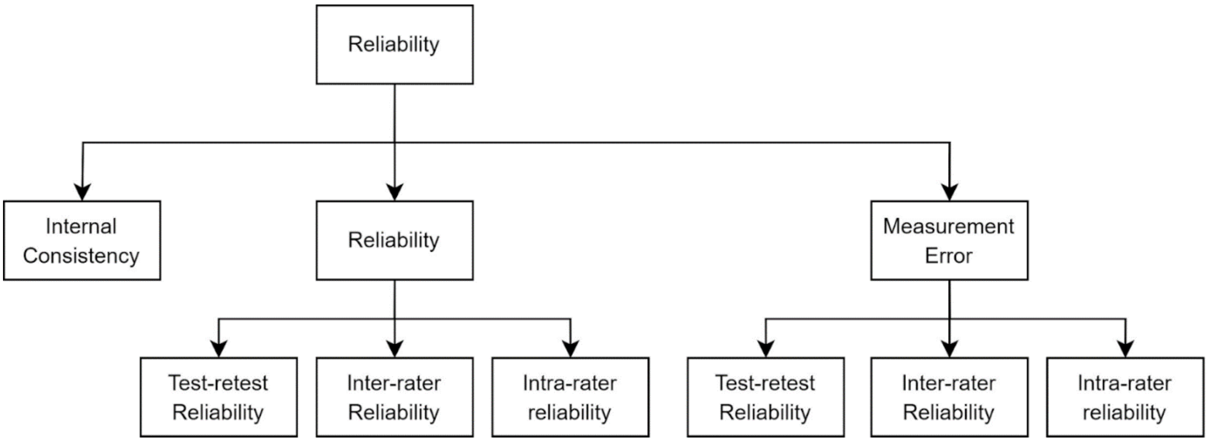
Test-retest Reliability. Test-retest reliability refers to the extent to which instrument scores are consistent when taken by an instrument on the same subject under the same conditions (Koo & Li, 2016). The scores from the Time 1 and Time 2 instrument administration are compared to evaluate the instrument for stability over time. For example, to assess test-retest reliability of a screen time questionnaire, individuals may be asked to complete the same questionnaire two times within a two-month interval so that the questionnaire results can be compared to assess the stability of the scores. An important point to consider when assessing test-retest reliability is the stability of the behaviour over time (Koo & Li, 2016).

Inter-rater Reliability. Inter-rater reliability refers to the extent to which scores from a measurement instrument remain consistent when the measurements are taken by two or more raters on the same subjects under the same conditions (Koo & Li, 2016). For example, if two researchers independently record screen time for the same group of children using the same direct observation instrument, inter-rater reliability would be assessed by comparing their scores to see how consistent they are.

Intra-rater Reliability. Intra-rater reliability refers to the extent to which scores from a measurement instrument remain consistent when the measurements are taken by the same rater using the same methods across different occasions (Koo & Li, 2016). For example, if a researcher records a child's screen time on two different days using the same method, intra-rater reliability would be evaluated by comparing the scores from both days to ensure the consistency of the rater's measurements. Figure 1.2 presents an overview of the reliability measurement properties.

Figure 1.2

Reliability Measurement Properties based on COSMIN Taxonomy of Measurement Properties



Using valid and reliable instruments is a crucial to research quality (Kimberlin & Winterstein, 2008). Establishing the validity and reliability of measurements in research ensures that measurements are replicable and accurate (Kimberlin & Winterstein, 2008). With higher validity, researchers can be more confident that the instrument scores are closely linked to the construct being measured (Kimberlin & Winterstein, 2008). With higher reliability, researchers can be more confident that the instrument scores are reproducible over a variety of conditions (Numnnally, 1978). Thus, a high-quality instrument measurement has both high validity and high reliability.

How has Children's Screen Use been Measured?

Despite the importance of improving our understanding of the impact of screen use on children's health and behaviours, currently, there is no consensus on the best way to accurately measure children's screen use (Jordan et al., 2007; Kaye et al., 2020). At present, the most common way to measure children's screen use is through self- and proxy-reported measurement instruments such as questionnaires and time-use diaries (Hale & Guan, 2015; Stiglic & Viner, 2019). Other studies have measured screen use through device-based measurement instruments such as smartphone usage applications, video recording, television monitors and accelerometers (Ellis et al., 2019; Reeves et al., 2020). Each method has strengths and limitations.

Self- and Proxy-Reported Measurements

Children's screen use is typically measured through self- and proxy-reported measurement instruments such as questionnaires and time-use diaries (Hale & Guan, 2015; Stiglic & Viner, 2019). Questionnaires require participants to estimate their screen use retrospectively. For example, in screen use questionnaires, participants are often asked questions such as "How many hours of screen time did you use in a typical day last week?" (Kaye et al., 2020). Time-use diaries record continuous actions through the data collection period (often a 24- or 48-hour period; Bauman et al., 2019). For example, some studies have asked participants to record their screen use by completing diaries by blocks of time (e.g., 15, 30 or 60 minutes) over the course of a day (Jordan et al., 2007). Self- and proxy-reported measurement instruments are cost-effective compared to other screen use measurement instruments and typically have low burden on participants (Hardy et al., 2013). However, self- and proxy-reported measurement instruments are susceptible to various forms of bias and inaccuracies due to recall bias, misclassification bias, and social desirability bias (Hardy et al., 2013). Additionally, when compared to device-based measurements of screen use, such as video observation and phone use tracking applications, self-reported measurements are often inaccurate and unreliable (Anderson et al., 1985; Perez et al., 2023).

Device-Based Measurements

Device-based measurement instruments provide an alternative approach that may overcome some of the challenges of self- and proxy-reported screen use measurement instruments. Direct observation (either by video or a researcher) is considered a gold standard method for measuring screen use (Anderson et al., 1985; Perez et al., 2023). Direct observation measurements typically involve video recording participants' television viewing at home (Anderson et al., 1985). However, few studies have used video and direct observation as they are time intensive and invasive, and often do not measure the content of the screen-based media device (Perez et al., 2023). Direct observation measurements are also susceptible to the Hawthorne effect (McCambridge et al., 2014). The Hawthorne effect is a phenomenon where participants change their behaviour in response to the awareness of being studied (McCambridge et al., 2014).

Developments of applications such as the Apple Screen Time Application give researchers access to objective measurements of screen use behaviours surrounding smartphone use (Ellis et al., 2019). Applications such as the Apple Screen Time Application record the user's time using smartphone applications based on categories such as 'productivity', 'information and reading', 'health and fitness', and 'entertainment' (Ellis et al., 2019). Such measurement instruments overcome the chance of misclassification bias and poor recall that is seen in self- and proxy-reported measurement instruments. However, these measurement instruments are limited to smartphone use and do not record screen time on other screen-based media devices. This is a significant limitation as while smartphones are widely used, they only represent a part of the broader digital environment. Evidence shows that other devices such as televisions and computers, play a crucial role in overall screen use, especially among children and adolescents (Thomas et al., 2019). Other researchers have built software that tracks participant's screen use over multiple screen-based media devices. For example, the Human Screenome project has recently built a software platform that records, encrypts, and transmits screenshots of a participant's screen-based media device every 5-seconds whenever the device is turned on (Reeves et al., 2020). The platform can be used on several screen-based media devices simultaneously and synchronises the screenshots from each device (Reeves et al., 2020). However, this type of measurement instrument is not accessible to most researchers and raises concerns about participant's privacy and surveillance (Kaye et al., 2020; Reeves et al., 2020). Further, device-based applications cannot account for device sharing, which often occurs within family homes (Perez et al., 2023).

Other device-based measurements used by researchers have relied on existing measurement technologies. For example, studies have used television monitors attached to every television in the home to detect television screen time (Robinson et al., 2006). However, this measurement is limited to only television viewing and does not measure other devices or the content being watched (Robinson et al., 2006). In the literature, other researchers have used accelerometers to measure sedentary time as a proxy for screen use

(Hardy et al., 2013). However, accelerometry measurement as a proxy for screen use has been shown to have poor reliability and validity (Mendoza et al., 2013). Accelerometer measurements also do not measure the context of the screen use behaviour (i.e., the type of device used, or screen media watched; Hardy et al. 2013). Table 1.2 presents a summary of the self- and proxy measurements, and device-based measurements.

Table 1.2

Methodology used to Measure Children's Screen Use

Type of measure	Instrument	Method	Strengths	Limitations
Subjective	Self-report questionnaires	Retrospective recall from participant	Low cost Accessible to researchers Provides information on the device and duration of screen use. Can be used to measure large samples	Recall bias is common. Susceptible to misclassification bias. Susceptible to social desirability bias. Not suitable for children. Low validity and reliability in some instruments.
	Proxy-report questionnaires	Retrospective recall from parent/guardian of the participant	Low cost Accessible to researchers Provides information on the device and duration of screen use. Can be used to measure large samples.	Recall bias is common. Susceptible to misclassification bias. Susceptible to social desirability bias. Not suitable for older children.

Type of measure	Instrument	Method	Strengths	Limitations
			More suitable to measure young children's screen use	Low validity and reliability in some instruments.
	Time-use diaries	Prospective/current	Low cost Accessible to researchers Provides information on the device and duration of screen use. Can be used to measure large samples.	Potential for participant reactivity Susceptible to social desirability bias Poor compliance Age limitation for memory Low validity and reliability in some instruments
Device-based	Smartphone apps	Prospective/current. Measurement captured by smartphone application.	Low cost Objective measurement Captures accurate screen time and content	Limited to the single device Only captures data on screen time and content
	Platform recording over multiple devices	Prospective/current. Measurement captured by screenshots of the device.	Objective measurement Captures accurate screen time and content.	Costly Not accessible for most researchers to use.

Type of measure	Instrument	Method	Strengths	Limitations
			Captures accurate data over multiple devices	Privacy and surveillance concerns Does not consider device sharing
	Video or Direct observation	Prospective/current. Observation with a video or in-person.	Considered a gold standard measurement.	Costly Time intensive Potential for participant reactivity (i.e., the Hawthorne effect) Can be considered invasive
	Television monitor	Prospective/current. Device attached to televisions within the home.	Objective measurement	Outdated as it does not measure other devices. Does not measure content. Potential for participant reactivity (i.e., the Hawthorne effect)

Type of measure	Instrument	Method	Strengths	Limitations
	Accelerometer	Prospective/current. Monitor worn by participants.	Objective measurement	Does not provide context of the screen use behaviour. Low validity and reliability

Methodological Issues in Screen Use Measurements

The inconsistent findings in screen use research may be due to the way screen use has been measured (Kaye et al., 2020). As I outlined above, there have been several methods used to measure children's screen use, including self-and proxy measurement instruments such as questionnaires and time-use diaries, smartphone usage applications, software platforms, video recording, television monitors and accelerometers (Ellis et al., 2019; Kaye et al., 2020; Reeves et al., 2020). Nevertheless, the limitations associated with these methods may contribute to the inconsistent findings in the screen use literature. In this section, I have discussed the limitations of the current screen use literature.

Reliance on Subjective Measurements. First, much of the research is based on self- or parent-report measurements of screen use (Hale & Guan, 2015; Stiglic & Viner, 2019). Self-reported estimates of sedentary behaviour have been shown to be prone to measurement error and may lead to bias in study results (Ainsworth et al., 2012; B. Clark et al., 2011). Moreover, child-reported measurements have been shown to not be the most valid and reliable approach to screen use measurement due to the limited cognitive capacity and increased recall bias among paediatric populations in research (Atkin, Ekelund, et al., 2013; Saunders et al., 2011). For example, Saunders et al. (2011) stated that in some cases, children have reported unrealistically high amounts of daily screen time (e.g., 13.5 hours per day) when estimating their screen time. Evidence suggests that parent-reported measurements of screen time may also not provide an accurate estimate of children's screen use. For example, when compared to an electronic television monitor, parents overestimated screen time by four hours per week if there was no television in the child's bedroom and underestimated by three hours per week if there was a television in the bedroom (Robinson et al., 2006). When parent-reported measurements were compared with applications that tracked children's mobile device use, 34.8% of parents overestimated and 35.7% underestimated their child's mobile device use (Radesky et al., 2020). These findings suggest that parent-reported measurements of screen use for children may be prone to misclassification bias due to parents not being aware of their child's actual screen use (e.g., children may engage with screens in separate rooms or at their friends' homes; Jordan et al., 2007). Device-based measurement instruments for measuring screen use such as smartphone usage applications and software for platforms, have demonstrated higher correlations with gold standard measurement instruments (Perez et al., 2023). However, many are still newly developed and require further validation before being used in the field.

Screen Use Measurements Lack Precision. Next, the majority of studies measure "total" screen time (i.e., aggregated) and often do not measure the type of content (e.g., educational game, children's television programme) or the context (e.g., alone vs with others and school vs home; Greitemeyer & Mügge, 2014; Latomme et al., 2018; Twenge et al., 2019). When studies have measured the context or content of screen use, they are often only investigated in isolation (e.g., TV violence) rather than in context (e.g., did the child watch alone or with a parental? Stiglic & Viner, 2019). This methodology does not take into account different contexts and types of content that may impact the effects of screen use on children's health and behaviour outcomes (Odgers & Jensen, 2020; Sanders et al., 2019). Most of the measurement instruments also do not consider media multitasking (Kaye et al., 2020). Media multitasking is when a user simultaneously uses two or more screen-based media devices (e.g., the simultaneous use of a television and a smart phone) or engaging in multiple activities on a single device (e.g., watching a video on a computer while online shopping; van der Schuur et al., 2018). This methodology may impact the overall measurement of children's screen use.

In summary, there are several complex issues in screen use measurement that may be contributing to the inconsistent findings in screen use research. First, much of the research is based on self- or proxy-reported measurements of screen use, which are prone to bias and inaccuracies (Ainsworth et al., 2012; B. Clark et al., 2011). Next, most studies used aggregated ‘total’ screen time measurements that do not take into account the different types of devices, content and contexts of screen use that may impact the effects of screen use on children’s health and behaviour outcomes (Odgers & Jensen, 2020; Sanders et al., 2019). This methodology limits our understanding of how the different types of screens, content, and context may have different effects on children’s outcomes. Failing to account for these factors could lead to incomplete or misleading recommendations for screen use guidelines and interventions, which may impact children’s developmental and health outcomes. Understanding these nuances is crucial for developing comprehensive and effective interventions that address both the harms and benefits of children’s screen use.

Automated Wearable Cameras as a Potential Solution

Automated wearable cameras present an opportunity to assess children’s screen use more accurately. Automated wearable cameras are a form of ecological momentary assessment, which involves repeated sampling of a participant’s behaviours and surrounding environment (Shiffman et al., 2008). Ecological momentary assessments capture real-time data in naturalistic settings to reduce recall bias and increase ecological validity (Shiffman et al., 2008). Thus, ecological momentary assessments can capture dynamic behaviours and contextual factors that might be missed by traditional retrospective self-reported methods. However, ecological momentary assessments are not without its limitations. The frequent data collection required by ecological momentary assessments can be burdensome for participants, which may increase non-compliance or alter participant behaviour due to the participant’s awareness of being monitored (Shiffman et al., 2008). Additionally, ecological momentary assessments can generate large amounts of data, which may complicate data management and statistical analysis methods (Shiffman et al., 2008).

In the context of automated wearable cameras, these cameras are typically worn on the chest and take first-person point-of-view images at fixed intervals (see Figure 1.3; Doherty, Hodges, et al., 2013). Consequently, these cameras can passively capture rich contextual data of the participants' behaviour.

Figure 1.3

Example of an Automated Wearable Camera used in Research (Model: Brinno TLC130)

Have Automated Wearable Cameras been used in Research before?

Automated wearable cameras were first developed in the field of digital lifelogging (Kelly et al., 2011). Digital lifelogging refers to the digital capture of everyday activities through a first-person point-of-view (Kelly et al., 2011). Since then, human behaviour research has increasingly used automated wearable cameras as the cameras have become more affordable, smaller, and are able to capture more data for extended periods of time (Barr et al., 2015; Beltran et al., 2018; Doherty, Kelly, et al., 2013; Signal, Smith, et al., 2017). In recent years, automated wearable cameras have been used to measure a range of health behaviours and exposures in children and adults. Such cameras have been used in studies investigating physical and sedentary behaviour in adults (Doherty et al., 2012; Kerr et al., 2013; Leask et al., 2015), diet in children and adults (Beltran et al., 2018; Gemming, Utter, et al., 2015), children's marketing exposure (Barr et al., 2015; Cowburn et al., 2016; Signal, Stanley, et al., 2017), and children's travel methods to school (Kelly et al., 2012). Despite these advances, automated wearable cameras remain a relatively new methodology in health behaviour research, and as such, may present opportunities and challenges that should be considered.

One challenge is the validity and reliability of the data captured by automated wearable cameras. While automated wearable cameras have been used in studies investigating the physical and sedentary behaviours, diet, children's marketing exposure and travel methods to school (Doherty et al., 2012; Kelly et al., 2012; Leask et al., 2015), there is limited evidence on how accurately these cameras reflect children's behaviours. For instance, one observational study used data from automated wearable cameras with accelerometers to assess the context of accelerometer-identified episodes of physical activity in a convenience sample of 52 adults (Doherty, Kelly, et al., 2013). The findings of the study demonstrated that automated wearable cameras were able to identify and categorise accelerometer episodes of physical activity in the context of the participant's daily life. However, they also highlighted the need for further research to validate the automated wearable camera measurements against other objective measurements, such as direct observation. Overdependence on automated wearable cameras without thorough validation could lead to misleading measurements of health behaviours, particularly when measuring dynamic behaviours and contextual factors.

Additionally, the presence of a camera may change the participant's behaviour via the Hawthorne effect. A feasibility study assessed the use of automated wearable cameras to document children's exposure to environmental determinants of obesity, such as food marketing (Barr et al., 2015). Findings from the study suggest that automated wearable cameras were an effective method for measuring children's exposure to food marketing. However, the authors noted that the camera's presence may change the participant's behaviour via the Hawthorne effect. Some studies have mitigated this by blinding participants to the study's purpose (Barr et al., 2015; Signal, Smith, et al., 2017). Barr et al. (2015) blinded the participants (convenience sample of 6 children aged 12 years old) from the main study aim of assessing the feasibility of using cameras to document children's exposure to food marketing in different settings to overcome this. Signal, Smith et al. (2017) blinded 168 participants aged 11-13 years to the primary food marketing focus of their study to reduce social desirability bias and camera reactivity. Although, qualitative evidence from Wilson et al. (2016), suggests that individuals wearing the camera often forgot about the camera's presence, remembering it only sporadically (Wilson et al., 2016). These findings highlight the complexity of mitigating the Hawthorne effect in research using automated wearable cameras, suggesting further research is needed to understand and mitigate this potential bias.

Ethical Challenges in Research

The use of automated wearable cameras also pose ethical challenges related to confidentiality, privacy and autonomy (Kelly et al., 2013; Mok et al., 2015). Such ethical challenges include handling inappropriate images, confidentiality, and protecting participants' and third parties' privacy (Kelly et al., 2013). In the literature, researchers often report privacy and ethical concerns about using these devices (Mok et al., 2015). To address ethical concerns, Kelly et al. (2013) developed an ethical framework that protects participants' and third parties' according to ethical principles for automated wearable cameras in health behaviour research. The recommended guidelines address the following: (1) informed consent (i.e., robust informed consent procedures that explicitly inform the participant of the nature of the data collection process); (2) privacy and confidentiality (i.e., data management procedures to enhance data security); (3) non-maleficence (i.e., participants should be prepared for questions about the camera by the public and be instructed to remove the device if any situation); and (4) autonomy of third parties (i.e., participants should obtain verbal consent from third parties and remove the device if third parties are uncomfortable). In this section, I discuss the ethical issues associated with automated wearable camera in research, and the procedures to mitigate potential harm.

Informed consent. A primary concern with the use of automated wearable cameras as a measurement instrument is the potential intrusion on participants' privacy and autonomy (Mok et al., 2015). Unlike traditional photography, where participants consciously choose when to take an image, image capture with an automated wearable camera is automatic and passive, meaning the participant does not have control over the timing and content of the image (Kelly et al., 2013). As a result, some images captured may be unwanted or unflattering. Thus, it is important to implement robust informed consent procedures that explicitly inform the participant of the nature of the data collection process (Kelly et al., 2013). Kelly et al (2013) states participants should be informed of; the passive nature of the image capture; the estimated volume of images that will be captured; and the potential for unwanted or unflattering image capture. Moreover, participants should be given tools to actively manage their privacy, such as the ability to review and delete images, and the ability to remove or disable the camera at any time (Kelly et al., 2013). Some studies use automated wearable cameras with a privacy button, where participants can disable image taking for a specific timeframe (i.e., 7 minutes; Kelly et al., 2012; Schrempft et al., 2017).

Privacy and confidentiality. Automated wearable cameras capture large amounts of data of the participants' environment, increasing the risk of accidental disclosure of sensitive information (Kelly et al., 2013). Unlike traditional participant-initiated photography, which may capture up to 30 images per day, automated wearable cameras are likely to capture between 2,000-4,000 images per day. Consequently, this will result in detailed information about the participant and their day being captured (i.e., identifiable features, such as faces or locations). Thus, the potential harm from the disclosure of automated wearable camera data is significant, particularly if the images can be accessed publicly (i.e., a third party finds a camera the participant lost) or used inappropriately (i.e., images are posted online). To protect participant confidentiality, researcher must use rigorous data management procedures to enhance data security. Common data management procedures to protect participant confidentiality include password-protecting devices and data files, such that only the authorised research team members have access to the data (Barr et al., 2015; Kelly et al., 2013; Smith et al., 2019). Additionally, when dissemination of research findings, researchers are required to de-identify any identifying features (i.e., faces or street signs; Kelly et al., 2013). Thus, researchers must implement robust privacy measures to mitigate the risks associated with automated wearable camera data

Non-maleficence. There is a risk of burden or harm to participant when wearing an automated wearable camera in the public (Kelly et al., 2013). For instance, a participant may be questioned by third parties who do not want to be recorded. A previous study investigating individuals' perceptions and reactions to being recorded by cameras in digital lifelogging found that third parties were mainly focused on ensuring that the recording had a valid purpose and on safeguarding their images and identities (Nguyen et al., 2009). Thus, participants should be prepared with a statement explaining the study purpose, and contact details of the researchers so third parties can request deletion of images they are in.

Autonomy of third parties. A unique challenge posed by automated wearable cameras is the capture of images of third parties who are not participants in the study. Third parties may include family members, coworkers, friends or people in the public. These individuals do not have the opportunity to provide informed consent, however; have their images collected as part of the study data. This issue raises concern over the potential violation of third-party privacy rights (Kelly et al., 2013). To address this issue, researchers must establish clear protocols for the treatment of third-party data. Kelly et al. (2013) suggests participants should be instructed to inform individuals who frequently interact with the participant, about the study and obtain their verbal consent, and if requested, the device should be removed. Additionally, participants should be advised to avoid wearing the camera in certain settings such as public bathrooms, school grounds and public swimming pools. By establishing these privacy measures, researchers can protect the privacy of third parties (Kelly et al., 2013).

Automated wearable cameras may present an opportunity to assess children's screen use more accurately. However, these benefits must be weighed against the ethical challenges. Unlike traditional photography research methodologies, automated wearable cameras generate more image data, increasing the risk of capturing unwanted images. Kelly et al. (2013) found that these ethical challenges could be adequately addressed through using appropriate procedures to protect confidentiality and privacy, informed consent, and respect for autonomy. Supporting this, a study investigated the ethical and practical implications of having children aged 12 years wear an automated wearable camera (Barr et al., 2015). The authors concluded that the privacy and practical implications of children wearing automated wearable cameras could be adequately addressed with a study protocol that addresses these participants' and third parties' privacy, confidentiality, and anonymity concerns. By implementing these privacy measures, it appears researchers can effectively protect the privacy, confidentiality and autonomy of participants and third parties.

Automated Wearable Cameras in Screen Use Measurement

Automated wearable cameras have been shown to be a feasible method for measuring adolescent screen use behaviour (Smith et al., 2019). Smith et al. (2019) conducted a study to assess the acceptability and feasibility of automated wearable cameras to measure pre-bedtime screen use in a sample of 15 participants aged 13-17 years. Participants wore an automated wearable camera for three evenings from 5:00 PM to bedtime. Smith et al. (2019) provides valuable insights into the use of automated wearable cameras and highlights several limitations that should be considered when using cameras to measure screen use.

As automated wearable camera research is a relatively new approach to measuring children's health behaviours there are no preexisting criteria for feasibility or acceptability. Smith et al. (2019) measured feasibility through the quality of the captured images, wear time, and the ability to code contextual factors, such as screen type and location. Acceptability was evaluated based on participant adherence to the protocol and feedback from exit interviews. They found that participants spent nearly half of their evening using screen-based media devices, with evidence of media multitasking for approximately 5% of the captured time. The most common scenario of media multitasking was using mobile phones with a television or laptop in the background. Additionally, they found that participants switched between screen-based media devices (e.g., a mobile phone to a television) or from a screen-based media device to no screen-based media device 10 times per hour. They also reported that they were able to identify the location of the screen use, where they found that nearly all television viewing was located in the living room, while laptop use often occurred in bedrooms. These results indicate that automated wearable cameras can capture screen use behavioural patterns and their context. However, their limited sample size raises concerns about the generalisability of the findings. Further research is needed to investigate the use of automated wearable cameras to measure screen use in a larger, more diverse sample.

Smith et al. (2019) also reported several limitations when using the cameras to measure screen use. First, the images blurred when the participant moved. The blurry images suggest that automated wearable cameras may be limited to only capturing sedentary screen use, not active screen use. This may limit automated wearable camera's ability to capture different types of screen use. Next, they found it challenging to code the content on mobile phones due to poor image resolution. Coder's inability to confidently identify content on mobile phones may limit the ability to code smaller types of screen-based media devices. Thus, automated wearable cameras may not be able to accurately capture the content on mobile phones, which is a common screen-based media device used among children and adolescents (Thomas, 2019). Finally, they found that wear time decreased from 78% on the first evening to 51% on the final evening, suggesting that participant compliance with wearing the camera may decrease over time. As automated wearable camera research is a new approach to measuring children's health behaviours there are no preexisting wear time criteria. However, maintaining high levels of adherence to protocols is essential for collecting reliable data (Kimberlin & Winterstein, 2008). The observed decrease in automated wearable camera wear time suggests that adherence to camera protocols needs to be improved.

The findings from the study indicate that automated wearable cameras may be a feasible and acceptable method for capturing the environmental context of adolescents' pre-bedtime screen behaviour. However, challenges such as image blurriness, poor resolution, and a decline in wear time should be considered when using automated wearable cameras to measure screen use.

Validity and Reliability of Automated Wearable Camera Measurements

A limitation of previous automated wearable camera studies is that there is limited research on the validity of automated wearable cameras in paediatric populations. The majority of studies published in this field are feasibility or pilot studies (Maddison et al., 2019), which have only investigated the reliability of automated wearable cameras (Beltran et al., 2018; Cowburn et al., 2016; Kelly et al., 2012; Smith et al., 2019). A recent scoping review investigating the use of automated wearable cameras to assist with the self-management of chronic disease and capture lifestyle behaviours in child and adult populations did not identify any studies investigating the validity of automated wearable cameras in child populations (Maddison et al., 2019). Throughout my literature review, I was only able to identify one study that investigated the validity of automated wearable cameras in a child population (Everson et al., 2019). In this study, the authors used automated wearable cameras to assess the concurrent validity of a health and lifestyle behaviours tool (Child's Health and Activity Tool; Everson et al., 2019).

While an instrument can be feasible and have high reliability, it does not equate to the instrument being valid (Kimberlin & Winterstein, 2008). Feasibility refers to how practical a measurement instrument is, and reliability relates to the consistency of a measure; however, validity relates to the accuracy of the measurement instrument. Therefore, it is possible for a measurement instrument with poor validity to be feasible and reliable. Further, there are many types automated wearable cameras, and each type of device are unique in physical characteristics (e.g., size or weight), features (e.g., image quality), and camera epoch. Consequently, no assumption regarding validity can be made from one device to another. Thus, the validity of automated wearable cameras needs to be examined for the study results to be credible (Sullivan, 2011). This thesis aims to examine the current evidence on the validity and reliability of automated wearable cameras in child and adolescent populations and investigate the validity of automated wearable camera measurements of children's screen use.

The research presented in this thesis builds on the skills and knowledge I developed when completing my Honours degree at the Australian Catholic University. For my Honours' thesis, I developed and tested the inter-rater reliability of a coding framework for coding images from automated wearable cameras to classify screen use. The framework assessed the type of device, the content of the screen, and both the physical (e.g., where the image was taken) and the social (e.g., who the child was with) environments. I then examined the extent to which independent coders could apply the framework consistently. The development of this framework was an important first step in investigating the validity and reliability of automated wearable measurements on children's screen use. This thesis aims to build on this research and contribute empirical evidence on using automated wearable cameras to measure children's screen use and whether estimates of screen use from device-based measurements are similar to self-report measurements.

Research Aims

The primary purpose of this thesis was to examine the use of automated wearable cameras to establish a more accurate measure of children's screen use. The secondary purpose was to examine whether estimates of screen use from device-based measurements (i.e., automated wearable cameras) are similar to self-report measurements.

Research Questions

To address the aims of this thesis, I present three studies that address the following questions:

1. Which health behaviours have automated wearable cameras been used to study in child and adolescent populations?
2. What is the evidence on the validity and reliability of using automated wearable cameras to capture children's and adolescent's health behaviours?
3. What are the methodological procedures for the instrument administration, device, wear time, coding protocol, data management and participant privacy when using automated wearable cameras to assess health behaviours in children and adolescent populations?

4. What is the convergent validity of automated wearable camera measurements for assessing screen-based behaviours in children aged 8-11 years old in a home setting compared to direct observation?
5. How does the camera epoch length impact estimates of screen use when using an automated wearable camera?
6. Are estimates of screen use from automated wearable camera measurements similar to self-report measurements?

Outline of the Thesis

- Chapter 1 (current chapter) provides an overview of the current knowledge and research on screen use and automated wearable cameras, which will establish the frameworks used for this thesis. This chapter also outlines the aims, objectives, and significance of each study contributing to this thesis.
- Chapter 2 presents the results of the systematic review (Study 1) of academic literature on the use of automated wearable cameras to measure health behaviours in child and adolescent populations.
- Chapter 3 details the development of the methodological procedures used in this thesis, including a brief description of the development of the original coding protocol and image processing (i.e., presented in my Honours' thesis) and the process of refining the coding protocol and image processing for the research presented in this thesis. This chapter also highlights ethical considerations for handling image-based data, and discusses the challenges of developing the coding protocol.
- Chapter 4 presents the results of the validation study (Study 2), which examined the convergent validity of automated wearable camera measurements for assessing children's screen-based behaviours compared to direct observation. This chapter also explores the impact of camera epoch lengths on screen use measurements.

- Chapter 5 compares automated wearable camera measurements (device-based) of screen use with a self- and proxy-reported measurement of screen use and examines whether estimates of screen use from device-based measurements are similar to self-report measurements (Study 3).
- Chapter 6 consolidates the significant findings from this thesis (all studies) and identifies limitations and recommendations for future research.

Chapter 2: A Systematic Review of Automated Wearable Camera Research to Measure Health Behaviours in Youth

Preface

Automated wearable cameras have been increasingly used in health behaviour research in recent years; however, the use of automated wearable cameras to measure health behaviours among young people is not well known. For this reason, the present study has been extended to include all health behaviours. The purpose of this study is to provide a more comprehensive overview on the current evidence on the use of automated wearable cameras to measure health behaviours in youth. This study will build on the growing body of literature in automated wearable camera research and will allow for the remaining studies in this thesis to focus more accurately on validating automated wearable cameras to measure children's screen use.

Introduction

A variety of health behaviours, including diet, physical activity, and sleep are critical for health and well-being. But, accurate measurement of these behaviours is challenging, particularly in young people. Much of the research on children's health behaviours relies on proxy- or self-reported questionnaires (Hale & Guan, 2015; Stiglic & Viner, 2019), which can be unreliable and prone to bias (Lubans et al., 2011). While measurement devices, such as heart rate monitors, accelerometers, and pedometers, all provide important measurements of some behaviours; these devices cannot account for the context surrounding these behaviours, such as where or with whom a behaviour occurs. Understanding the context of behaviours is essential as it influences how and why certain behaviours take place. Moreover, the measurement of context may facilitate the development of successful interventions by identifying specific-context factors that can be targeted for interventions (Short & Mollborn, 2015). While some methods, such as direct observation, can measure context, they are time-consuming, expensive, or impractical (Hardy et al., 2013).

Automated wearable cameras may offer a new solution that provides an accurate measurement that can account for context but with relatively little respondent burden. Automated wearable cameras are typically worn on the chest or collar and take first-person point-of-view images at fixed intervals (Doherty, Hodges, et al., 2013). These cameras were first developed in the field of digital lifelogging (Kelly et al., 2011). Since then, human behaviour research has increasingly used automated wearable cameras as the cameras have become more affordable, smaller, and able to capture data for a longer period of time (Doherty, Hodges, et al., 2013). While a multitude of health behaviours have been studied, such as sedentary behaviour and physical activity (Doherty et al., 2012; Kerr et al., 2013; Leask et al., 2015), screen use (Smith et al., 2019), and diet (Gemming, Utter, et al., 2015), to date no review has collated the existing evidence on the use of automated wearable cameras among young people. The purpose of this review was to provide an overview of the current evidence on the use of automated wearable cameras to measure health behaviours in youth to answer the following questions:

1. **Research Question 1:** Which health behaviours have automated wearable cameras been used to study in child and adolescent populations?
2. **Research Question 2:** What is the evidence on the validity and reliability of automated wearable cameras to capture child and adolescent health behaviours?
3. **Research Question 3:** What are the common methodological procedures when using automated wearable cameras to assess health behaviours in child and adolescent populations?

Methods

Protocol Registration

I prospectively registered this systematic review with PROSPERO (#CRD42021213532) and reported the findings according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement (Page et al., 2021).

Eligibility Criteria

I included articles if they: (1) had a sample with a mean age between 5 and 18 years old; (2) measured at least one health behaviour on one or more occasions using an automated wearable camera; and (3) reported qualitative or quantitative findings. I chose the age range of 5 to 18 years to capture critical development stages, from early childhood through to adolescence (Rosen et al. 2014). I excluded articles with children under 5 years due to this age groups distinct ethical and practical challenges, such as participant safety, comprehension, and camera positioning. For the purpose of this study, I followed Gochman's (1997) definition of health behaviours when assessing study eligibility (i.e., "behavioural patterns, actions and habits that relate to health maintenance, to health restoration and to health improvement"; p. 3).

I excluded articles if they: (1) only used active image capture (i.e., participants had to take the photo manually); or (2) were non-empirical articles such as expert opinions, editorial letters, and review articles. Published and unpublished studies, including dissertations, were eligible for inclusion. There were no time restrictions or language restrictions. The search strategy was only conducted in English. Consequently, this may have restricted the inclusion of articles written in non-English languages in the search strategy.

Information Sources and Search

In order to minimise subjectivity and maximise sensitivity in the search strategy, I used an 'objective approach' (Hausner et al., 2015, 2016, 2012), deriving the search terms from a broad initial set of target articles. To do this, I conducted a preliminary search to identify previous reviews on automated wearable cameras and health behaviour research. I located seven previous reviews (Allé et al., 2017; Bell et al., 2020; Bell et al., 2017; Boushey et al., 2017; Gemming, Utter, et al., 2015; Maddison et al., 2019; Silva et al., 2018); however, only three of the articles included in these reviews included child or adolescent samples (Cowburn et al., 2016; Kelly et al., 2012; Signal, Smith, et al., 2017). All other studies focused on adult samples. To identify additional articles, I conducted a precise search strategy using concepts such as "wearable cameras" or "SenseCam" or "Autographer" and "children" or "adolescents" in Google Scholar, PsycINFO, CINAHL and MEDLINE. I then conducted bidirectional screening using these articles to generate a set of 29 articles. These articles were randomly divided into a development set (14 studies) and a validation set (15 studies). The development set was entered into word frequency analysis software (Clark et al., 2020). Terms that were present in at least 20% of the articles in the development set were selected to develop the search strategy. The initial draft of the search strategy identified 13 of the 14 articles in the development set when searching in CINAHL, PubMed and PsycINFO. I refined this search strategy to identify all articles in the development set. I then tested the search strategy against the validation set. The final search strategy identified all of the articles in the validation set and was then adapted for each database used in the searches. Further details on the development of the search strategy are available in Appendix A.

I searched PsycINFO, Scopus, CINAHL, ProQuest Dissertations and Theses Global, Web of Science Core Collection, ACM Digital Library, PubMed and SPORTDiscus databases in October 2020. I searched ClinicalTrials.gov, the EU Clinical Trials Register and the Australian New Zealand Clinical Trials Registry (ANZCTR) in December 2020. I searched the reference lists of included studies, relevant reviews, and conference papers for eligible studies. I also asked the authors of the included studies for any additional and grey literature that may be eligible for inclusion. Six out of fifteen of the contacted authors responded.

Study Selection

All search results were first imported into EndNote (EndNote X9.3.1), duplicates were removed, and then uploaded to Covidence systematic review management software (Veritas Health Innovation, Melbourne, Australia). Two independent reviewers, including myself, screened titles and abstracts. Reviewers then independently screened full-text articles for eligibility in duplicate. I resolved discrepancies between reviewers by discussion. If we could not reach a consensus by discussion, we consulted a third reviewer.

Data Extraction

Following the full-text screening another reviewer and I extracted data from the included studies using a custom data extraction form. I resolved discrepancies between reviewers by discussion and re-examination of the study’s full text. Data extracted included first author, year of publication, country, study design, sample size and characteristics, automated wearable camera details (e.g., model, camera epoch), wear time characteristics, ethical and privacy details, type and time for image coding, data management details, health behaviour examined, measurement property (e.g., reliability, measurement error and validity), methods of measurement, and results. I gathered details on the automated wearable camera (i.e., weight and field of view) from manufacturers' websites that were missing from the included studies.

Quality Assessment

Reviewers independently rated the methodological quality of the included studies in duplicate using the COSMIN Risk of Bias Checklist (Mokkink et al., 2020, 2018; Prinsen et al., 2018; Terwee et al., 2018). I resolved any discrepancies through discussion. The COSMIN Risk of Bias checklist uses a 4-point scale where validity, reliability, and measurement error were rated as very good, adequate, doubtful, or inadequate. I used the 2018 COSMIN Risk of Bias Checklist to assess validity (Mokkink et al., 2018; Prinsen et al., 2018; Terwee et al., 2018). In the pre-registered protocol, I planned to use the 2018 COSMIN checklist for reliability and measurement error. However, during the study an updated version was published which included items more relevant to the assessment of automated wearable cameras (Mokkink et al., 2020). I therefore chose to use the updated version, as a deviation from the pre-registered protocol.

I used the worst-score-counts principle to determine the overall quality of the study for the measurement property (Terwee et al., 2012). For example, the measurement property in a study would be rated as inadequate if the study is scored as inadequate on one of the items. I applied the COSMIN items that were relevant to automated wearable cameras (i.e., Items 1-4 were excluded because they are specific to questionnaires). I assessed the methodological quality of the included studies only if the study assessed the validity, reliability, or measurement error of automated wearable camera measurements (i.e., I did not assess those that only describe the methodological procedures of automated wearable cameras).

Synthesis of Results

I conducted a narrative synthesis to examine the study results. A meta-analysis could not be undertaken due to the heterogeneity across the studies. If data had allowed for a meta-analysis, I would have used a random-effects model to pool reliability and validity results for each health behaviour and coding method. This approach would have provided a more precise estimate of the overall effect size. Additionally, if sufficient data were available, I would have conducted a moderator analysis to explore the influence of different methodological procedures (i.e., wear time and instrument administration) on the validity and reliability of automated wearable camera measurements. The moderator analysis would have offered insights into which methodological factors may impact the validity and reliability of automated wearable camera measurements in youth populations.

I followed COSMIN recommendations and evaluated each measurement property based on the COSMIN taxonomy of measurement properties (Mokkink, Terwee, Patrick, et al., 2010), regardless of the terms used by the authors in the articles. I adapted Mokkink, Terwee, Patrick, et al. (2010) definition of reliability, measurement error, and validity (including construct, content, and criterion validity) to suit automated wearable camera measurement outcomes. Reliability was defined as the extent to which the scores for participants who have not changed are the same for repeated measurements under several conditions, including over time (test-retest), by the same person on different occasions (intra-rater), or by different persons on the same occasion (inter-rater; Mokkink, Terwee, Patrick, et al., 2010). Measurement error was defined as the random and systematic error of a participant's score that cannot be attributed to the true changes in the construct it aims to measure (Mokkink, Terwee, Patrick, et al., 2010). Construct validity was defined as the extent to which scores of an instrument measurement were consistent with the hypothesis (i.e., internal relationships (structural validity), relationships to scores of other instruments that it should theoretically be related (convergent validity), or differences between relevant groups (Mokkink, Terwee, Patrick, et al., 2010). Content validity was defined as the extent to which the content of a measurement instrument adequately reflects the construct it aims to measure (Mokkink, Terwee, Patrick, et al., 2010). Criterion validity was defined as the extent to which the scores of a measurement instrument adequately reflect those of a "gold standard" measurement instrument (Mokkink, Terwee, Patrick, et al., 2010).

I considered parameters of reliability as kappa (k) and intraclass correlation coefficient (ICC). I interpreted a kappa above .80 or ICC above .75 as acceptable, and a kappa of .60 to .79 or ICC .50 to .74 as borderline (Koo & Li, 2016; McHugh, 2012). Measurement error parameters included percentage agreement, change in the mean or mean difference, limits of agreement (LOA), and standard error of measurement (SEM). I considered measurement error outcomes acceptable when the percentage agreement was above 80%, and borderline when between 60% to 79% (McHugh, 2012). I considered parameters of validity as correlations and receiver operating characteristics (ROC). Since the validity of automated wearable cameras is the focus of this systematic review, the cameras were not considered a gold standard measure for criterion validity. Studies that used automated wearable cameras as the criterion measure were assessed as convergent validity.

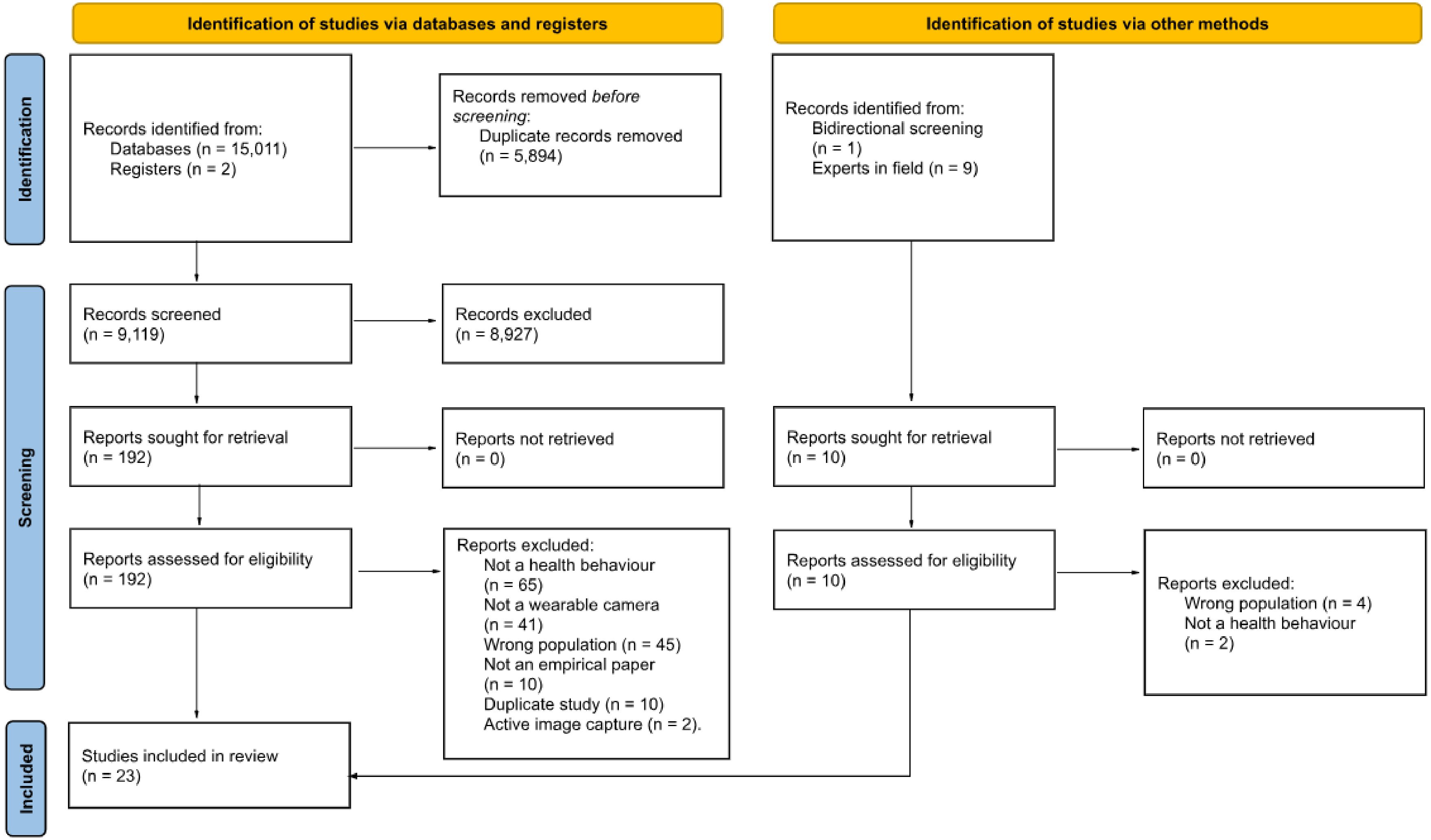
Results

Study Selection

The database and trial registry search identified 15,013 articles. After deduplication, 9,119 articles remained for title and abstract screening, and 192 articles were moved to full-text screening. One article was added from bidirectional screening, and nine articles were added from recommendations from the included authors to full-text screening. In total, 23 articles were included in this review (see Figure 2.1). A list of excluded full-text articles, including the reasons for exclusion, can be found in Appendix B.

Figure 2.1

PRISMA Flow Diagram



Study Characteristics

Details of the included studies are summarised in Table 2.1. The included studies were published between 2012 and 2021, with most (82%) since 2017. One paper had two independent samples of participants, and these were treated as separate studies (Beltran et al., 2016). Seven countries were represented, including New Zealand (52%), the United States (17%) and England (13%). Of the 23 studies included, 11 studies analysed primary data, and 12 studies analysed existing data. The studies were predominantly feasibility and pilot studies (43%) followed by observational (39%), and methodological studies (30%). Sample sizes ranged from 8 to 166. Mean ages ranged from 9.8 (*SD* = 0.4) to 15.8 years; most studies (86%) sampled children (aged 5-13 years).

Table 2.1

Summary of Characteristics of the Include Studies

First author (year)	Study design	Country	Sample size	Population and age	Health behaviour	Automated wearable camera method	Camera wear time	Measurement property assessed
Beltran et al. (2016a)	Pilot	USA	21	Children Mean age: 10.48 (*SD* = 1.24) years	Diet	Cloth strap on the front of the shirt at chest height during a meal	NR	Reliability
Beltran et al. (2016b)	Pilot	USA	10	Children Mean age: 10.67 (*SD* = 1.5) years	Diet	Adjustable lanyard and secured with a magnet placed inside the participant's shirt for one day at home and school	Mean: 9 hours Range: 4-13 hours	Reliability, Measurement error
Beltran et al. (2018)	Feasibility and Reliability	USA	30	Children Mean age: 11.9 (*SD* = 1.27) years	Diet	On the chest for two consecutive days from waking until bedtime.	NR	Reliability, Measurement error
Cowburn et al. (2016)	Feasibility	England	22	Adolescent Age range: 13-15 years	Diet	Lanyard around the neck for four consecutive school days from waking until bedtime	NR	Reliability
Everson et al. (2019)	Validation	Wales	14	Children Mean age: 11 (*SD* = 0.4) years	Diet, screen use, physical activity, sedentary behaviour, travel behaviour, sleep	Lanyard around the neck for two days (one weekend & one school day) from waking until bedtime	Mean: 10.36 (*SD* = 3.16) hours per day Mean placement time: 07:37 (*SD* = 0:27) a.m.	Validity

First author (year)	Study design	Country	Sample size	Population and age	Health behaviour	Automated wearable camera method	Camera wear time	Measurement property assessed
							Mean removal time: 6:17 (*SD* = 3.15) p.m.	
Freeman et al. (2020)	Observation	New Zealand	74	Children [a] Mean age: 12.6 years	Green space recreation	Lanyard around the neck for four consecutive days (Thursday to Sunday) from waking until bedtime	Mean: 7.2 leisure hours per participant	N/A
Gage et al. (2017)	Method and Reliability	New Zealand	100	Children [a] Age range: 10-14 years	Sun-protective behaviour	Lanyard around the neck for four consecutive days (Thursday to Sunday) from waking until bedtime	NR	Reliability, Measurement error
Gage et al. (2018)	Observation	New Zealand	12	Children [a] Age range: 11-13 years	Sun-protective behaviour	Lanyard around the neck for four consecutive days (Thursday to Sunday) from waking until bedtime.	NR	N/A
Gage et al. (2018)	Observation	New Zealand	15	Children [a] Age range: 11-13 years	Sun-protective behaviour	Lanyard around the neck for four consecutive days (Thursday to Sunday) from waking until bedtime	NR	N/A
Gage et al. (2019)	Feasibility	New Zealand	15	Children [a] Age range: 11-13 years	Sun-protective behaviour	Lanyard around the neck for four consecutive days (Thursday to Sunday) from waking until bedtime	NR	Measurement error
Gage et al. (2020)	Observation	New Zealand	158	Children [a] Mean age: 12.6 (*SD* = 0.5) years	Diet	Lanyard around the neck for four consecutive days (Thursday to Sunday) from waking until bedtime	Mean: 8.7 (95% CI [9.0, 9.4]) hours per participant. Wear time was higher for Thursdays (6.1 hours, 95% CI [5.7, 6.5]) than	Measurement error

First author (year)	Study design	Country	Sample size	Population and age	Health behaviour	Automated wearable camera method	Camera wear time	Measurement property assessed
							Saturdays (2.7 hours, 95% CI [2.4, 3.1]).	
Hänggi et al. (2020)	Method and Reliability	Switzerland	14	Children Mean age: 10.3 (*SD* = 0.6) years	Sedentary behaviour	Around the neck for seven consecutive days during leisure time out of school.	Mean: 6.2 (*SD* = 1.5) hours per day	Reliability
Kamar et al. (2019)	Feasibility	England	10	Adolescent Median age: 13.5 years	Diet	Around the neck for three days from waking until bedtime, both within and outside of school hours depending on school consent.	NR	N/A
Kelly et al. (2012)	Feasibility and Validation	England	17	Adolescent Age range: 13-16 years	Travel behaviour	Around the neck for one week to and from school	Wear time was separated into duration for journey stages. Mean: 13.48 minutes per journey stage	Reliability, Measurement error, Validity
McKerchar et al. (2020)	Feasibility	New Zealand	37	Children [a] Age range: 11-13 years	Diet	Lanyard around the neck for four days from waking until bedtime	NR	Measurement error
McKerchar et al. 2020)	Observation	New Zealand	37	Children [a] Mean age: 12.6 years	Diet	Lanyard around the neck for four days from waking until bedtime	NR	Measurement error

First author (year)	Study design	Country	Sample size	Population and age	Health behaviour	Automated wearable camera method	Camera wear time	Measurement property assessed
Narayanan et al. (2020)	Validation	New Zealand	15	Children Mean age: 10 (*SD* = 2.6) years	Sedentary behaviour and physical activity	Lapel for two hours	2 hours	N/A
Pearson et al. (2017)	Observation	New Zealand	166	Children [a] Age range: 11-13 years	Blue space recreation	Worn for four days (Thursday to Sunday) from waking until bedtime.	NR	N/A
Raber et al. (2018)	Feasibility	USA	31	Children Age range: 9-13 years	Diet	Collar for two days from waking until bedtime	NR	Reliability, Measurement error
Robinson et al. (2017)	Observation	Tonga	72	Children [b] Age range: 10-13 years	Activities of daily living	Lanyard around the neck for three consecutive days (Friday to Sunday) from waking until bedtime	NR	Measurement error
Smith et al. (2019)	Feasibility	New Zealand	15	Adolescent Mean age: 15.8 years	Screen use	Adjustable lanyard on their upper chest secured with Velcro for three days (two weekdays, one weekend day) at home from 5:00PM until ready to sleep.	Mean: 267 (*SD* = 97) minutes per evening. Range: 54-419 minutes per evening. Wear time decreased over time (296 minutes per evening on Day 1 compared to 244 minutes on Day 3).	Reliability, Measurement error

First author (year)	Study design	Country	Sample size	Population and age	Health behaviour	Automated wearable camera method	Camera wear time	Measurement property assessed
							Mean time of first image: 5:36 (*SD* = 1:04) p.m. Mean time of last image: 10:04 (*SD* = 1:07) p.m.	
Smith et al. (2019)	Observation	New Zealand	158	Children [a] Mean age: 12.6 (*SD* = 0.5) years	Diet	Worn for four days (Thursday to Sunday) from waking until bedtime.	Mean: 10 hours per day Placement range: 6:00 a.m. to 12:00 a.m.	Measurement error
Veatupu et al. (2019)	Observation	Tonga	35	Children [b] Mean age: 10.7 years	Diet	Lanyard around the neck for three consecutive days (Friday to Sunday) from waking until bedtime	Mean: 10 hours per day	Measurement error
Zhou et al. (2019)	Feasibility and Validation	China	52	Children Mean age: 9.8 (*SD* = 0.4) years	Diet	Collar with a metal clip and elastic lanyard for seven consecutive days from waking until bedtime.	Median duration weekdays: 13.0 (IQR 12-13.8) hours per day Median duration weekends: 10.5 (IQR 9.1–11.9) hours per day	Measurement error, Validity

Note. NR, not reported; N/A, not applicable; IQR, interquartile range.

[a] Sample from Kids'Cam study sample.

[b] Sample from Kids'Cam Tonga study sample.

Health Behaviours Assessed by Automated Wearable Cameras

Automated wearable cameras have been used to assess a wide range of health behaviours in children and adolescent populations. The most frequently measured behaviours were related to diet, including food identification, portion size estimations, and beverage purchasing and consumption (Beltran et al., 2016, 2018; Everson et al., 2019; Raber et al., 2018; Cowburn et al., 2016; Kamar et al., 2019; McKerchar, Smith, Gage, et al., 2020; McKerchar, Smith, Stanley, et al., 2020; Smith et al., 2019; Veatupu et al., 2019; Zhou et al., 2019). These behaviours were captured in various locations, including homes, supermarkets and service stations.

Three studies measured screen use behaviours (device and activity type; Smith et al., 2019; Hänggi et al., 2020; Everson et al., 2019). Six studies identified non-screen sedentary behaviours (e.g., reading, writing, playing quietly; Everson et al., 2019; Freeman et al., 2020; Hänggi et al., 2020; Narayanan et al., 2020; Robinson et al., 2017; Smith et al., 2019).

Physical activity (i.e., duration and travel methods) were measured by automated wearable cameras in five studies ((Everson et al., 2019; Freeman et al., 2020; Kelly et al., 2012; Narayanan et al., 2020; Pearson et al., 2017).

Other health behaviours measured by automated wearable cameras included sun-protective behaviours (e.g., third party shade use, sun-protective clothing worn; Gage, Barr, et al., 2018; Gage et al., 2017, 2019; Gage, Leung, et al., 2018), sleep hygiene (Everson et al., 2019; Robinson et al., 2017), and dental behaviour (Everson et al., 2019).

Nearly all of the included studies measured other contextual factors alongside the health behaviours, most commonly the location of the health behaviour (e.g., home, community venue, supermarket) and the social environment (e.g., alone or with other people).

Reliability of Automated Wearable Camera Measurements for Assessing Health Behaviours

Eight studies reported the reliability of the automated wearable camera measurements (Beltran et al., 2016, 2018; Cowburn et al., 2016; Gage et al., 2017; Hänggi et al., 2020; Kelly et al., 2012; Raber et al., 2018; Smith et al., 2019). The methodological quality of the reliability studies included was adequate to very good (see Appendix C). We rated one study as doubtful due to not reporting specific category agreements (Raber et al., 2018).

Inter-rater reliability of the coding protocols was the only type of reliability assessed. There was heterogeneity between coding protocols used in the studies as each protocol was specifically designed for its research purpose. As a result, there are differences between coding protocols which may vary results for the measurement of health behaviours. Five studies used Cohen's kappa (Cowburn et al., 2016; Hänggi et al., 2020; Kelly et al., 2012; Raber et al., 2018; Smith et al., 2019), and four used intraclass correlation coefficients (Beltran et al., 2016, 2018; Gage et al., 2017; Kelly et al., 2012). One study reported the inter-rater reliability of three different raters using a mean kappa (Hänggi et al., 2020). Another study reported the inter-rater reliability of up to four coders using intraclass correlation coefficients (Beltran et al., 2016). Only four studies purposely assessed the inter-rater reliability of the coding protocol (Beltran et al., 2016, 2018; Gage et al., 2017; Hänggi et al., 2020). The remaining studies assessed the inter-rater reliability of the coding protocol as part of a coding quality check (Cowburn et al., 2016; Kelly et al., 2012; Raber et al., 2018; Smith et al., 2019).

Reliability of Dietary Behaviour Measurements

Four studies assessed inter-rater reliability for dietary behaviours (Beltran et al., 2016, 2018; Cowburn et al., 2016; Raber et al., 2018). The results indicate moderate to high levels of agreement among raters when assessing dietary behaviours, however; this varied depending on the diet related behaviour being assessed. My findings suggest that reliability of automated wearable camera measurements can be improved through participant verification interviews and visual estimation techniques, particularly for complex behaviours such as food preparation and portion size estimates.

Beltran et al. (2016) found that inter-rater reliability was higher when dietitians visually estimated portion sizes from images compared to using a 3-dimensional wire mesh (.98 and .53, respectively). The inclusion of participant verification interview data slightly improved the reliability for visual estimation (ICC = .98) but only marginally for the 3-dimensional wire mesh (ICC = .62). Similarly, Beltran et al. (2018) reported a borderline ICC of .67 between two dietitians when estimating calories in food portion sizes, which decreased when assessing inter-rater agreement between dietitians and child-parent verification interviews. The authors recommended that researchers complete a following-day verification interview with the child and parent to ensure the estimates were complete.

Cowburn et al. (2016) found an acceptable kappa statistic between food diaries and image-based assessment for food purchase (k = .96), though the agreement for food consumption was borderline (k = .77). Raber et al. (2018) reported a borderline kappa statistic of .67 between two coders for various food-related activities, including meal preparation and food selection. Most of the disagreements in coding centred around the 'observing' category for food preparation of adults and food selection by peers.

Reliability of Sedentary Behaviour Measurements

Two studies assessed the inter-rater reliability for sedentary behaviours (Hänggi et al., 2020; Smith et al., 2019). Sedentary behaviour was broadly defined as any waking behaviour characterised by having low energy expenditure (≤1.5 metabolic equivalents; Tremblay et al., 2017). I found that automated wearable cameras can reliably categorise sedentary behaviours (screen-based and non-screen based), particularly when assessing the location of the behaviour. However, the reliability for specific sedentary behaviour categories, such as reading, writing and playing quietly, can vary, suggesting that further refinement of coding protocols may be needed.

Hänggi et al. (2020) developed a protocol to categorise children's automated wearable camera data into the sedentary behaviour components for screen-based sedentary behaviour (e.g., television, computer, gaming console, mobile phone, and other screen use), non-screen sedentary behaviour (e.g., reading, writing, playing quietly, relaxing, and personal care) and the location of the behaviour (e.g., nature, home, school, and shops). They reported an overall acceptable inter-rater reliability between three coders across all categories (k = .85, 95% CI [.83, .87]), with particular high agreement for the location of the behaviour (k = .91, 95% CI [.88, .93]). However, the agreement was borderline for screen-based sedentary behaviour categories (k = .72, 95% CI [.62, .82]) and non-screen sedentary behaviour (k = .69, 95% CI [.65, .72]).

Smith et al. (2019) conducted a feasibility study using automated wearable cameras to measure pre-bedtime screen use in adolescents. Smith et al. (2019) reported acceptable inter-rater reliability between coders for identifying the type of screen-based media device (k = .81, $p < .001$) and location of the screen use (k = .85, $p < .001$).

Reliability for Other Behaviour Measurements

One study assessed the inter-rater reliability of adolescents' travel behaviour to school (Kelly et al., 2012). Kelly et al. (2012) reported an acceptable kappa statistic of 1.0 between coders when assessing journey mode and an acceptable ICC of .98 (95% CI [.985, .992]) between coders when assessing journey duration.

Another study assessed the inter-rater reliability between coders for two methods for assessing sun protection of clothing coverage (Gage et al., 2017). Gage et al. (2017) reported that both methods had acceptable ICC with the CAP field observation method having a greater score compared to the Lund & Browder Chart method (ICC = .992, 95% CI [.990, .994] compared to ICC = .971, 95% CI [.961, .979]).

There were no differences between the reliability estimates of children or adolescents.

Measurement Error of Automated Wearable Cameras for Assessing Health Behaviours

Thirteen studies reported the measurement error of automated wearable camera measurements for assessing health behaviours (Beltran et al., 2016, 2018; Gage et al., 2020, 2017, 2019; McKerchar, Smith, Gage, et al., 2020; McKerchar, Smith, Stanley, et al., 2020; Raber et al., 2018; Robinson et al., 2017; Smith et al., 2019; Smith et al., 2019; Veatupu et al., 2019; Zhou et al., 2019). I rated one study as very good methodological quality (Gage et al., 2017), three as adequate (Beltran et al., 2016, 2018; Smith et al., 2019), two as doubtful (Gage et al., 2019; Raber et al., 2018), and seven as inadequate (Gage et al., 2020; McKerchar, Smith, Gage, et al., 2020; McKerchar, Smith, Stanley, et al., 2020; Robinson et al., 2017; Smith et al., 2019; Veatupu et al., 2019; Zhou et al., 2019). Most of the inadequate scores were due to the authors only reporting percentage agreement thresholds (McHugh, 2012). Inter-rater agreement (i.e., percentage agreement) of the coding protocol was the most common measurement error assessed, used in 11 studies (Beltran et al., 2016, 2018; Gage et al., 2020; McKerchar, Smith, Gage, et al., 2020; McKerchar, Smith, Stanley, et al., 2020; Raber et al., 2018; Robinson et al., 2017; Smith et al., 2019; Smith et al., 2019; Veatupu et al., 2019; Zhou et al., 2019). Four studies assessed the measurement error of automated wearable camera measurements compared to other methods (Beltran et al., 2016, 2018; Kelly et al., 2012; Zhou et al., 2019). In this section, I have not discussed the results from studies where percentage agreement was reported as a threshold (e.g., coders achieved 90% concurrence with model answers) or when percent agreement was reported alongside kappa statistics for inter-rater reliability.

Measurement Error of Dietary Behaviour Measurements

Two studies compared automated wearable camera measurements of dietary behaviours to verification interviews (Beltran et al., 2016, 2018). Overall, my findings from this review suggest that while automated wearable cameras can provide useful data on dietary behaviours, their accuracy improves when supplemented with participant verification interviews.

Beltran et al. (2016) reported an acceptable percentage agreement for food identification in images before conducting participant verification interviews. However, when compared to child verification interviews, the agreement dropped significantly, with coders correctly identifying only 50.7% of the food items.

In a follow-up study, Beltran et al. (2018) found a borderline percentage agreement of 60.5% between two dieticians’ coders for food identification in images. The dietician’s codes and child-parent verification interviews improved to a 77% agreement with the food identified in the image. The dieticians identified 5.4% of food that the child-parent verification interviews could not identify, while child-parent verification interviews identified 12.4% of food items that the dietician could not identify in the images.

For calorie intake, Beltran et al. (2018) found no bias between dieticians reporting of calorie intake; however, found a small bias in calorie intake reported by the dieticians when compared to the intake estimated from the child-parent verification interviews (-42.5 kcal mean difference compared to 287.8 kcal mean difference, respectively).

Validity of Automated Wearable Camera Measurements for Assessing Health Behaviours

Three studies reported the convergent validity of automated wearable camera measurements for assessing health behaviours (Everson et al., 2019; Kelly et al., 2012; Zhou et al., 2019). I rated the methodological quality of two studies as very good (Zhou et al., 2019) and adequate (Kelly et al., 2012). I rated one study as inadequate due to using inappropriate statistical methods (Everson et al., 2019), as the author's used percentage agreement to examine the convergent validity between automated wearable cameras and a web-based questionnaire. Two studies used automated wearable cameras as a proxy for direct observation to validate another measurement instrument (Kelly et al., 2012; Zhou et al., 2019). Overall, my results suggest that automated wearable cameras can be a valuable tool for validating other measurement instruments, particularly when used as a proxy for direct observation. However, there is limited research on validating automated wearable cameras measurements itself.

Zhou et al. (2019) used automated wearable cameras to investigate the convergent validity of a 3-day dietary recall by comparing results obtained from the dietary recall with and without the assistance of the images captured by the cameras. They found that dietary recall without camera assistance (i.e., recall without viewing the captured images) was lower than dietary recall with camera assistance. However, there were strong correlations between the two methods, indicating that camera assistance improved 3-day dietary recall accuracy.

Kelly et al. (2012) used automated wearable cameras to investigate errors in adolescents' self-reported school travel journey duration. They found strong correlations between methods for both within-subject and between-subject journey duration ($r = .89$, 95% CI [.84, .93], and $r = .92$, 95% CI [.79, .97], respectively), indicating that automated wearable cameras provide a reliable method of validating self-reported travel data.

Device and Instrument Administration

Across the studies reviewed, seven different types of cameras were used, with the most common being the Autographer. Camera weight ranged from 19 grams (Narrative Clip 2) to 101 grams (Brinno TLC120). Camera field of view ranged from 86 degrees (Narrative Clip 2) to 136 degrees (Autographer). Camera epochs ranged from 4- to 30-seconds, with most studies using a 7-second camera epoch (56%). One study used video recording instead of images (Narayanan et al., 2020). I did not find any evidence on the most accurate camera epoch.

Most studies had the participants wear the camera around their neck on a lanyard (69%). Other studies had participants wear the cameras on their shirt collar or lapel (17%), or on their chest attached with a cloth strap (8%). However, wearing the camera on a lanyard often led to excessive movement, necessitating more secure attachment methods (Beltran et al., 2016, 2018; Everson et al., 2019; Smith et al., 2019). Cameras attached at collar bone level (e.g., on the shirt collar) had to be adjusted according to the participant's height and was unsuitable for capturing dietary behaviours (Beltran et al., 2016; Raber et al., 2018). I did not find any evidence on the most effective attachment method.

Six studies reported issues with the automated wearable cameras' battery life (Beltran et al., 2016; Everson et al., 2019; Kamar et al., 2019; Narayanan et al., 2020; Veatupu et al., 2019; Zhou et al., 2019). The battery life of the devices varied according to the device used and the frequency of the camera epoch (Hänggi et al., 2020). In most cases, the battery capacity was insufficient for a whole day of data collection. To mitigate this, participants were asked to charge the cameras overnight or during periods when they did not need to wear the camera. One study gave each participant two cameras and changed the camera twice per day (Zhou et al., 2019).

Wear time

Eleven studies reported the participants' camera wear time (Beltran et al., 2016; Everson et al., 2019; Freeman et al., 2020; Gage et al., 2020; Hänggi et al., 2020; Kelly et al., 2012; Narayanan et al., 2020; Smith et al., 2019; Smith et al., 2019; Veatupu et al., 2019; Zhou et al., 2019). For whole-day assessments, the average camera wear time ranged from 8.7 hours to 13 hours, from as early as 6:00 a.m. to removing the camera as late as midnight. One study investigating evening screen use in adolescents found that average wear time decreased over the study period from 4.9 hours on the first day to 4.0 hours on the third day of data collection (Smith et al., 2019).

Data Management and Coding Protocols

The total number of images collected in a study ranged from 719 to 739,162 images, with 3% to 35% of images deemed uncodable. Images were uncodable from being blurry, dark, blocked (i.e., camera lens blocked by hair), and unsuitable camera placement (i.e. camera recorded the ceiling when lying down). Common reasons for missing data included camera malfunctions (i.e., camera not fully charged) and user error (i.e., the participant forgetting to press record).

All studies used manual coding protocols to analyse the images. The average time for coding images varied depending on the health behaviour analysed, the amount of detail coded from the images, and the number of images required to be coded. Simple coding protocols averaged around 30 minutes per participant (Kelly et al., 2012), while more complex coding protocols ranged from 40 minutes to 9 hours (Beltran et al., 2016; Zhou et al., 2019).

Privacy

To protect the privacy of participants and third parties, several strategies were implemented across studies. The most common procedures to protect participants' and third parties' privacy included camera functions (i.e., switching the camera off or using built-in privacy buttons on cameras), having clear information in booklets for when cameras should be removed or turned off (e.g., public spaces, bathrooms), having informative scripts participants can give to third parties with details of the research, providing participants with the opportunity to view or delete images at the end of the study period, de-identifying images by blurring faces or identifying information, and encrypting SD (Secure Digital) cards to restrict access to the images.

Three studies reported having privacy concerns expressed by participants or parents (Beltran et al., 2016, 2018; Everson et al., 2019). The most common privacy concerns from participants (or parents) included experiencing unwanted attention in public, feeling self-conscious in public when wearing the camera, and other family members not wanting to be filmed. Only one study had a participant drop out due to privacy concerns (Beltran et al., 2018). Another study reported that four images were removed by parents and one by the researcher during data screening due to religious reasons, explicit images, and parent unease (Everson et al., 2019).

Discussion

This systematic review provided an overview of the current evidence on the use of automated wearable cameras to measure health behaviours in youth to (1) identify health behaviours that automated wearable cameras have measured, (2) provide an overview of the evidence on the validity and reliability of wearable camera measurements, and (3) provide an overview on standard automated wearable camera procedures. I identified 23 studies, with most published since 2017. Most studies were feasibility or pilot studies.

I identified a range of health behaviours assessed by automated wearable cameras in youth, with dietary behaviours being the most common (56%). Nearly all of the included studies measured other contextual factors alongside the health behaviours, most commonly the location of the health behaviour and whether the participant was alone or with other people. The inclusion of contextual factors aligns with the recognition in the literature of the importance of ecological validity in health behaviour research (Shiffman et al., 2008). However, the predominance of dietary behaviours raises questions about the broader applicability of automated wearable cameras when measuring behaviours. While automated wearable cameras have been shown to be useful in capturing dietary behaviours, their effectiveness in measuring other health behaviours such as sedentary, physical activity, sleep hygiene, dental, travel, and sun-protective behaviours remains relatively unknown.

Overall, I found acceptable reliability results for travel and sun-protective behaviour and mixed results for sedentary and dietary behaviours. I found acceptable reliability results for screen-based sedentary behaviours (device and activity type) but borderline results for non-screen sedentary behaviours (e.g., reading, writing, relaxing). The difference in reliability results between these sedentary behaviour domains may be due to the increased difficulty in coding specific activities, relying more on the subjective input from the coder. Moreover, only two studies assessed the inter-rater reliability of screen-based and non-screen sedentary behaviours. The differences in the results may be due to different coding protocols. These findings highlight the need for ongoing refinement of coding protocols and improving coder training to improve reliability across different health behaviours. This is consistent with previous research emphasising the importance of standardised procedures in observational studies (Chorney et al., 2015).

Moreover, I found acceptable reliability results for food consumption and purchase behaviours but borderline reliability for portion size estimates. Studies by Beltran et al. (2016, 2018) highlight that dietitians often misidentified foods, suggesting that verification interviews should be used alongside automated wearable camera images when assessing portion sizes and calorie metrics. This approach aligns with best practices in dietary measurement, which advocate for the combination of multiple methods to mitigate the limitations of any single approach (Gemming, Utter, et al., 2015). This limitation suggests that while automated wearable cameras may offer insights into food consumption patterns, they may not be reliable for detailed assessments, such as portion sizes and calorie counts.

Two studies assessed the convergent validity of automated wearable camera measurements, where the camera was used as a proxy for direct observation to validate another measurement instrument (Kelly et al., 2012; Zhou et al., 2019). Travel journey duration to school was found to have strong correlations with self-reported measurements for within-subject and between-subject journey duration (Kelly et al., 2012). Studies using adult samples have reported similar results (Kelly et al., 2014). Camera-assisted dietary recall was found to have moderate to strong correlations with a 3-day dietary recall for energy, macronutrient, micronutrient intake (Zhou et al., 2019). Zhou et al. (2019) reported that daily dietary intakes from dietary recall without camera assistance were lower than daily dietary intakes from dietary recall with camera assistance, suggesting dietary recall was more accurate with camera assistance than without. These findings are consistent with a systematic review that reported image-assisted methods can enhance self-report measurements by revealing unreported food and identifying misreporting errors among adults (Gemming, Utter, et al., 2015).

I found limited privacy or ethical concerns from participants. Researchers often report privacy and ethical concerns about using these devices (Mok et al., 2015). However, I found limited concerns expressed by participants in the included studies. Individuals expressed concern were primarily non-participants during recruitment (Cowburn et al., 2016). Given the importance of participant and third-party privacy, specifically involving young participants, future research should take the necessary precautions to protect individuals' privacy, such as applying established ethical frameworks (Kelly et al., 2013).

My results suggest that certain camera positions may better capture some health behaviours than others. Cameras positioned on the chest were better for capturing a range of screen use behaviours (Hänggi et al., 2020; Smith et al., 2019), while those positioned at the collarbone level were better for capturing food marketing exposures that required a higher camera angle (Raber et al., 2018). The camera placement for other health behaviours (e.g., food consumption) may depend on the participant's height (Beltran et al., 2016). This finding is consistent with studies that emphasise the need for individualised camera position to account for variances in participant's heights to ensure that the camera captures the behavioural contexts (Gemming, Utter et al., 2015).

There is a clear need for devices with a longer battery life. Some cameras' limited battery capacity prevented researchers from collecting a full day of data. Participants were required to charge the cameras at different points in the day, suggesting that some behaviours may have been missed, such as evening meals (Beltran et al., 2016; Veatupu et al., 2019).

All included studies used manual image coding, which was time and resource intensive. Findings from this review revealed that manual image coding can range from around 30 minutes per participant to more than nine hours per participant per day. The current manual techniques limit the size of studies using automated wearable cameras. Other studies on adult populations have developed data algorithms and object classification methodology to reduce the amount of image coding (Doherty et al., 2011; Rosenberg et al., 2017). These studies use machine learning techniques for automated image recognition (Biswas et al., 2017; Krizhevsky, Sutskever, et al., 2015). Additional automated methods need to be developed and tested for wearable camera measurements of young people's health behaviours to allow for larger sample sizes to be collected.

Limitations and Future Directions

A limitation of this review was the small number of studies included. There were limited results on some health behaviours (e.g., physical activity and sedentary behaviours). Therefore, I included studies of low methodological quality. Additionally, nearly half of the included studies analysed data from an automated wearable camera study conducted in New Zealand (Signal, Smith, et al., 2017), increasing the chance of double-counting effects. Furthermore, some studies measured health behaviours as a one-off measurement, which may have introduced measurement error. A single measurement could capture atypical behaviour rather than the child's usual patterns. Additionally, the reliability and validity of automated wearable camera measurements may be compromised due to the lack of repeated data points. As a result, findings from this review need to be interpreted with caution. I could not quantitatively assess publication bias, and given the nature of this research, it is possible that studies showing unfavourable outcomes (i.e., unacceptable reliability or validity) may not have been published.

Another limitation is that most of the included studies were feasibility or pilot studies. As a result, few studies stated an objective to investigate the validity or reliability of the automated wearable camera measurements, and instead, they were primarily included as study quality checks. This finding is consistent with a scoping review, which did not identify any studies investigating the validity of automated wearable cameras to assist with chronic disease the self-management (Maddison et al., 2019). Additional studies should be conducted examining the reliability and validity of automated wearable camera measurements in child and adolescent populations.

There was also substantial heterogeneity between studies. As automated wearable cameras are a relatively new method in health behaviour research, there are no best practice principles regarding methods and reporting. As a result, substantial differences in the methods and results across the studies included, made it difficult to draw broad conclusions regarding the reliability or validity of automated wearable cameras for assessing different health behaviours. Future research should focus on developing clearer guidelines for using automated wearable cameras and the consistency of coding frameworks.

A limitation highlighted by a reviewer was the date of the search strategy. In this review, the information sources were searched in December 2020. To address this limitation, I ran the search strategy in PubMed database in July 2024 to capture articles published after the original search. The updated search identified 541 articles. After screening the articles, I found four articles that were eligible to be included in this review. Of the four articles, two articles assessed dietary behaviours among children and adolescents (Idris et al., 2021; Jobarteh et al., 2023), one article assessed physical activity and sedentary behaviours among adolescents (Andriyani et al., 2022), and one article assessed screen use behaviours among adolescents (Thomas et al., 2022). The results from these eligible articles would not substantially influence the results of this systematic review. For more robust findings, a future systemic review using more recent database search results could be beneficial.

Implications in Research and Policy

Previous research suggests that automated wearable cameras may provide a new solution that provides an accurate measurement of children's health behaviour while accounting for context (Doherty, Hodges, et al., 2013). Findings from this current study suggest that these devices may offer valuable insights into a range of health behaviours, but also highlights implications for policy development. The current focus on dietary behaviours indicates a gap in the broader applications of automated wearable cameras to other health behaviours such as physical activity, sleep, screen use, and sun protection. Moreover, substantial differences in the methods and results across the studies included highlights the need to further refine automated wearable cameras procedures to enhance reliability and validity. Such improvements may produce more reliable data that policymakers can use to make more informed decisions when establishing health policies. An issue in public health interventions is that often a 'one size fits all' approach is used when targeting health behaviours; however, there is little evidence that suggest the current interventions effectively address the complex interplay between health behaviours, illness, and wider determinants of health (Jepson et al., 2010). By incorporating the contextual insights provided by automated wearable cameras, future interventions may be more tailored to individual needs. The shift towards more contextually informed and evidence-based interventions has the potential to improve the effectiveness of public health policies and better address factors that influence children and adolescent health.

Conclusion

This study provided an overview of the current evidence on the use of automated wearable cameras to measure health behaviours and summarised the current evidence on the validity and reliability of automated wearable cameras in child and adolescent populations. My results suggest that automated wearable cameras may provide a reliable method for measuring specific health behaviours, but further studies are needed on the validation of automated wearable cameras in child and adolescent populations. More specifically, my results from this study demonstrated that automated wearable cameras may provide a reliable method for measuring the type of device and location of screen use behaviours. Researchers undertaking future research with automated wearable cameras should consider several factors when using the cameras among youth, including ethical and privacy considerations, image quality, camera placement, battery life and time for manual image coding. Given the challenge of accurately measuring young people's health behaviours, these cameras have the potential to measure multiple health behaviours.

The remainder of this thesis demonstrates how I developed and tested a new method using automated wearable cameras to establish a more accurate measurement of children's screen use. I then examine whether estimates of screen use from this new device-based method are similar to self-reported measurements.

Chapter 3: Developing the Coding Protocol

Introduction

Automated wearable cameras present an opportunity to assess children's screen use more accurately. However, using automated wearable cameras in screen use research depends on our ability to analyse the collected data. In the same way a questionnaire can be used to measure anxiety, coding protocols can be used to measure behaviours captured in image-based data (Chorney et al., 2015). Thus, for automated wearable cameras to be able to measure children's screen use accurately, there must first be a valid and reliable image coding protocol to analyse the collected data (Chorney et al., 2015).

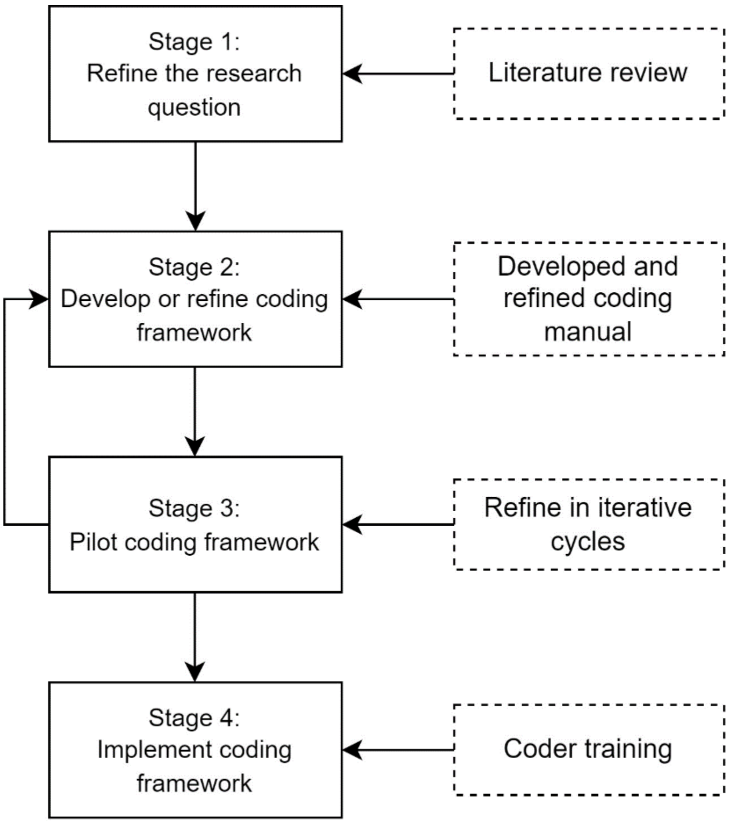
My findings from the systematic review (Chapter 2) demonstrated that automated wearable cameras may provide a reliable method for measuring screen-based sedentary behaviours (type of device, duration, and location of screen use behaviours). However, I found borderline results for non-screen sedentary behaviours (e.g., reading, writing, relaxing). The difference in reliability results between these sedentary behaviour domains may be due to the increased difficulty in coding specific activities or the use of different coding protocols. These findings highlight the need for ongoing refinement of coding protocols and improving coder training to improve reliability when measuring health behaviours (Chorney et al., 2015). Further, for my Honours' thesis I developed and tested the inter-rater reliability of a coding framework for coding images of automated wearable cameras to classify the type of device, content, location, and social environment of children's screen use. I have built on this research to create a comprehensive coding protocol that researchers can use to identify the type of device, content and context of children's screen use captured in automated wearable camera research. In this Chapter, I have demonstrated how I developed and tested the coding protocol and procedures to measure children and adolescents (ages 5-18 years old) screen use, highlighted ethical considerations for handling image-based data, and discussed the challenges of developing the coding protocol.

The Coding Protocol

The original version of the coding protocol (i.e., presented in my Honours' thesis) was developed based on a step-by-step guideline for developing and modifying behaviour coding protocols in paediatric populations (Chorney et al., 2015). I followed this guideline to ensure the coding protocol was created systematically to increase inter-rater reliability and validity, identify all relevant codes, and decrease the chance of bias occurring through the development phase. I followed the same guidelines to refine the coding protocol as part of this thesis (see Figure 3.1). Consistent with the guideline, the coding protocol modification consisted of four stages: (1) refining the research question, (2) refining the coding protocol, (3) piloting the coding protocol, and (4) implementing the coding protocol.

Figure 3.1

Process for Developing the Coding Protocol based on Chorney et al. (2015) Guidelines



Stage 1: Refine the Research Question

The first stage of the image coding protocol development involved refining the research question. In the literature, this is often considered to be the conceptual phase where the construct intended to be measured should be clearly defined (de Vet et al., 2011).

As part of my prior work, I reviewed the literature to determine screen use factors that may influence children's and adolescent's health outcomes. The results from this literature review identified four key factors: device (e.g., the type of device), content (e.g., what the child is viewing), context (e.g., who the child is with and how they interact), and viewing duration (e.g., time spent on the device). Thus, the coding protocol was designed to identify the type of device, duration of the screen exposure (i.e., screen time), content, location, associated behaviours (i.e., co-occurring behaviours) and social setting. There were no changes in the coding protocol's purpose between the original version and the refined version presented in this thesis (i.e., the protocols purpose remains to measure the device, content, and context of children's screen use).

Stage 2: Refining the Coding Protocol

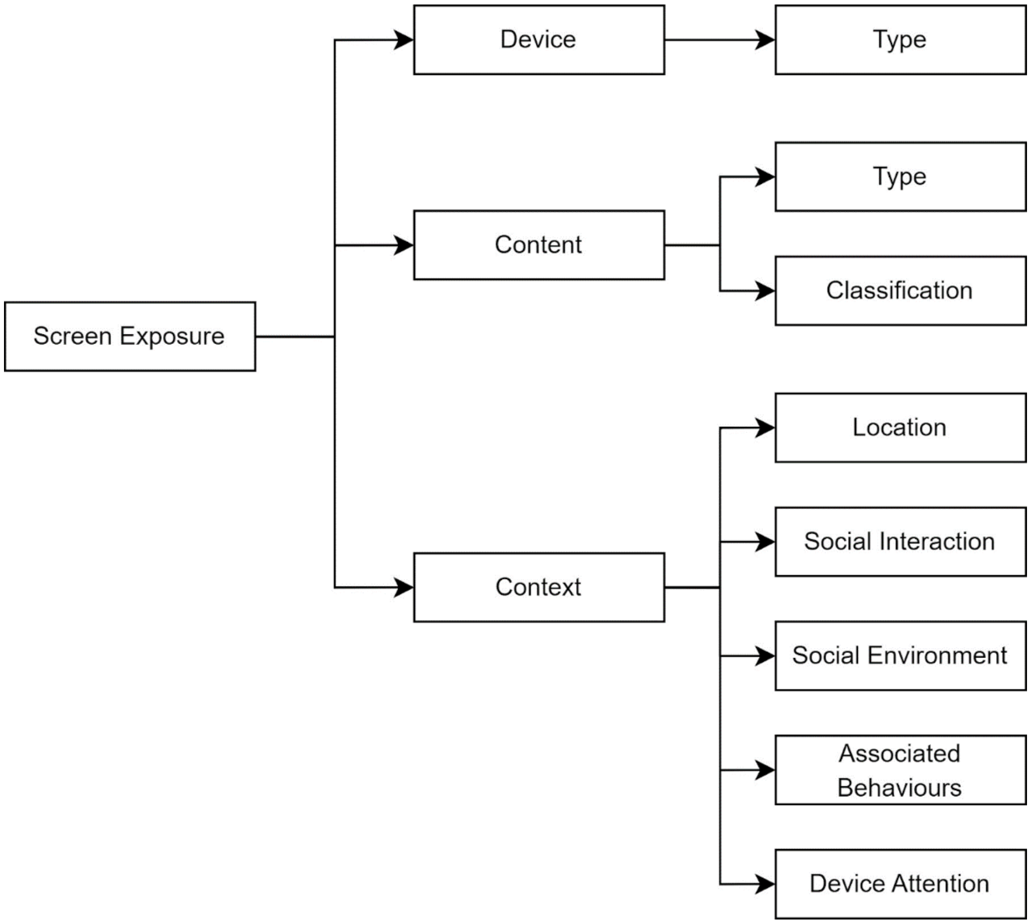
The second stage of the coding protocol development involved developing and refining the coding manual. In this stage, I developed a list of codes and operational definitions and then organised codes into facets, categories, and subcategories within each core dimension (i.e., the device, content, and context) to form the structure of the coding protocol (see Figure 3.2). I used this strategy for three reasons. First, the hierarchical structure makes it easier for the coder to decide on the most appropriate code (Floyd et al., 1998). Second, it allows coders to choose a higher code (e.g., a category code rather than a subcategory code) if they cannot determine the correct subcategory. Third, the structure's flexible nature allows for other researchers to adapt the protocol to suit their research needs.

Further, each category has been organised according to the screen use and screen time definitions presented in Chapter 1 (Table 1.3). For example, the content facet, ‘Content type’, was made up of two categories (passive screen media and interactive screen media) to reflect the definitions in the literature for passive screen time (i.e., passively receiving screen-based media information) and interactive screen time (i.e., cognitively engaging in screen-based media activities). I used the definitions in the literature to ensure the categorisation of screen use and screen time definitions were consistent with the established research. Additionally, the coding structure and operational definitions were discussed with supervisors to confirm that they aligned with the existing literature and were applicable to automated wearable camera measurements.

The codes and structure of the protocol were then refined during testing for this thesis. Some of the most valuable refinements to the coding protocol occurred during testing. In this section, I have outlined the process of developing and refining each category of the coding protocol below.

Figure 3.2

Overview of the Dimensions and Facets of the Coding Protocol



Screen Exposure Definition

An important part of this stage was creating a definition to define when the coders would code an image as screen use. Most research examining screen use do not consider media multitasking. The few studies that have examined the impact of media multitasking have suggested that media multitasking may have adverse effects on learning and sleep (Cain et al., 2016; Calamaro et al., 2009; van der Schuur et al., 2018). For this reason, I have incorporated media multitasking into my screen exposure definition. I defined screen exposure as an event or episode where a person is in the presence of one or more active screen-based media devices, regardless of whether or not the person is consciously attending to the device. An argument against including media multitasking in the definition of screen exposure is that it may have a different impact than traditional screen exposure. To overcome this issue, I added a facet to the coding framework to allow coders to identify the device attention (i.e., the attention the participant appears to be giving the screen-based media device) of the screen-based media device to ensure that the primary device screen exposure (i.e., the device being that appears to capture the majority of the child's attention during the screen exposure) and secondary device screen exposure (i.e., the second device being used during media multitasking) are analysed separately. For example, if a television were on in the background while the participant used a mobile phone, the television would be coded as the secondary device. The mobile phone would be coded as the primary device. I have provided further detail about the device attention facet under the Context section in this chapter.

Type of Device

I created the device categories based on the Screen Based Media Use Scale (SBMUS), a measure of the different types of devices adolescents use (Houghton et al., 2015). The SBMUS includes newer devices such as smartphones and tablets, which are often neglected in previous screen use research (Houghton et al., 2015). I then searched electronic store websites to identify any newer devices not included in the scale. Newer devices in the list that previous screen use measurements have not measured included wearable devices such as smartwatches, digital signage, interactive whiteboards, and projectors. These additional devices were included after discussions with supervisors. However, some of these devices are yet to be observed in my testing. For instance, I have not coded for a smartwatch, digital signage, interactive whiteboard, or projector in the data I have collected for this thesis. Table 3.1 presents the type of device categories and definitions.

Table 3.1

Device Type and Corresponding Definitions

Type	Device	Definition
Non-portable	Television	A device shaped like a box or rectangle with a screen that receives electrical signals and changes them into moving images (Cambridge University Press, n.d.-l). Can stand alone or be mounted to a wall.
	Desktop Computer	A computer that fits on a desk but is not easily moved from place to place (Cambridge University Press, n.d.-b). Has a monitor, keyboard, mouse, and tower.
	Interactive Whiteboard	A large electronic screen linked to a computer. It is often used in classrooms to show information and can be written on or controlled by touching the screen with a finger or special pen (Cambridge University Press, n.d.-d).
	Projector	A device for showing films or images on a projection screen or other surfaces such as walls (Cambridge University Press, n.d.-h). Commonly used in movie theatres and schools.
	Digital Signage	A screen-based media device that is in the form of a small to large billboard composed of LCD, LED, or a similar display system (Rouse, 2014). Includes digital sign boards, interactive direction signboards, electronic menus, billboards, and similar display devices used for displaying visual information, promotional content, and advertisements in public areas (Rouse, 2014).
Portable	Laptop Computer	A computer that is battery operated and has an integrated screen. Indicated by an inbuilt keyboard.
	Mobile Device	A handheld device that can be used as a small computer, connect to the internet, and run applications. Includes

Type	Device	Definition
		smartphones, feature phones and multi-purpose devices (e.g., Apple iPod Touch).
	Tablet	A small, flat computer that is controlled by touching the screen with one's finger or a special pen (Cambridge University Press, n.d.-k). Does not require a keyboard or mouse. Includes e-readers.
	Handheld game console	Portable, self-contained devices that have a built-in screen, game controls and speakers (*Tech Encyclopedia Index*, n.d.)
Wearable	Smartwatch	A watch that has an electronic screen with many of the features of a smartphone or a computer (Cambridge University Press, n.d.-i). Does not include fitness trackers, such as a Fitbit.

Identifying the different types of devices in the images was relatively straightforward. Televisions were identified as a device shaped like a box or rectangle with a screen and could be stand-alone or mounted on a wall (Cambridge University Press, n.d.-l). Typically, televisions were characterised by being located in a living room setting. Figure 3.3 shows an example of an image that would be coded as 'television'.

Figure 3.3

Example of an Image that would be coded as 'Television'

Note. The Australian Catholic University Ethics Committee (2017-317H) approved the dissemination plan for this study, which allows for the use of non-identifiable images.

Desktop computers and laptops were also easy to identify. Due to laptops' portability, desktop computers and laptops were categorised as separate devices. The desktop computer code was used for computers that typically fit on a desk and cannot be easily moved. The laptop code was used for a computer that is battery operated and has an integrated screen. Figure 3.4 shows an example of an image that would be coded as a 'desktop computer'.

Figure 3.4

Example of an Image that would be coded as 'Desktop Computer'

Devices such as smartphones, feature phones and multi-purpose devices (e.g., Apple iPod Touch) were categorised as 'mobile devices'. Initially, I had these categorised separately; however, during testing I observed that these devices performed similar functions in terms of the content the children were consuming. For example, children used these devices to send messages. In some cases, it was also challenging to distinguish between smartphones and multi-purpose devices. In consideration of these factors, I modified the mobile device category to include all handheld devices that can be used as small computers to run applications but were not classified as tablets. Tablet devices were easier to identify due to their larger size compared to smartphones. During testing I observed that most participants used tablets rather than smartphones to play games and engage in educational content. Thus, I categorised tablets separately from other mobile devices. Figure 3.5 shows an example of an image that would be coded as a 'mobile device'. To differentiate the mobile device from a remote control, I reviewed the surrounding images for context cues (i.e., hand position and screen media depicted on the device). Figure 3.6 shows an example of an image that would be coded as a ‘tablet’.

Figure 3.5

Example of an Image that would be coded as ‘Mobile Device’ and ‘Desktop Computer’

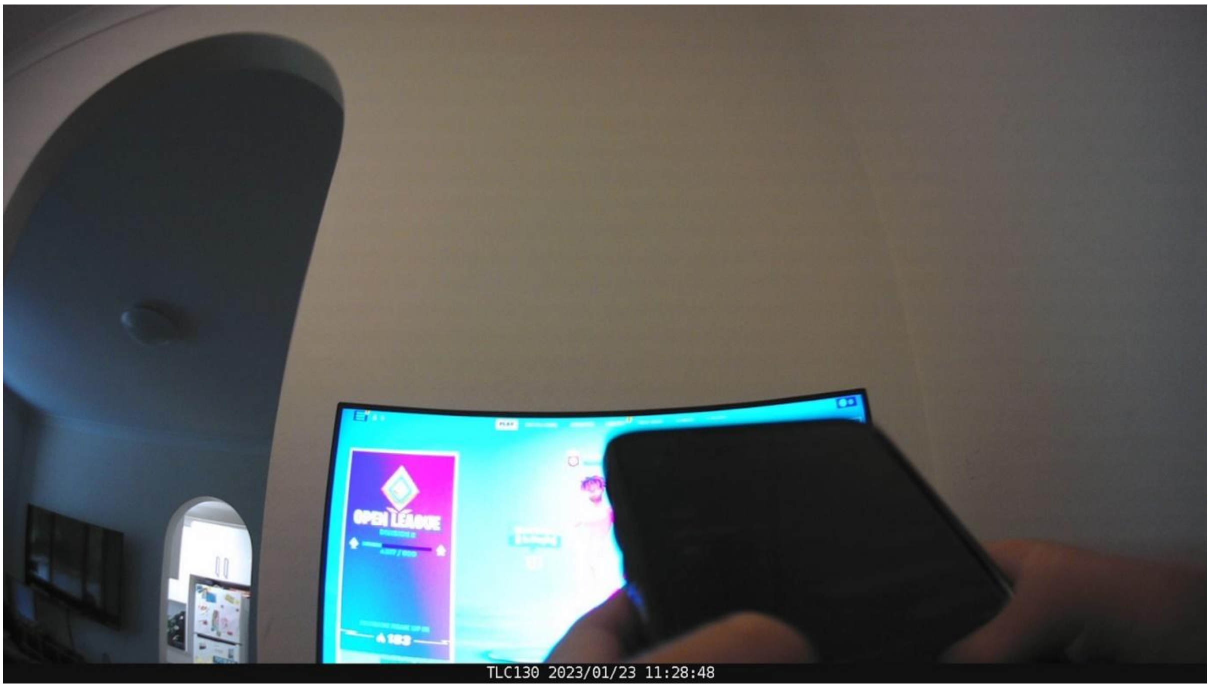
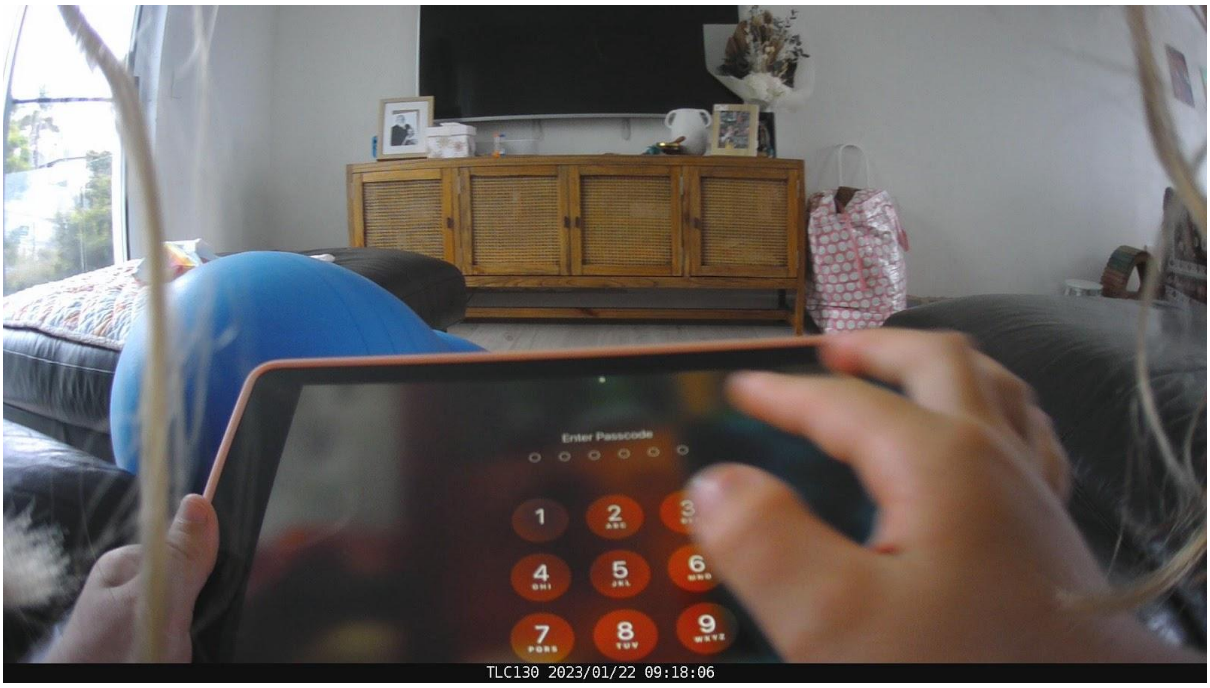


Figure 3.6

Example of an Image that would be coded as 'Tablet'



Handheld game consoles were also categorised separately from mobile devices. They were characterised by having game controls with a built-in screen. Handheld game consoles can be connected to and played through television screens. In those instances, the screen use was coded as television. Figure 3.7 shows an example of an image that would be coded as a 'handheld game console'.

Figure 3.7

Example of an Image that would be coded as 'Handheld Game Console'

During testing, I observed that it was common among children to use multiple devices at once (i.e., media multitasking). To capture this data in the protocol, I adjusted the protocol to code multiple devices at once. For example, if a child was playing on a tablet while watching television, coders could code the tablet and television. I defined media multitasking as having two or more active screen-based media devices visible in the image.

Content

Content refers to the screen media during the screen exposure. The content dimension is made up of two facets: 'Content type' and 'Content classification'. Tables 3.2 and 3.3 present the categories and corresponding definitions for 'Content type' and 'Content classification', respectively.

Content Type. 'Content type' refers to the type of media the child is exposed to during the screen exposure. I created the 'Content type' facet based on the few studies that have measured content or content-related screen-based activities (e.g., using social media or searching the internet) during children and adolescent screen use (Houghton et al., 2015; Rideout, 2015; Sweetser et al., 2012). The ‘Content type’ facet was made up of three categories (passive screen media, interactive screen media and social media), with subcategories in each category. I created these categories and subcategories to include different screen use activities.

‘Passive screen media’ was used to code screen media that required no input or interaction during screen exposure, with the viewer only receiving screen-based information (Sweetser et al., 2012). The ‘Passive screen media’ category has the subcategory ‘Programme’. Initially, I had included more detailed subcategory codes for ‘Programme’ such as coding whether the programme was live-action or cartoon and child- or adult-orientated. However, during testing I found that the inter-rater reliability between coders decreased when coding the detailed subcategory codes. This may have been due to two reasons. First, the detail coders were required to code from the images was labour intensive and could be prone to bias through observer fatigue and drift (Haidet et al., 2009). The risk of bias and measurement error can be increased when observers are required to analyse complex and simultaneously occurring behaviours, which is often seen in children’s screen use (e.g., media multitasking; Haidet et al., 2009). Second, the detail coders were required to code from the images largely relied on a subjective judgement from the coder. For example, to code whether the content was child- or adult-orientated, the coder required pre-existing knowledge of the content programme. If the coder was not familiar with the content the child was viewing, inter-rater reliability most likely decreased. For these reasons, I decided to remove the detailed subcategories and instead had a broad subcategory called ‘Programme’.

‘Interactive screen media’ was used to code screen media that required real-time input from the child during the screen exposure (e.g., playing video games; Sweetser et al., 2012). The ‘Interactive screen media’ category included subcategories such as ‘Internet use, 'Communication', 'Game', and 'Creation'. The subcategory 'Internet use' included browsing the internet. The subcategory 'Communication' included messaging applications and phone calls. The subcategory 'Game' included all types of video games (e.g., game consoles and mobile phone games). The subcategory 'Creation' included screen media used for creative purposes such as photo editors, writing applications and camera applications on mobile phones.

'Social media' was used to code websites, applications and computer programs that allowed users to communicate and share information on the internet (e.g., Facebook, Twitter, and Instagram). The codebook was developed for children and adolescents aged 5-18 years old. 'Social media' category included subcategories such as 'Facebook', 'Instagram', and 'Snapchat' as these were the most popular screen media applications during the coding protocol development. This category was most common on a mobile device. During testing, I observed that the 'Social media' category was challenging to code as it required subjective input from the coder. For example, when coding for 'Social media', coders had to be able to identify the different interfaces of social media applications, such as Snapchat and Instagram. To overcome this challenge, I included examples of popular applications in the coder training sessions. However, while I included examples of these applications in my coder training, it is difficult to have coders remember every screen interface. Therefore, using coders with have prior knowledge of various current technologies may be advantageous to researchers looking to examine screen media.

Table 3.2

Content Type and Corresponding Definitions

Category	Subcategory	Definition
Passive screen media	Programme	Any form of TV Show, movie, or video. Includes online videos (e.g., videos being viewed on YouTube).
Interactive screen media	Internet	Includes all internet-based activities other than those for social media, gaming activity or watching online videos. Browsing is characterised by scrolling through the screen media and searching things up. Includes online shopping, using google and searching.
	Games	Includes when a participant is playing a video game, or watching another person play a video game.
	Creation	Creation screen media refers to visual content on screen-based media devices that has been created by the child (Rideout, 2015).
	Communication	Communication screen media refers to screen media that has the primary purpose of communicating with other people (Rideout, 2015).
Social media	Social media	Websites, applications, and computer programs that allow users to share information on the internet. Includes Facebook, Twitter, Instagram, Snapchat, and blogs (Cambridge University Press, n.d.-l).

Content Classification. 'Content classification' refers to whether the screen media was educational, recreational, or social. I created the content classification facet based on previous studies investigating the impact of content on children's and adolescent's health outcomes (Orben & Przybylski, 2019; Sanders et al., 2019). As I previously highlighted in

Chapter 1, most studies measuring children's screen use do not measure the content viewed on the screen-based media device. The few studies that have investigated the impact of content on health outcomes show that educational, entertainment and social content may have different impacts on health outcomes (Orben & Przybylski, 2019; Sanders et al., 2019). For this reason, the 'Content classification' facet was made up of three categories: 'Educational', 'Recreational', and 'Social'. 'Educational' was used to code screen media with the purpose of educating, informing and enlightening the viewer (Kirkorian & Anderson, 2008). 'Educational' screen media included creation applications, educational video games, programmes, and internet-based activities where it appears that the screen media's purpose is to educate the viewer. For example, screen media was coded as "Educational" if the child was completing homework online. 'Educational' screen media was typically characterised by being viewed on a laptop or tablet with a keyboard attached. 'Recreational' was used to code screen media with the purpose of entertaining the viewer, with no intention of educating, informing or enlightening the viewer (Tremblay et al., 2017). 'Recreational' screen media often included programmes, video games and other internet-based activities where it appeared that the screen media was not being used for educational purposes. For example, in most cases, screen media was coded as 'Recreational' if the child was watching a television programme or playing a video game. 'Social' was used to code screen media with the purpose of communicating with others (Rideout, 2015). 'Social' screen media included social media and communication applications. A limitation of content classification is that the classification of the content is subjective to the coder's judgement. For example, a coder may recognise a video game such as Minecraft as educational, while another coder may recognise the video as recreational. To overcome this, coders only coded the content if they were certain of the content and context in which it was used.

Table 3.3

Content Classification and Corresponding Definitions

Content classification	Definition
Educational	Screen media created with the purpose to educate, inform, and enlighten the viewer (Kirkorian & Anderson, 2008).
Recreational	Screen media created with the purpose to entertain the viewer, with no intention to educate, inform or enlighten the viewer (Gemming, Doherty, et al., 2015)
Social	Screen media created with the purpose to communicate with others (Rideout, 2015).

Context

The context dimension of the protocol codes the environment that the participant is in when the screen exposure occurs. The content dimension is made up of five facets: device attention, location, social environment, social interactions, and associated behaviours.

Device Attention. 'Device attention' refers to the level of attention the participant appears to be giving to the screen-based media device. I alluded to the purpose of this facet earlier on in this chapter when I presented the definition of screen exposure. I created the device attention facet to distinguish engagement levels during media multitasking. Initially, I had this section made up of three categories: 'Primary', 'Secondary' and 'Background'. I coded the device as the a 'Primary' device when the screen-based media device appeared to capture a large amount of the participant's attention. Only one screen-based media device per image could be coded as a primary device. I coded the device as the a 'Secondary' device when there was more than one screen-based media device in the image, and the screen-based media device appeared not to capture the participant's full attention. For example, a television was turned 'on' in the background while the participant used their mobile phone. Multiple screen-based media devices could be coded as a secondary device in the same image. I coded a device as a 'Background' device when the screen-based media device was in the image, but the participant appeared not actively engaged. For example, when a participant walks into a room with a screen-based media device being used by another person. During testing, I observed that in most cases, the 'Background' device was not being used by the participant and therefore did not provide meaningful data. I decided to remove this category from the framework. Table 3.4 presents the device attention categories and definitions.

Table 3.4

Device Attention and Corresponding Definitions

Device attention	Definition
Primary	When a screen-based media device appears to capture a large amount of the participant's attention. Only one screen-based media device per image can be coded as primary.
Secondary	When there are more than one screen-based media devices visible in the image and the screen-based media device appears to not capture full attention of the participant. Multiple screen-based media devices can be coded as secondary in the same image.

Location. 'Location' refers to the place or type of surroundings where the screen use occurred (Watkins et al., 2018). I created the location categories based on Signal, Smith, et al. (2017) and Watkin et al. (2018) previously developed coding protocols for automated wearable cameras. Signal, Smith, et al.'s (2017) and Watkin et al.'s (2018) coding protocols were developed as part of the Kids'Cam study. The Kids'Cam study was a cross-sectional study of 168 children measuring children's exposure to food marketing. As Signal, Smith, et al. (2017) and Watkins et al. (2018) measured children's exposure to food marketing, they included a large variety of codes that identified the participant's location. For example, these codes included transportation, school, community venues and home. I adapted these codes based on the few studies that measured the location of children's screen use to ensure that all relevant codes for the context of screen use were identified. For example, Signal, Smith, et al. (2017) and Watkins et al. (2018) use the code 'Home' to code the participant's location. However, the literature suggests that children's location within the home during screen use may influence the duration and content consumed (Atkin, Corder, et al., 2013). For this reason, I further disaggregated the 'Home' code to include specific places in the home, including the bedroom, living room, kitchen/dining room, and outside. Table 3.5 presents the location categories and definitions.

Table 3.5

Location Categories and Corresponding Definitions

Category	Subcategory	Definition
Home	Bedroom	A room or space used for sleeping in (Cambridge University Press, n.d.-a). Indicated by the presence of a piece of furniture for sleep or rest, typically a frame with a mattress.
	Living room	A room or space that is used for relaxing in and entertaining guests (Cambridge University Press, n.d.-f). Indicated by furniture such as lounge or coffee table.
	Kitchen / Dining Room	A room or space where food is kept, prepared, cooked, and eaten (Cambridge University Press, n.d.-e). Indicated by the presence of cooking appliances, such as a stove, refrigerator, microwave. Indicated by the presence of a flat surface such as a table, on which meals are served on.

Category	Subcategory	Definition
	Outside	A space outdoors within the home boundaries (e.g., backyard and front yard; Watkins et al., 2018). Indicated by grass or pavement within the home boundaries.
	Other	A room or space in the home (e.g., an office).
Public	Street	The public areas or roads of a town, suburb, or city. Includes roads, footpaths and outside of private properties, community venues and retails (Watkins et al., 2018).
	Community Venue	A building or room where members of the community can meet or use. Includes public libraries, recreation centres/community halls or churches (Watkins et al., 2018).
	Retail	A place/space where goods are sold to the public. Includes general product retailers (e.g., stores, supermarkets, service stations and shopping malls)
	Food Retail	A place where meals are prepared and sold (e.g., restaurants, cafes, and bakeries).
	Recreational Space	A space located outside rather than inside a building. A space where individuals participate or watch organised sports. Includes parks, beaches, rivers, walking tracks, sports stadiums, and sports grounds (Watkins et al., 2018).
Transport	Private Transport	Inside a truck, van, or car (Watkins et al., 2018).
	Public Transport	Inside a train, bus, ferry, or aeroplane (Watkins et al., 2018).

In most cases, it was relatively straightforward to identify the participant's location in the images. The 'Home' category codes were identified as the place where one lives (Cambridge University Press, n.d.-c). Typically, the 'Home' category codes were characterised by spaces within the home boundaries and included all private residences (i.e., friends and extended family residences; Watkins et al., 2018). The kitchen and dining room were categorised as a single code. Initially, I had these categorised separately; however, during testing I observed that it was difficult to distinguish between the two locations as the dining area was often in the same room as the kitchen. The 'Public places' codes were identified as indoor and outdoor spaces shared with other community members (Watkins et al., 2018). The 'Public' code categories included community venue, retail, food retail and recreational space. The 'Transport' codes were identified as the use of vehicles for getting from one place to another (Watkins et al., 2018). The 'Transport' code categories included private transport, such as inside a privately owned car, van, or truck, and public transport such, as a bus, train, aeroplane or ferry (Watkins et al., 2018).

Social Setting. 'Social setting' refers to the social context of the screen exposure. I created the social setting categories based on Gemming et al.'s (2015) previously developed coding protocol for automated wearable cameras. Gemming et al.'s (2015) coding protocol was developed as part of a study that used automated wearable cameras to assess eating episodes' social and environmental context. I adapted Gemming et al.'s (2015) codes based on the literature investigating the impact of co-viewing and co-playing. For example, Gemming et al.'s (2015) used the code 'Social interaction' to code the interaction between two individuals during an eating session. However, the literature suggests that parent-child co-viewing is associated with children watching more television and greater exposure to adult television programmes (Latomme et al., 2018). For this reason, I further disaggregated the code 'Social interaction' to include co-viewing and co-participating. I then created another facet of the coding protocol to code who the child was with during this interaction. This facet was called the ‘Social environment’. Tables 3.6 and 3.7 present the social interaction and social environment categories and definitions, respectively.

Table 3.6

Social Interaction Categories and Corresponding Definitions

Social interaction	Definition
None	No one else is present in the image.
Co-participating	When the child and another person are actively using a screen-based media device together (e.g., a parent playing a video game with their child).
Co-viewing	When the child and another person are watching a screen-based media device together. Indicated by other people’s body positioning facing the device.

Table 3.7

Social Environment Categories and Corresponding Definitions

Social environment	Definition
Alone	No one else is visible in the image.
Single adult	One person who appears to be over 18 years of age is visible in the image.
Single child	One person who appears to be under 18 years of age.
Adults only	People who appear to be over 18 years of age are visible in the image only (must be multiple people).
Children only	One or more people who appear to be under 18 years of age.

Social environment	Definition
Mixed ages	Multiple people in the image who are children and adults.

During the testing, I encountered some challenges when identifying the social environment and social interaction in the image. Social environment refers to who the child was with during the screen exposure. Initially, this facet was made up of four categories: 'Alone', 'Adult', 'Child', and 'Adult/Child unclassifiable'. The 'Child' code referred to one person who appeared to be under 18 years of age. The 'Adult' code referred to one person who appeared to be over 18 years of age. The coder was required to code each individual in the screen exposure separately. For example, if four people were present in the image, coders were required to code the four people individually; if there was only one person in the image, the coder only had to code one person. During testing, I found that the reliability of the coders decreased as the number of people in the image increased. The decrease in the coder's reliability may be due to the increased variance in the image. For example, it was easier to code the location of an image as it does not change until the participant moves, while for social environment, the coding depended on the actions of the people surrounding the participant. For this reason, I broadened the social environment categories to 'Alone', 'Single adult', 'Adults only', 'Children only', and 'Mixed ages'. In doing this, coders were not required to track the actions of each individual in the image.

Further, during the development of the social environment categories, I limited the categories to only include people, and did not include pets or animals. This decision was informed by the limited research on the impact of pet interaction during children's screen use (Charmaraman, 2022). Additionally, the inclusion of pets or animals may increase the difficulty in image coding as the camera may not adequately capture animals and their movements due to the camera position and field of view. However, researchers can adapt the protocol to include pets or animals to suit their research needs.

Social interaction refers to the exchanging information, interactivity, or lack of exchange between people. Initially, this facet was made up of four categories: ‘None’, ‘Co-participating’, ‘Co-viewing’, and ‘Background’. The ‘Background’ category was used to code when another person was present in the image but was not co-participating or co-viewing. For example, when a person was in the background of the image or not looking at the screen-based media device. During the testing, I observed that the ‘Background’ category was not practically relevant to screen use research and decided to remove this category from the coding protocol. I removed this category to ensure the coding protocol was concise and only captured meaningful data. Thus, the refined coding protocol included ‘None’, ‘Co-participating’, and ‘Co-viewing’. Coders coded the social interaction in the image based on indicators such as another person being visible in the image and that person’s body language or movements (Gemming, Doherty, et al., 2015). For example, I coded an image as ‘Co-viewing’ if the image depicted another person looking at a television during the child’s television screen exposure. I coded an image as ‘Co-participating’ if the image depicted another person holding a gaming controller and using the same screen-based media device as the child.

Associated Behaviours. Associated behaviours refer to co-occurring actions or tasks the participant undertakes during screen exposure. I created the associated behaviour categories based on Kerr et al.’s (2013) previously developed coding protocol for automated wearable cameras. Kerr et al.’s (2013) coding protocol was developed in a study that used automated wearable cameras to assess a range of sedentary behaviours in adults. For example, in Kerr et al.’s (2013) coding protocol, they coded for multiple behaviours simultaneously occurring in an image, such as eating while watching television. I used Kerr et al.'s (2013) definitions to identify multiple behaviours occurring in an image. I then searched the literature to identify other behaviours associated to children's screen use. This facet was made up of two categories: 'Eating' and 'Multitasking'. Initially, I had both categories further split into subcategories. For example, the 'Eating' category was further split into the subcategories 'Snack', 'Meal', and 'Beverage'. The 'Multitasking' category was split into the subcategories 'Writing', 'Reading', and 'Hobby'. However, during testing I observed that the additional subcategories were not practically relevant to screen use research and were time-intensive to code. For these reasons, I decided to remove these subcategories from the coding protocol. Thus, the refined coding protocol included two categories: 'Eating' and 'Multitasking'.

It is important to note that while I did not include the additional subcategories, the coding protocol's hierarchical structure allows other researchers to adapt it to suit their research needs. For example, suppose a researcher wants to compare children's snacking behaviours and meal behaviours during screen use, they could adapt the coding protocol to incorporate the subcategories 'Meal' and 'Snacks' under the broader category of 'Eating'. Table 3.8 presents the associated behaviours categories and definitions.

Table 3.8

Associated Behaviours and Corresponding Definitions

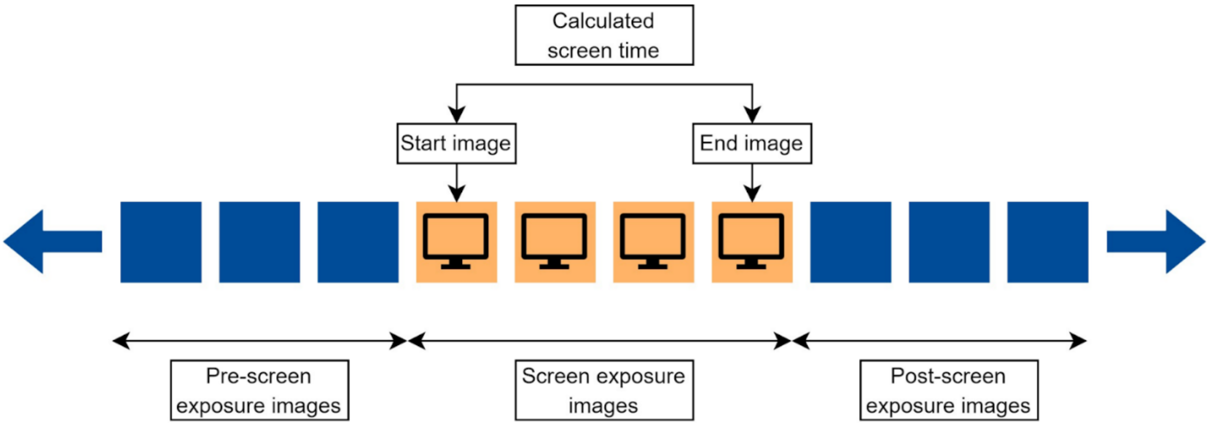
Behaviour	Definition
Eating	The presence or observed consumption of food or drink in an image (Kerr et al., 2013).
Multitasking	When a participant engages in another task or multiple tasks during screen exposure (Chinchanachokchai et al., 2015).

Screen Time Duration

Screen time refers to the duration of the screen exposure. Screen exposure was measured from the first image containing an active screen-based media device to the last image containing an active screen-based media device (see Figure 3.8). Screen time duration was then measured by multiplying the number of images coded with a behaviour by the camera epoch. For example, if a screen exposure consisted of 30 images, with an image interval of 2 seconds, the calculated screen time duration would be 1 minute.

Figure 3.8

Example of how Screen Time was Measured using the Coding Protocol



Stage 3: Pilot Coding Protocol

The third stage involved testing the coding protocol. To test the coding protocol, I coded images that were included in my Honours' thesis study and captured by research team members who wore the cameras during camera testing. I repeated stages 2 and 3 in an iterative cycle until no more changes were made to the coding protocol and coding manual. During testing I encountered several challenges when applying the protocol. To overcome these challenges, I developed rules for coding the images. The coding rules were based on previously developed rules in other automated wearable camera studies. In this section I have outlined the process of developing and refining the coding rules during testing.

Blocked Image Rule

The most important coding rule I developed was the blocked image rule. During testing, I observed that within some screen exposures, some images were blocked or did not capture the screen-based media device being used. For example, in some cases, a blanket blocked the camera, or the camera was pointed towards the ceiling due to the participant lying down. I applied the blocked image rule when a screen-based media device was not present in the image, but the coder had a high degree of certainty that the participant was engaging in screen use. The blocked image rule allowed coders to code an image that did not contain an active screen-based media device as screen exposure if the image before or after the image showed a screen-based media device and showed the same location (e.g., the participant did not move to another room). An image was coded as 'Uncodable' if the location of the image changed, the screen-based media device changed, or the coder was not confident that the participant remained in the screen exposure. An example of the blocked image rule is depicted in Figure 3.9, where the images 1 to 5 are blocked but coded as screen exposure due to the images before and after the blocked images depicting screens.

Figure 3.9

Example of Blocked Image Rule

The blocked image rule was developed based on previous wearable camera research. The coding rule was first developed by Lowe (2017) when using wearable cameras to assess the nature and extent of New Zealand children's screen time. According to Lowe's (2017) coding rule, images that did not contain a screen-based media device were coded as a screen exposure if the image before and after the image show a screen-based media device, and there were no more than eighteen images (approximately 4.5 minutes) between them. I adapted the coding rule to include the physical setting rule to avoid overestimating the duration of the screen exposure. For example, if a blocked image had an image of the participant watching the television before and after, the blocked image would be coded for as screen exposure. However, if the blocked image had an image of the participant watching the television before and then was followed by an image of another room and then the image of the television, the blocked image would not be coded as screen exposure as this indicated the participant had briefly left the room and was not exposed to the screen-based media device during that time.

Proportion of Device Visible in the Image

During testing, I observed that sometimes only a small proportion of the screen-based media device was present in the image. For example, some images only depicted the corner of the screen-based media device. I coded images only showing a small proportion of the device as screen exposure. I coded these images as screen exposure as the participant was likely to be engaging in screen use. Coders were asked to use context from other images surrounding these images to gain certainty in their coding decisions. If a coder found reasonable evidence of the presence an active screen-based media device, then the image was coded according to the coding protocol.

Context Rule

During testing, I observed that I could code an image more accurately after viewing the screen exposure episode before I began coding. Viewing the screen exposure episode before I began to code an image gave me a better understanding of what was happening during the screen exposure. For example, when I viewed the screen exposure episode before I began coding, I observed more characteristics of the participant’s location, which allowed me to code the images more accurately. The context rule allowed coders to code an image based on the context of the screen exposure episode if the coder was more than 50% certain that the context of the before or after images is the same context of the image being coded.

Uncodable Images

During testing, I observed that I could not confidently code some images due to poor image quality. For example, when the participant moves, the images can become blurry. Consequently, making it hard to observe what is happening in the image. I developed the ‘Uncodable’ code to ensure coders coded the image as accurately as possible when they could not apply the context or eighteen image rules. Common reasons for an image being coded as ‘Uncodable’ included blurry or poor lighting that made it challenging to observe the context of the image. Images were only coded as ‘Uncodable’ if all aspects of the image or set of images could not be confidently determined (i.e., the was not uncodable if the background of the image was dark but you can still code the device).

Unclassifiable Images

I developed the unclassifiable code after I observed that I could not confidently code specific categories accurately but was able to identify broader categories. For example, I coded the content of screen-based media devices as ‘unclassifiable’ when the content displayed on the device was blurry and I could not apply the context or blocked image rule. Moreover, the ‘unclassifiable’ code was also applied to an image that could not be coded at the highest level. For example, if the coder could not determine if the device was Portable > Tablet or Portable > Mobile device, then the coder coded the image as Portable > Unclassifiable.

Stage 4: Implement Coding Protocol

The fourth stage involved implementing the coding protocol. In this stage, I generated the image data systematically and ethically. This stage involved (1) downloading the images, (2) running them through a face blurring software, (3) storing the images in a secure location,

(4) uploading and viewing the images in specialised software, (5) completing coder training, and (6) coding the images.

In Stage 4, I refer to some ethical considerations I have not discussed yet. I have discussed these ethical considerations in the next section. It is also important to note that Stage 4 occurs after a participant has worn the automated wearable camera. I have not addressed the procedures for wearing the automated wearable camera as they depend on the study aims (e.g., how long the participant wears the camera). The specific study protocols for wearing camera are outlined in Chapters 4 and 5.

Downloading, Blurring and Storing the Images

This section described the process of downloading, blurring, and storing images. Steps 1-5 occurred at the participants' homes, and step 6 occurred at the research lab.

1. The encrypted SD card is removed from the camera and inserted into the laptop. The laptop is only accessible to the members on the research team in accordance with ethical guidelines (Kelly et al., 2013).
2. The encrypted SD card is unlocked using specialised software, and the folder with the timelapse videos captured by the camera is moved onto the laptop desktop. This step may take several minutes to move all the data as it depends on the amount of data captured by the automated wearable camera. For example, a participant who wore the camera for one day will have a shorter offload duration compared to a participant who wore the camera for four days.
3. Once downloaded, the timelapse videos were run through a face blurring software that was developed and tested for the KidVision study (Chapter 5; Sanders, 2023). The face blurring software converts the timelapse videos to images and uses a machine learning model to detect faces in these images. A gaussian blur is applied to regions identified as faces before these blurred images are combined into a timelapse video

for the participants to review (see step 4). This is done in accordance with ethical guidelines to protect third party privacy (Kelly et al., 2013). This step may take up to 40 minutes, depending on the amount of data captured.

4. Participants were given the opportunity to review the blurred images in a timelapse video and delete any images they did not want the researchers to view. The cameras capture a large amount of data (i.e., > 2000 images) and going through each image individually during data collection was not feasible for the participant or researchers. Therefore, the images were made into a timelapse video to streamline the image reviewing process.
5. Once the images were reviewed by the participant, the approved images were saved onto the laptop desktop and the original time lapse video files and images were deleted.
6. As soon as the researcher arrived at the research lab, the saved images were uploaded to the Australian Catholic University (ACU) secure server. This server is only accessible to the research team members in accordance with ethical guidelines (Kelly et al., 2013). Images were stored on the server in a folder labelled with the participant's number.
7. Once the images were stored on the ACU secure server, they were then coded.

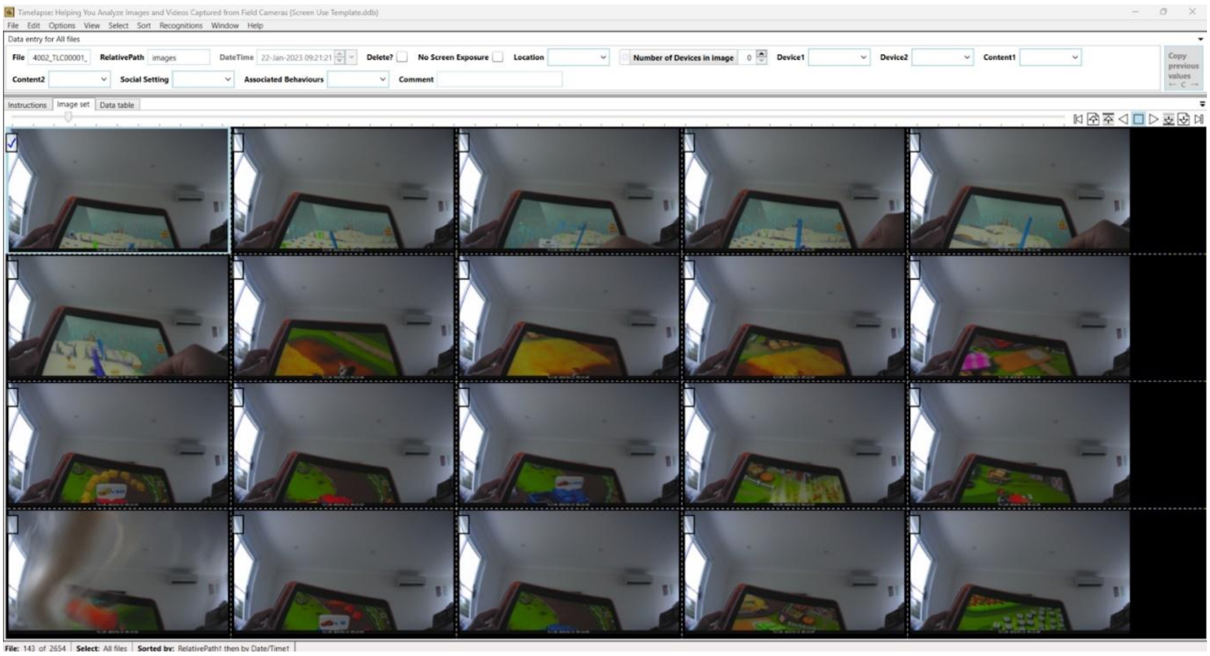
Preparing the Software

I coded the images using the Timelapse 2.0 Image Analyser software developed by Greenberg (2023). I followed existing protocols created by Greenberg (2023) on how to use the Timelapse 2.0 Image Analyser software. This section described the process of preparing the images using the Timelapse 2.0 Image Analyser software.

1. The coder copied and saved the Screen Use Template .tdb file (i.e., a template file I created that defines the analysis codes for the software) into the participant's image folder on the ACU server.
2. Coders then open the Timelapse 2.0 software and load the Screen Use Template .tdb in the appropriate folder by following the prompts provided by the software.
3. Coders then made the software window full screen to code the images. While testing the coding software I established that images were coded more accurately when the images were displayed as large as possible. See Figure 3.10 for an example of the interface.

Figure 3.10

Example of Timelapse 2.0 Interface using the Screen Use Template File



Coding the Images

Coder Training. Two coders completed a coder training session prior to coding the images. Coders included a researcher who had no experience coding images and me. The coder training session provided information on the ethical guidelines for handling image-based data, and familiarisation with the coding manual, software and popular screen media used by the participant's age group. Coder's then practised image coding on a set of images (n=2,654) according to the coding protocol. During this practice session coders can ask questions about image coding (e.g., coding rules or advice on how to code an image). After the set of images were coded, I calculated the image-by-image inter-rater reliability with model answers using percentage agreement. The percentage agreement statistic is a common method used to assess inter-rater reliability as it is easily calculated and interpretable (McHugh, 2012). The percentage agreement statistic does not consider the possibility that raters guessed scores and may overestimate the true agreement among raters (McHugh, 2012). Thus, I used percentage agreement for practical implications during the training session, and not an assessment of image coding quality check. Coders needed to reach a 90% agreement with model answers on the subset of images before they began image coding. Coders repeated this process until they achieved a score of ≥ 90% agreement (McHugh, 2012). Once, coders completed the training session, they began image coding.

Image Coding. This section described the process of coding the images after the coding training session was complete. The coding protocol can be found in Appendix D.

1. After the images were uploaded into Timelapse 2.0, coders coded each image in chronological order. To gain an understanding of what was happening in the image, coders were allowed to browse through the screen exposure episode before coding the image.
2. Coders then determined if the image was codable or contained a screen exposure. If the image was codable, they continued coding the image. If the image could not be coded or did not include a screen exposure, they then coded the image as 'Uncodable'.

3. Coders then coded the image for the location, number of devices, device type, content, social setting, and associated behaviours, respectively.
4. Timelapse 2.0 software creates a dsql database file in the same folder and automatically saved the recorded data into the database. Once coding was complete, the coders exported the data to a .csv file and saved the file on the ACU server.

Ethical Considerations for Data Storage and Management

In Chapters 1 and 2 of this thesis I highlighted that privacy and ethical concerns are an important issue in automated wearable camera research. Automated wearable cameras take first-person point-of-view images that capture rich contextual data of participant's behaviours. Consequently, these cameras capture detailed information about participant and the people around them. In the literature, privacy and ethical concerns about the use of camera in research are often debated (Mok et al., 2015). However, in Chapter 2 I found limited privacy or ethical concerns from participants. Individuals who did express concern were primarily non-participants (i.e., third parties who may have been captured in an image). Given the importance of participant and third-party privacy, especially involving children, I have taken great consideration into how to address the ethical concerns on data storage and management of image-based data. Thus I developed the data storage and management procedures based on the ethical framework for automated wearable cameras in health behaviour research (Kelly et al., 2013).

Access to Data

Following the recommendations from the ethical framework (Kelly et al., 2013), I stored all image-based data on a secure project-specific server on the ACU network drive to ensure that only members of the research team had access to the data. Further, I used encrypted SD cards during data collection to ensure only research team members had access to the image-based data. In this way, lost devices are not a risk to participants and third-party privacy and confidentiality. Additionally, this meant images captured by the cameras were not accessible to the participants, preventing participants from saving or distributing images depicting third parties in an image.

Inappropriate or Unwanted Images

Automated wearable camera image capture is passive (i.e., the camera takes an image independent of the wearer's actions), thus, it is possible inappropriate or unwanted images of participants or third parties may be captured (Kelly et al., 2013). Participants are provided information on how to turn the camera 'off' and to remove the camera for certain circumstances. For example, when getting changed or going to the bathroom. However, it is common for participants to forget they are wearing the camera, and consequently may capture inappropriate images (Kelly et al., 2013). To overcome this, participants are given the opportunity to review and delete any images they did not want the researchers to view (Kelly et al., 2013).

Third Parties

Third parties who encounter the participant while wearing the camera may have their image taken, knowingly or unknowingly, without the opportunity to provide informed consent (Kelly et al., 2013). To protect third-party privacy, the images are run through a face blurring software developed and tested for the KidVision study (Chapter 5; Sanders, 2023). The face blurring software converts the timelapse videos into images and uses a machine learning model to detect faces in these images. In this way, third parties captured by the automated wearable camera are unidentifiable to the participant and research team.

Coder Training

Before accessing any images, all coders completed training that provided information on the ethical guidelines for handling image-based data. Coders had to read the Australian Catholic University Research Code of Conduct (Australian Catholic University, n.d.), the ethical framework for automated wearable camera research (Kelly et al., 2013), the ethical guidelines in the coding manual, and sign an ethical declaration form. Additionally, all coders were required to have a Working with Children's Check when handling any data involving child participants.

Challenges

I have presented this process sequentially in this Chapter; however, in practice, the development and refinement of the coding protocol was an iterative process. The results from each stage informed the next stage and were repeated in an iterative cycle until no more changes were made to the protocol. Throughout this process I encountered many challenges that informed decisions on the coding protocol. In this section I have discussed the challenges I encountered while developing the coding protocol.

Subjective Input

The process of coding automated wearable camera data to measure screen-based behaviours is based on decisions made by the coder (Hänggi et al., 2020). Consequently, the accuracy of data is dependent on the coder's pre-existing knowledge of the devices, content, location, and other contextual features. For example, for a coder to identify a television they must know what a television looks like and how it is typically used. If a coder does not have this pre-existing knowledge, then they will not be able to identify the device being used as a television.

This issue was most apparent when I tested the content dimension of the coding protocol. During testing, I observed that coders required specific pre-existing knowledge to identify the type and classification of content that was not required in the other dimensions. I believe this was due to the coders having better pre-existing knowledge of the other dimensions and categories. For example, a coder was more likely to be able to recognise a television due to their interactions with these devices in their daily lives than recognise the specific television programme the participant was watching. To minimise this issue, I removed detailed subcategories (i.e., child- or adult-orientated) and instead had a broad subcategory (i.e., programme). I also included the code, ‘unclassifiable’, so when coders could not confidently code specific categories, they were able to identify broader categories.

Further, to address this issue, coder training, included familiarisation of popular screen-based media devices and screen media used by the participant’s age group. Coder training included showing coders images of popular screen-based media devices and screen media, and examples of what the screen-based media devices and screen media looked like in the automated wearable camera data. While the coder training included examples of these screen-based media devices and screen media, it is difficult to have coders remember every detail. Therefore, using coders who have prior knowledge of a range of different current technologies may be advantageous to researchers using automated wearable cameras to measure children’s screen use.

Trade-off Between Depth Coding and Coding Time

The process of image coding is time and resource intensive. Therefore, researchers need to be able to balance the quality of the data with the amount of labour required to code each image. For example, coders using a detailed coding protocol (i.e., having to code lots of individual behaviours) will be more time intensive compared to a coding protocol that has broader codes (i.e., grouping certain behaviours). This was demonstrated in Chapter 2, where I found that simple coding protocols averaged around 30 minutes per participant (Kelly et al., 2012), while more complex coding protocols ranged from 40 minutes to 9 hours (Beltran et al., 2016; Zhou et al., 2019).

I encountered this challenge throughout the initial tests of my coding protocol. In the original version of the coding protocol, I had coders code each image with more detail than the coding protocol presented in this thesis. For example, coders were required to code whether the food a participant was consuming was a meal, snack, or beverage. However, I observed coding to this level of detail was too time intensive for larger datasets. Coders were only able to code 7.7 images per minute and took 11.7 hours to code 4,497 images.

To address this issue, I collapsed specific subcategories into broader categories to minimise the amount of labour required to code each image. For example, the subcategory codes 'meal', 'snack' and 'beverage' were collapsed into a broader category called 'eating'. I then structured the coding protocol to allow other researchers to easily adapt the coding protocol to suit their research needs. For example, if a researcher wanted to code specific dietary behaviours during screen use, they can further modify the coding protocol to include subcategories under the 'eating' category. In this way, researchers using this coding protocol can code to the level of detail they require.

Further, I used an object classification model to reduce the amount of data needing to be coded. Briefly, object classification is a computer vision technique used to determine if an object of interest is present in each image (Chen et al., 2015). The model is trained to identify important objects by being provided with pre-coded data with the object already identified. These studies use machine learning techniques to undertake automated image recognition (Biswas et al., 2017; Krizhevsky, Sutskever, et al., 2015). The data collected in Chapter 4, was used to train an object detection model, based on the You-Only-Look Once model architecture (Bochkovskiy et al., 2020), to identify images containing electronic screens (with power on). This model was used to reduce the amount of image coding in Chapter 5.

Briefly, another way to limit the amount of data captured would be to reduce the automated wearable camera epochs. For instance, a camera set at an epoch of 2-seconds will capture more data than a camera set at an epoch of 30-seconds. However, reducing camera epochs may impact on the accuracy of the measurement. For example, I observed during testing that mobile phone use was often more sparingly and in shorter duration. Therefore, mobile phone use that was under 30-seconds (i.e., sending a message) may be missed. In Chapter 2, I found little evidence on the best camera epoch to capture screen use behaviours. I address this gap in the literature in Chapter 4.

Conclusion

In this Chapter, I have demonstrated how I developed and tested the coding protocol and procedures to measure children and adolescents (ages 5-18 years old) screen use, highlighted ethical considerations for handling image-based data, and discussed the challenges of developing the coding protocol. To develop the coding protocol, I followed a guideline for developing and modifying behaviour coding protocols in paediatric populations to decrease the chance of bias occurring through the development phase (Chorney et al., 2015). The first stage involved refining the research questions to clearly define the constructs of screen use (i.e., device type, screen time, content, social environment, and associated behaviours). This stage established a foundation for the coding protocol, ensuring that it aligned with key factors identified from the literature that influence children's health outcomes.

The subsequent stages involved the developing and refining the coding protocol in an iterative cycle until not more changes were made to the coding protocol and coding manual. During these stages I encountered several challenges, such as blocked images and the need for specific coder knowledge, when applying the protocol and developed coding rules to overcome these challenges. For example, a major challenge I faced was coding blocked images (i.e., camera lens was blocked by a blanket) when it was apparent that screen use was occurring, but a screen-based media device was not present in the image. I developed the blocked image rule to allow coders to code images that were blocked as screen exposure to avoid underestimating screen exposure. Refinements such as, the blocked image rule, were instrumental in improving the protocols' reliability and applicability.

The final stage involved implementing the coding protocol, which included addressing ethical considerations and coder training. Procedures for secure data storage, managing inappropriate images, and protecting third-party privacy were implemented based on established ethical frameworks (Kelly et al., 2013). Coder training was implemented to ensure all team members adhered to ethical guidelines, were familiar with the coding manual, and coded images in agreement with a set of model images. Additionally, in this stage I implemented techniques to balancing coding depth with time efficiency. While the refined protocol demonstrated improvements in capturing detailed information about screen use, the process of coding the images remined labour-intensive. The balance between data quality and coding efficiency was an important consideration, which lead to the use of an object classification model to streamline image coding.

Chapter 4: Evaluating the Validity of Automated Wearable Cameras to Assess Children's Screen Use

Preface

I originally planned to recruit 30 participants in this study. The study would have examined the convergent validity of all categories in the coding protocol. Briefly, the rationale was that this measurement instrument has not been validated in children. Therefore, I was going to conduct this study to investigate the convergent validity of the coding protocol. However, when restrictions related to COVID-19 pandemic banned all face-to-face data collection, I was not able to begin data collection when I had originally planned to. When the face-to-face data collection ban was lifted, I then experienced slower than anticipated recruitment. To ensure I could complete this study before my thesis submission deadline and still conduct an adequately powered study, I reduced the number of categories assessed in the coding protocol and sampled 10 participants.

Introduction

Screen-based media devices, such as televisions, tablets, and smartphones, are now ubiquitous in childhood and adolescence (Stiglic & Viner, 2019). These devices have transformed how children and adolescents learn, socialise, spend their leisure time, and interact with the world (The Lancet Child Adolescent Health, 2018). There is now a concern that these devices may be detrimental to children and adolescent's health and well-being outcomes. However, evidence to support these concerns is lacking, with systematic reviews showing inconsistent findings (Carson et al., 2016; Costigan et al., 2013; Stiglic & Viner, 2019; van Ekris et al., 2016).

One reason for the inconsistent findings may be the quality of the instruments used to measure children's screen use. Much of the research on children's screen use relies on unvalidated self- or parent-reported measurements (Cain & Gradisar, 2010; Hale & Guan, 2015), which lack precision (Atkin et al., 2013; Saunders et al., 2011), are prone to bias (Ainsworth et al., 2012), and often do not account for newer types of screens (Jordan et al., 2007). Most of these measurement instruments aggregate screen time and often do not measure the context of the screen exposure or the type of content, which are factors that may affect health and well-being outcomes (Greitemeyer & Mügge, 2014; Latomme et al., 2018; Twenge et al., 2019). When compared to the limited objective evidence that does exist, results indicate that self- or proxy- reported data has low validity (Robinson et al., 2006). While objective measurements of screen time—such as electronic television monitors and video observation—provide valid data, these measurements are now outdated due to the increased diversity of screen types (e.g., smartphones, tablets, laptops) and do not measure the content or context of screen use.

The "gold standard" in screen use measurement is direct observation (Perez et al., 2023). Direct observation measurements typically involve video recording or observing a participant's screen use in their home (Anderson et al., 1985). Results from direct observation provide the most comprehensive, reliable and valid measurement of screen use to date (Perez et al., 2023). However, direct observation is time-consuming, expensive, or impractical (Hardy et al., 2013).

Automated wearable cameras may offer a solution that provides an accurate measurement and is more practical than the current screen use methodology. In Chapter 2, I sought to examine the evidence on the validity and reliability of automated wearable cameras in child and adolescent populations. I found that automated wearable cameras are a feasible and reliable method for measuring the type of device and location of screen use behaviours. However, I found limited research examining the validity of automated wearable cameras for measuring children's health behaviours. This is consistent with a recent scoping review, which did not identify any studies investigating the validity of automated wearable cameras to assist with the self-management of chronic disease (Maddison et al., 2019).

While a measurement instrument can have high reliability, it does not equate to the measurement instrument being valid (Kimberlin & Winterstein, 2008). Therefore, it is possible for automated wearable camera measurements to be very reliable but have poor validity. Given the importance of improving the quality of research on children's screen use, it is important to consider the validity of automated wearable camera measurements.

There are many types automated wearable cameras, and each type of device are unique in physical characteristics (e.g., size or weight) and features (e.g., camera epochs or image quality). Consequently, no assumptions regarding validity can be made from one device to another. Therefore, information regarding validity must be matched to each individual device, coding protocol and population under study (de Vet et al., 2011).

Previous studies assessing screen-based behaviour (i.e., the type of device, duration and location of screen use) using automated wearable cameras have used independent coding protocols with differing automated wearable cameras and camera epoch lengths (Hänggi et al., 2020; Smith et al., 2019). To date, no coding protocols exist to extract reliable and valid measurements of the type of device, content and context of children's screen use captured in automated wearable camera research. The primary purpose of this study was to evaluate the convergent validity of automated wearable camera measurements using the coding protocol in Chapter 3 with data captured by direct observation.

Further, there is limited information on the most accurate camera epoch to capture screen use behaviours. In Chapter 2, I found that camera epochs ranged from 4- to 30-seconds, with most studies using 7-second epochs. Screen use behaviours such as mobile phone use are often used more sparingly and in shorter duration (i.e., using a mobile phone to send a message) than other behaviours. Therefore, shorter screen use behaviours may be missed with longer camera epoch lengths. The secondary purpose was to examine the impact of the camera epoch on estimates of screen use when using an automated wearable camera.

Research Questions and Hypotheses

Research Question 1

What is the convergent validity of automated wearable camera measurements for assessing screen-based behaviour in children aged 8-11 years old in a home setting compared to direct observation?

Hypothesis 1

Based on previous literature, I hypothesised that there would be substantial to almost-perfect agreement ($k \geq .60$, ICC = $\geq .5$) and strong correlations ($r = \geq .7$) between automated wearable camera measurements and direct observation measurements of screen time, type of device, and location. I further hypothesised that there would be substantial agreement ($k = \geq .60$, ICC = $\geq .5$) and moderate correlations ($r = \geq .4$) between automated wearable camera measurements and direct observation measurements of content, social environment, and associated behaviours. These behaviours require more subjective input from image coders and, therefore, may be more likely to be incorrectly coded compared to other screen use behaviours.

Research Question 2

How does camera epoch length affect estimates of screen use when using an automated wearable camera?

Hypothesis 2

I hypothesised that the correlations and levels of agreement between direct observation and automated wearable camera measurements would decrease as the epoch length increased. That is, longer epoch lengths would result in lower accuracy than shorter epoch lengths.

Methods

This study was guided by the COSMIN taxonomy measurement properties and definitions and the study design checklist and reported according to the COSMIN reporting guidelines for studies on measurement properties (Gagnier et al., 2021; Mokkink et al., 2018)

Participants and Recruitment

Participants in this study were children aged 8 to 11 years old from the community in New South Wales Central Coast and Sydney (Australia). Though screen time recommendations state that all young people and children aged 5 to 17 years have similar screen time limits, to reduce variability due to age (e.g., camera position due to body size), I limited the eligible age range to 8-11 years old.

Children were eligible to participate if they: (1) were aged between 8-11 years old; (2) located in New South Wales Central Coast and Sydney (Australia); (3) had personal access to at least one screen-based media device, such as a television, smartphone, tablet, or laptop, at home; and (4) a parent/caregiver was available for the direct observation session in the participant's home. Children were excluded if they did not have access to any screen-based media devices at home.

Participants were recruited by convenience sampling using online social media advertising and word of mouth. Upon completion of the study, participants received a $25 Woolworths Essential Gift Card as a token of appreciation and a summary report of the child's movement behaviours, including sleep duration, physical activity, and sedentary behaviour. I used the gift card and summary report to aid participant recruitment.

Ethical Approval

Written consent was obtained from all the participants before the commencement of the study (see Appendix E for the participant information letter, child information letter, consent form and assent form). The procedures for this study were guided by international ethical guidelines for automated wearable cameras (Kelly et al., 2013) and were approved by the Australian Catholic University Ethics Committee (Approval #2017-317H). See Appendix F for ethics approval.

Measures and Procedures

I recruited participants and collected data between October 2022 and July 2023. Following recruitment, I completed two home visits with each participant. During each home visit, participants were asked complete an observation session where they were asked to wear an automated wearable camera on a chest harness and perform screen-based and non-screen-based behaviours while being observed by a researcher, which served as the comparison measure.

At the first home visit, I collected basic demographic information, measured height and weight, completed the two-hour semi-structured observation session, and a content recall interview to confirm the content viewed during the observation session. At the second home visit (approximately one week following the first visit), I completed the second two-hour semi-structured observation session and the content recall interview.

Demographics

I asked the participating child's parent to complete a Family Information Questionnaire (see Appendix G for the Family Information Questionnaire). The Family Information Questionnaire included the participant's age, sex, parental socioeconomic status, and the highest level of education attained by the parent.

Anthropometry

I measured the participant's weight using digital scales (UC-321, A&D Company LTD, Tokyo, Japan) and height using stadiometers (Surgical and Medical Products No. 26SM, Medtone Education Supplies, Melbourne, Australia). I measured the participant's weight and height as the automated wearable camera position on the chest harness may be impacted due to body size (i.e., certain behaviours may not be captured).

Accelerometer

I collected accelerometer data to provide participants with an activity report as a recruitment tool. I asked participants to wear a GENEActiv accelerometer on their non-dominant wrist for seven days. The GENEActiv (Activinsights Ltd., Cambridge, United Kingdom) is a wrist-worn, battery-powered, tri-axial accelerometer. I collected data in 5-second epochs at a rate of 85.7 Hz. The accelerometer data was not related to the study research questions. The accelerometer was given to the participant at the first home visit.

Automated Wearable Camera

The primary measure of interest was the automated wearable camera. To ensure I selected an automated wearable camera appropriate to measure children's screen use, I created a selection criterion based on my findings from Chapter 2, advice from my supervisors and other researchers in the field, and my experience from testing the coding protocol in Chapter 3. I then trialled several automated wearable cameras in both a lab and home setting and selected the camera that most aligned with the selection criteria. After trialling several cameras, I chose the Brinno TLC130 (Brinno TLC130, Brinno, 55mm x 55mm x 28.2mm, 74.5g, 138° field-of-view) due to its functionally, battery and storage capacity, image quality, and weight of the camera. A summary of the selection criteria and results from the camera trials can be found in Appendix H.

I asked the participants to wear the Brinno TLC130 on a GoPro chest harness during a two-hour direct observation session in their homes. Participants and their parents were given the flexibility to choose the timing of the data collection (i.e., on a weekend, after school, or before school). Before each visit, the camera was programmed using the Brinno phone application. The camera was set to timelapse mode and took images every 2-seconds (no video or audio). The camera timestamp setting was turned on. The camera data was collected using an encrypted SD card.

At each visit, the researcher collected the camera data and converted the timelapse data into images (.jpeg). Participants and their parents were then given the opportunity to review the images and delete any images they did not want the researchers to view, as per ethical guidelines (Kelly et al., 2013). After each visit, the timelapse data was run through a face blurring software (Sanders, 2023). The face blurring software converts the timelapse videos to images and uses a machine learning model to detect faces in these images. The blurred image data (as well as all other data) was stored on a secure network drive. Access to this drive was restricted to members of the research team. This was done in accordance with ethical guidelines to protect participant and third-party privacy (Kelly et al., 2013).

Direct Observation

I compared the automated wearable camera data captured during the observation session to the direct observation data. I structured the direct observation sessions around the participant's usual screen use interactions. The purpose of structuring around the participant's usual screen use was to encourage them to interact with the screen-based media devices as if it was a typical day. At the beginning of each direct observation session, I asked participants what devices they typically used, what type of content they have access to, and what location they used the device in the house. For example, if a participant typically used a tablet, I asked them the following questions:

- Can you show me how you use your tablet?
- How do you sit when using the tablet?
- Where do you usually use your tablet?
- What type of things do you play on your tablet?
- Do you use it for fun or educational purposes?
- What type of things do you do when you use your laptop, like do you ever get a drink or eat food?

I then asked the participants to complete tasks from a standardised list of tasks that centred around their usual screen use interactions to replicate regular behaviour (Appendix I). I also asked participants to complete non-screen sedentary behaviour, such as reading a book, playing with a toy quietly and writing, to differentiate between screen and non-screen behaviours. Each activity was separated by a 1-5 minute rest period to prevent observer fatigue and drift (Haidet et al., 2009). Similar study designs have been used to assess the validity of physical activity assessments in children and adolescent populations (Lyden et al., 2014; Sirard & Pate, 2001). I chose this study design to ensure all relevant screen use behaviours were observed in a home setting while minimising the costs associated with direct observation.

I conducted the direct observation using continuous sampling, where each behaviour change was coded. To code the direct observation data, I used an event-logging software, Behavioral Observation Research Interactive Software (BORIS), which allows researchers to code human behaviour from live or video-recorded observations (Friard & Gamba, 2016). During the observation, I used the BORIS software on a laptop and continuously entered the observation whenever the participant's screen-based behaviour changed while the software simultaneously calculated the elapsed time spent in each behaviour. I recorded for the type of device (i.e., television, computer, mobile device, tablet, handheld game console or smartwatch), location (i.e., bedroom, living room, kitchen/ dining, outside or other room), social environment (i.e., alone, single adult, single child, adults only, children only or mixed ages), associated behaviours (i.e., food or multitask) and non-screen behaviour (i.e., any time the participant was not exposed to a screen-based media device). The observation coding protocol followed the same category definitions as the image coding protocol (described below) to ensure there were no discrepancies. A screen use behaviour change was coded as soon as the participant started or stopped a behaviour. For example, I coded a screen exposure from the moment the participant engaged with a screen-based media device until the participant disengaged with the screen-based media device. The clock on the laptop was synchronised with the automated wearable camera clock. Additionally, at the beginning of each observation session, I showed the time on the laptop to the camera to align the timepoints on the camera and laptop.

Content Recall Interview

Immediately following the observation session, participants completed a content recall interview where they reviewed the images captured within the two-hour observation session. During this interview, I asked participants what type of content they were viewing in the image (i.e., educational, social, or recreational) and the name of the content (i.e., the name of the programme or game). The content recall data was then matched to the direct observation data during data processing. I collected this data as the screen-based media device's content may be hard to observe due to the position of the screen-based media device or may be unfamiliar to recognise during the live observation coding.

Observer Training

Observer training included three stages: (1) familiarisation of the observation protocol, (2) coding practice using still images, and (3) completing practice observation sessions (Myers & Wells, 2015). In Stage 1, I memorised the observation codes and familiarised myself with the event-logging software (BORIS; Friard & Gamba, 2016). Once I memorised all the codes and felt confident using the event-logging software, I completed Stage 2, where I practised coding by viewing images of children using screen-based media devices. In Stage 3, I completed a practice observation session using a member from the research lab. During Stage 3, I discussed any discrepancies in coding with my supervisors and finalised the criteria for observation coding.

Image Coding Protocol

Two coders split the image coding and coded the images using the open-source software Timelapse 2.0 Image Analyser software (ver. 2.3.0.6; Greenberg, 2023). The software allows users to create coding protocols and generates the database while the images are coded. The software allows a maximum of 80 images to be viewed in the image viewer and coded simultaneously. The number of images viewed in the image viewer and coded simultaneously was based on the coder's discretion. The advantages of using coder discretion were that coders could code longer screen exposure episodes in less time than image-by-image coding but could code more complex screen exposures via image-by-coding when necessary.

Images were coded using sections of the coding protocol in Chapter 3. I reduced the number of categories assessed in the coding protocol to ensure each category was adequately powered. In this section, I describe each category of the coding protocol used in this study.

Coders coded each image for a screen exposure. Screen exposure was defined as an event or episode where a person is in the presence of one or more active screen-based media devices, regardless of whether or not the person is consciously attending to the device. Images that did not contain a screen device could still be coded as screen exposure if they occurred in a screen exposure episode and had an image before or after depicting a screen-based media device. For each identified screen exposure, coders were required to code the image for the location, number of devices, type of devices, content classification, social environment, and associated behaviours, respectively. Images were coded in sequential order. Coders were required first to determine if the image was codable (i.e., could be coded as screen exposure). If the image was codable, coders continued coding the image. If the image was determined as uncodable (i.e., did not contain a screen exposure), the coders coded the images as ‘Uncodable’ and moved on to the next image.

Location referred to the place or type of surroundings where the participant’s screen exposure occurred (Watkins et al., 2018). To code location, coders used visual cues from the images to identify where the screen exposure occurred. For example, a common visual cue included the type of furniture depicted in the image. The location codes were limited to areas within the home and included the bedroom, living room, kitchen/ dining room, outside and other rooms.

Coders coded the number of devices present in the image and the type of device. The type of device included television, computer (desktop and laptop), tablet, mobile device (mobile phones and multipurpose devices), handheld game console, and smartwatch.

Content classification referred to whether the screen media was educational, recreational, or social. Educational screen media included creation applications, educational video games, programmes, and internet-based activities where the screen media appears to be educational (e.g., compelling homework online). Recreational screen media included programmes, video games and internet-based activities where the screen media appears to be recreational (e.g., playing video games). Social screen media included social media and communication applications. Coders coded the content classification as unclassifiable if they could not determine the type of content.

Social environment referred to who the child was with during the screen exposure. Social environment codes included alone, single adult, single child, adults only, children only, and mixed ages. To code the social environment, coders used visual cues from the images to identify who the participant was with during the screen exposure. For example, a common visual cue included images depicting a person walking into a room where the participant was located and sitting down.

Associated behaviours referred to co-occurring actions or tasks the child undertakes during the screen exposure. Associated behaviour codes included eating and multitasking. To code associated behaviours, coders used visual cues from the images. For example, a common visual cue to code 'eating' included images depicting food on a plate. Table 4.1 presents an overview of the coding protocol used to code the images.

Table 4.1

Brief overview of the coding protocol used to code the images captured by the automated wearable camera

Device	Content	Location	Social environment	Associated behaviours
Television	Educational	Bedroom	Alone	Eating
Computer	Recreational	Living room	Single adult	Multitask
Mobile device	Social	Kitchen/ dining	Single child	
Tablet	Unclassifiable	Outside	Adults only	
Handheld game console		Other room	Children only	

Device	Content	Location	Social environment	Associated behaviours
Smartwatch			Mixed Ages	

Coder Training

Prior to coding the images, coders completed a coder training session. The coder training session provided information on the ethical guidelines for handling image-based data, familiarisation with the coding manual, software, and popular screen-based media devices and screen media used by the participant's age group. Coders then practised image coding on a set of images (n = 2,654) according to the coding protocol. During this practice session, coders could ask questions about the image coding process. After the set of images were coded, I calculated the image-by-image inter-rater reliability with model answers using percent agreement. The percentage agreement statistic is a standard method to assess inter-rater reliability as it is easily calculated and interpretable (McHugh, 2012). The percentage agreement statistic does not consider the possibility that raters guessed scores and may overestimate the true agreement among raters (McHugh, 2012). Thus, I used percentage agreement for practical implications during the training session and not an image coding quality check.

I recalculated percent agreement after discussing code disagreements, correcting codes, and clarifying the coding protocol and visual cues. I repeated this process until coders achieved a score of ≥ 90% agreement (McHugh, 2012). Once coders completed the training session, they began image coding.

Sample Size Calculation

I calculated the sample size based on the number of observations required to detect a moderate correlation of .6 between the two methods using a sample size calculator for Hypothesis Testing using Pearson's Correlation (Arifin, 2023). In each observation, the camera data and direct observation data are simultaneously captured and compared. A final sample size of 19 observations provided 80% power to detect a moderate correlation of .6 between the two methods, assuming a significance level of .05. I aimed to recruit 10 participants to complete two observations per participant amounting to 20 total observations.

Statistical Analysis

I conducted all analyses using R (v4.3.1; R Core Team, 2023). Descriptive statistics for participant characteristics were calculated using the Base R functions (v4.3.1; R Core Team, 2023).

To address my first research question of the convergent validity of automated wearable camera measurements compared to direct observation measurements, I matched the observation and automated wearable camera data by timestamps. To examine the second research question of whether camera epoch length impacts the measurement of screen use, the data was subset to represent a 10-second, 20-second, and 30-second camera epoch. To do this, I filtered out rows in the automated wearable camera dataset that would correspond with the set camera epoch length. Each filtered dataset was then matched with the observation dataset by timestamps. I excluded any observations that were not matched in all datasets.

To assess the level of agreement between the methods for continuous variables (i.e., duration), I calculated the intraclass correlations (ICC) using a two-way mixed effects single measurement model (Koo & Li, 2016) using the R-package *irr* (v0.84.1; Gamer et al., 2019). The intraclass correlation coefficient is a common statistical method used to assess the agreement of continuous data ranging from 0 to 1, where 1 indicates perfect reliability (Koo & Li, 2016). I interpreted the intraclass correlation coefficients as ICC values less than .5 indicate poor agreement, values between .5 and .75 indicate moderate agreement, values between .75 and .9 indicate good agreement, and values greater than .9 indicate excellent agreement (Liljequist et al., 2019). Further, I compared the agreement between the methods using a Bland–Altman plot (Bland & Altman, 1986, 1995). The Bland-Altman plot is a graphical methodology that compares two different measurement techniques and plots the differences between techniques against the averages of the two techniques. The Bland-Altman plot is used to identify a relationship between the pattern (systematic or proportional) of differences and the magnitude (bias) of measurements and can also identify outliers (Bland & Altman, 1986, 1995). I created the Bland-Altman plot using the R-package *ggplot2* (Wickham, 2016).

To assess the level of agreement between methods for the categorical variables, I calculated the Fleiss' Kappa (*k*) using the R-package *DescTools* (Signorell, 2023). Fleiss' Kappa is a variation of Cohen's Kappa that allows including two or more categories when using nominal data (Fleiss, 1971). I interpreted the Fleiss' Kappa coefficients as kappa values less than .2 indicate poor agreement, values between .21 and .4 indicate fair agreement, values between .41 and .6 indicate moderate agreement, values between .61 and .8 indicate substantial agreement, and values between .81 and 1 indicate almost perfect agreement (Landis & Koch, 1977).

I assessed the correlation of automated wearable camera measurements compared to the direct observation measurements using Pearson correlation coefficient (r) using the *Base R functions* (v4.3.1; R Core Team, 2023). The Pearson correlation coefficient is a common statistical method used to assess the strength of the linear association between two continuous variables ranging from -1 to 1, where 1 or -1 indicates perfect correlation and 0 indicates no correlation (Liu et al., 2016). The sign of the r indicates the direction of the correlation. For instance, a negative r indicates that the variables are inversely related. I interpreted the Pearson correlation coefficient as values less than .1 indicate negligible correlation, values between .1 and .39 as weak correlation, values between .4 and .69 as moderation correlation, values between .7 and .89 as strong correlation, and values between .9 and 1 as very strong correlation (Schober et al., 2018). I excluded participants from the correlation and ICC analysis if they did not perform the behaviour. Due to insufficient data, I excluded smartwatches, handheld game consoles, and educational content categories from all analyses. For all tests, the alpha level was set at .05.

Image Coding Quality Check

I conducted an image coding quality check to assess the level of agreement between coders for each dimension of the coding protocol. Images were repeat coded for two randomly selected visits (n = 6,128 images). Inter-rater reliability was assessed using Fleiss's kappa (R-package DescTools; Signorell, 2023). Interrater reliability between image coders for all dimensions was above .81, indicating an almost perfect level of agreement between coders (type of device: k = .98, 95% CI [.96, 1], location: k = 96, 95% CI [.94, .98], content classification: k = .97, 95% CI [.95, .99], social environment: k = .99, 95% CI [.97, 1.0], and associated behaviours k = .89, 95% CI [.87, .91]).

Results

A total of 46,049 images were collected across 16 observations from 10 participants aged 8 to 11 years (*M* = 10.1, *SD* = 1.10). Six participants (four male) completed two direct observation sessions, and four participants (two male) completed one direct observation. All missed observation sessions were due to time constraints. After matching the automated wearable camera data to the observation data, 43,176 images were included in the analysis.

The average observation session was 102.6 minutes (*SD* = 18.1), with 26.4 minutes (*SD* = 21.6) being non-screen-based behaviours and 76.2 minutes (*SD* = 21.2) being screen-based behaviours. The average number of images captured per visit was 2878 images (*SD* = 578). Nine observations sessions occurred in the afternoon during the school holidays, four observation sessions occurred on a weekday after school, and three observation session occurred on the weekend in the afternoon.

All participants had access to a television, and 90% of participants had access to a tablet device. All participants used home-based screen-based media devices to access recreation content, while only 30% used home-based media devices to access educational content. All parents of the participants had completed a high school certificate, with 70% of parents having a higher education qualification. Almost all participants (90%) had a sibling, with 50% having more than one sibling. Among the participants, there were two sets of siblings: two siblings in one pair, and two siblings in another pair. I have presented the participant characteristics and descriptive data in Table 4.2.

Table 4.2

Participant characteristics and observation descriptive data

Variable	*n*	%	*M*	*SD*
Sex				
Males	6	60		
Females	4	40		
Number of siblings				
None	1	10		
One	4	40		
Two	2	20		
Three	3	30		
Highest education of parents				
School certificate or lower	0	0		
High school certificate	3	30		
Diploma	3	30		
Bachelor’s degree	1	10		
Master’s degree or higher	3	30		
Access to types of content				
Educational content	3	30		

Variable	*n*	%	*M*	*SD*
Recreational content	10	100		
Access to types of screen-based media devices				
Television	10	100		
Tablet	9	90		
Handheld game console	1	10		
Mobile devices	5	50		
Computer	4	40		
Smartwatch	1	10		
Images coded as screen exposure	34,145	74		
Images coded with two screen-based media devices	1,894	4		
Height (cm)			145.8	6.11
Weight (kg)			35.95	8.10
Movement behaviours				
Average daily sleep (hr/day)			8.9	0.68
Average daily sedentary (min/day)			580.36	67.2
Average daily MVPA (min/day)			36	21.53
Images per Visit 1			2990	530
Images per Visit 2			2691	565

Convergent Validity of Automated Wearable Camera Measurements

To address Research Question 1 of the convergent validity of automated wearable camera measurements for assessing screen-based behaviours within children compared to direct observation measurements, I have reported the results of each dimension separately. I have interpreted the results to answer Research Question 1 based on the 2-second camera epoch length.

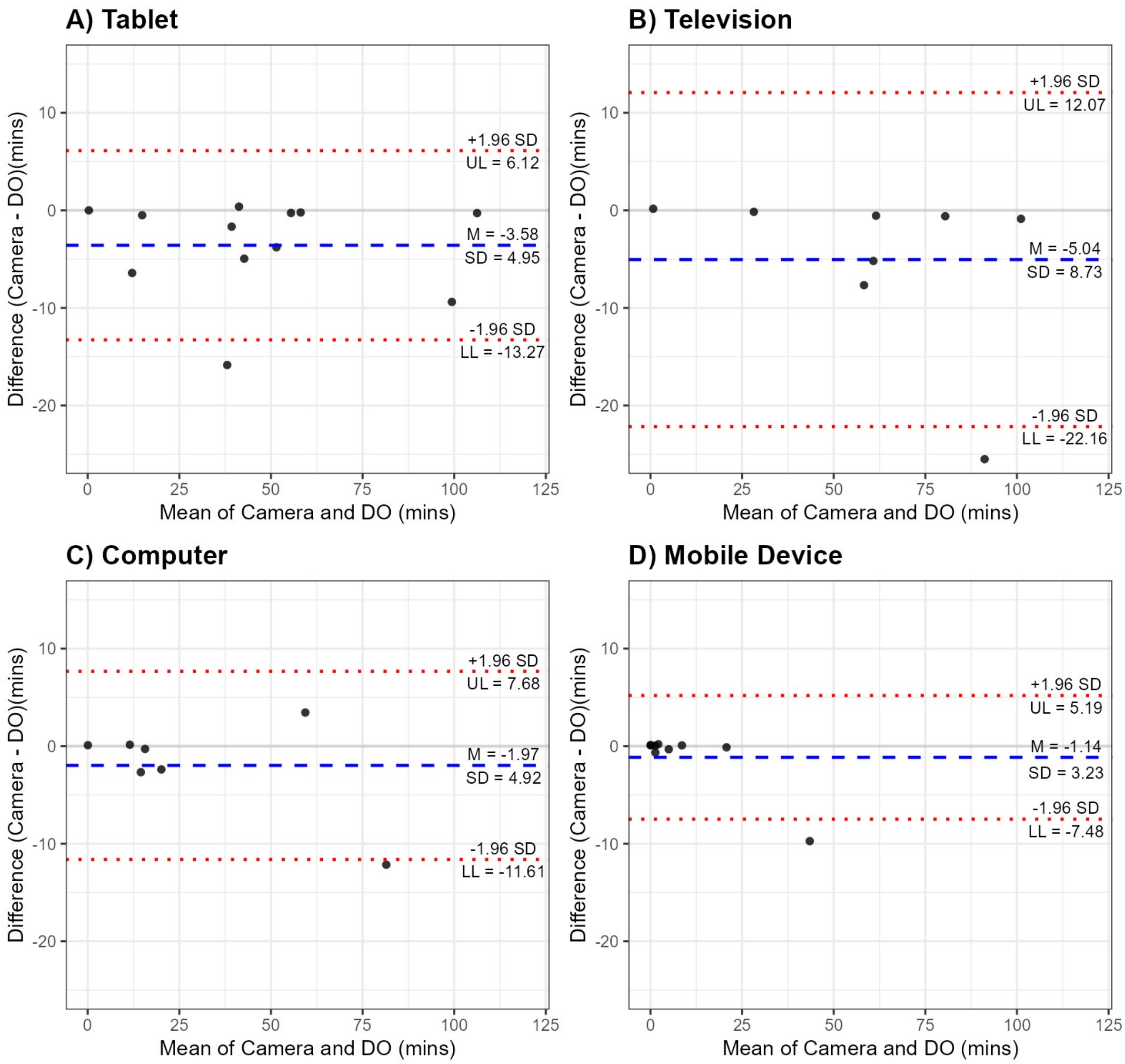
Screen Exposure

For the total screen time duration (i.e., aggregated screen exposure events), I found excellent agreement and very strong correlations between direct observation and camera measurements, ICC = .92, 95% CI [.67, .97] and $r(14) = .94$, $p < .001$, respectively. I found a small systematic bias of -5.12 minutes ($SD = 6.97$) and a wide limit of agreement for camera measurements of total screen time compared to direct observation measurements (upper limit of agreement was 8.55 minutes (+1.96 *SD*) and the lower limit of agreement was -18.78 minutes (-1.96 *SD*).

For device specific screen time duration, I found excellent agreement and very strong correlations between direct observation and camera measurements of the duration of tablet (ICC = .98, 95% CI [.91, .99] and $r(10) = .99$, $p < .001$), television (ICC = .96, 95% CI [.80, .99] and $r(6) = .97$, $p < .001$), mobile device (ICC = .97, 95% CI [.90, .99] and $r(7) = .99$, $p < .001$), and computer screen time (ICC = .99, 95% CI [.93, .99] and $r(5) = .99$, $p < .001$). Further, tablet, television, mobile device, and computer screen time kappa values were above 0.81 indicating an almost perfect agreement level of agreement between methods for categorical variables, .97 [95% CI .96, .98], .99 [95% CI .98, .1], .98 [95% CI .97, .99], .98 [95% CI .97, .99], respectively. As seen in Figure 4.1, I found a small and consistent negative systematic bias between camera and direct observation measurements of tablet, television, mobile device and computer, -3.58 minutes ($SD = 4.95$), -5.04 minutes ($SD = 8.73$), -1.14 minutes ($SD = 3.23$), and -1.97 minutes ($SD = 4.92$), respectively. The results show that television has the widest limits of agreement, followed by tablets, computers, and mobile devices.

Figure 4.1

Bland-Altman plot of Camera and Direct Observation measurements of tablet, television, computer, and mobile device measurement difference against the mean (both in minutes)



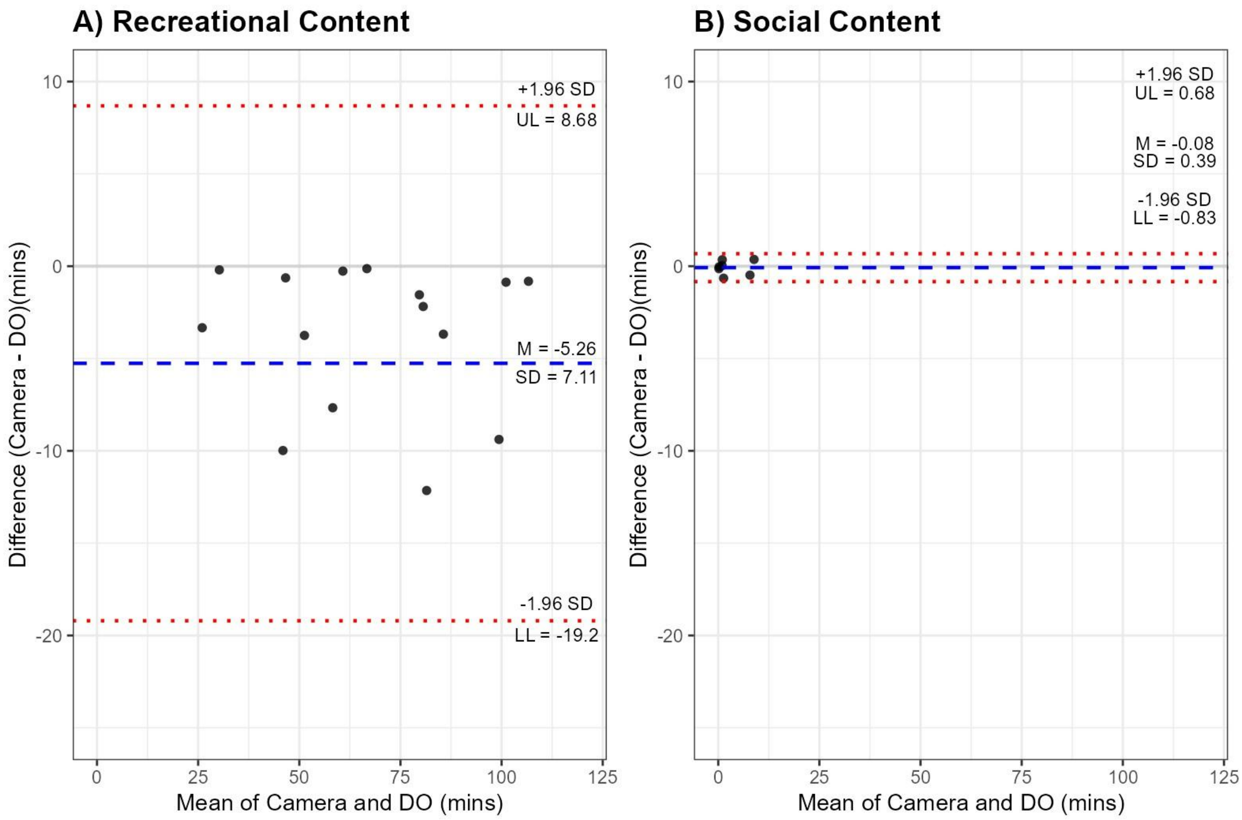
Note. There is one marker for each paired observation. Each marker below the y = 0 line suggests camera measurements were under-reported and each marker above suggests camera measurements were over-reported in comparison to direct observation measurements. A) Tablet has 12 paired observations. B) Television has eight paired observations. C) Computer has seven paired observations. D) Mobile Device has nine paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement, DO = Direct Observation.

Content

For the content of screen use, I found excellent agreement and very strong correlations between direct observation and camera measurements of recreational and social content, ICC = .94, 95% CI [.75, .98] and r(14) = .96, p <.001, and ICC = .99, 95% CI [.97, .99] and r(5) = .99, p <.001, respectively. Further, recreational content and social content kappa values were above .81 indicating an almost perfect agreement level of agreement between methods for categorical variables, .99 [95% CI .98, .1] and .96 [95% CI .95, .97], respectively. As seen in Figure 4.2, I found a small systematic bias between camera and direct observation measurements for recreational content and social content, -5.26 minutes (*SD* = 7.11) and -0.08 minutes (*SD* = 0.39), respectively. Recreational content had wider limits of agreement compared to social content.

Figure 4.2

Bland-Altman plot of Camera and Direct Observation measurements of recreational content and social content measurement difference against the mean (both in minutes)



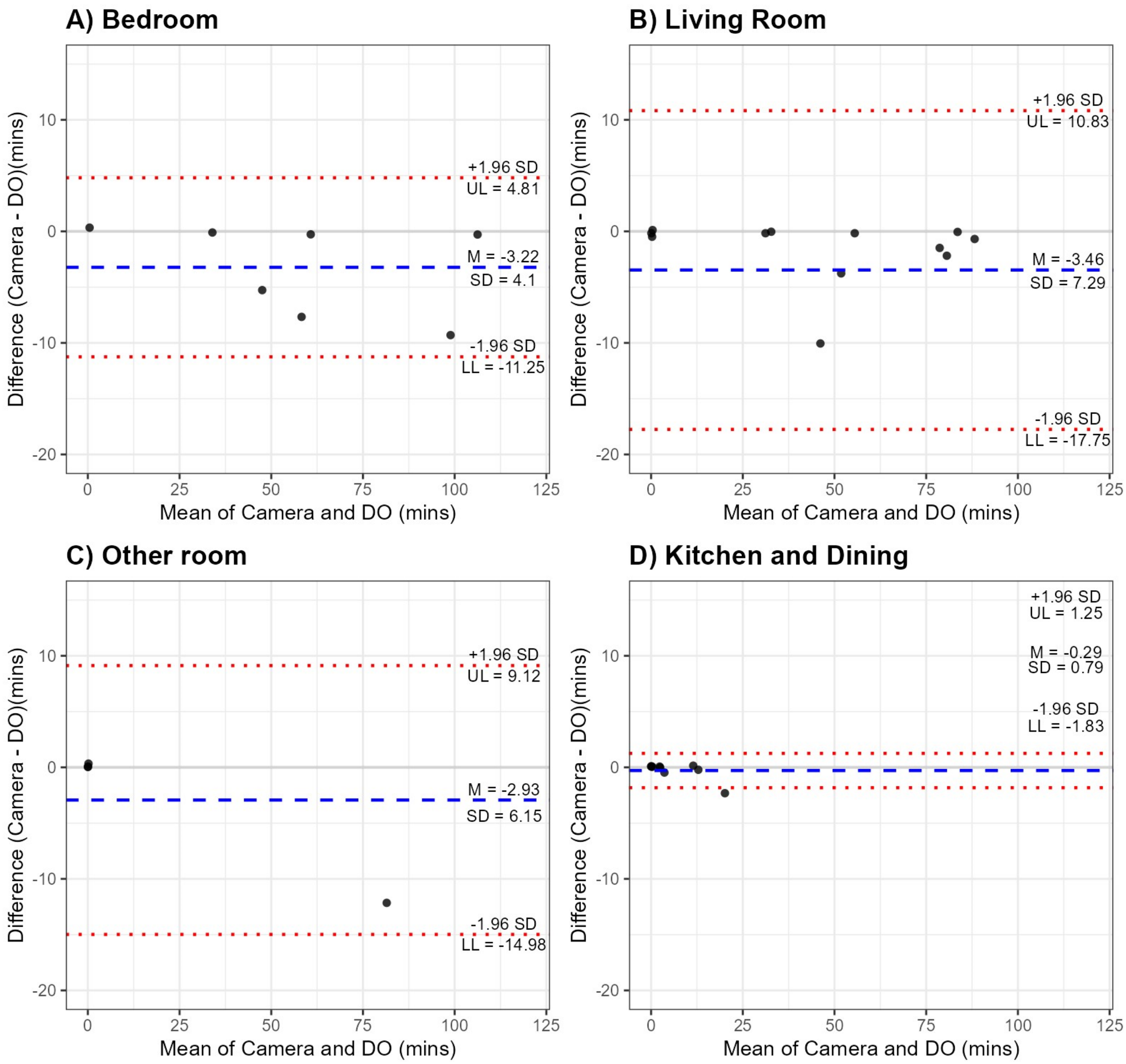
Note. A) Recreational Content has 16 paired observations. B) Social Content has seven paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement, DO = Direct Observation.

Location

For the location of screen use, I found excellent agreement and very strong correlations between direct observation and camera measurements for overall location (ICC = .98, 95% CI [.96, 1] and r(14) = .99, p <.001), bedroom (ICC = .98, 95% CI [.91, 1] and r(5) = .99, p <.001), living room (ICC = .97, 95% CI [.91, .99] and r(11) = .98, p <.001), other room (ICC = .99, 95% CI [.89, 1] and r(2) = .99, p <.001) and kitchen and dining room (ICC = .99, 95% CI [.97, 1] and r(7) = .99, p <.001). Further, overall location, bedroom, living room, other room and kitchen and dining room kappa values were all above .81 indicating an almost perfect agreement level of agreement between methods for categorical variables, .99 (95% CI [.99, 1]), .99 (95% CI [.98, 1]), .99 (95% CI [.99, 1]), .99 (95% CI [.99, 1]), and .98 (95% CI [.97, .99]), respectively. For overall location, I found a small systematic bias between camera and direct observation measurements of -1.62 minutes (*SD* = 2.9), where the upper limit of agreement was 4.06 minutes (+1.96 *SD*), and the lower limit of agreement was -7.3 minutes (-1.96 *SD*). I have presented Bland-Altman plots for each location category in Figure 4.3.

Figure 4.3

Bland-Altman plot of Camera and Direct Observation measurements of bedroom, living room, other room and kitchen and dining measurement difference against the mean (both in minutes)



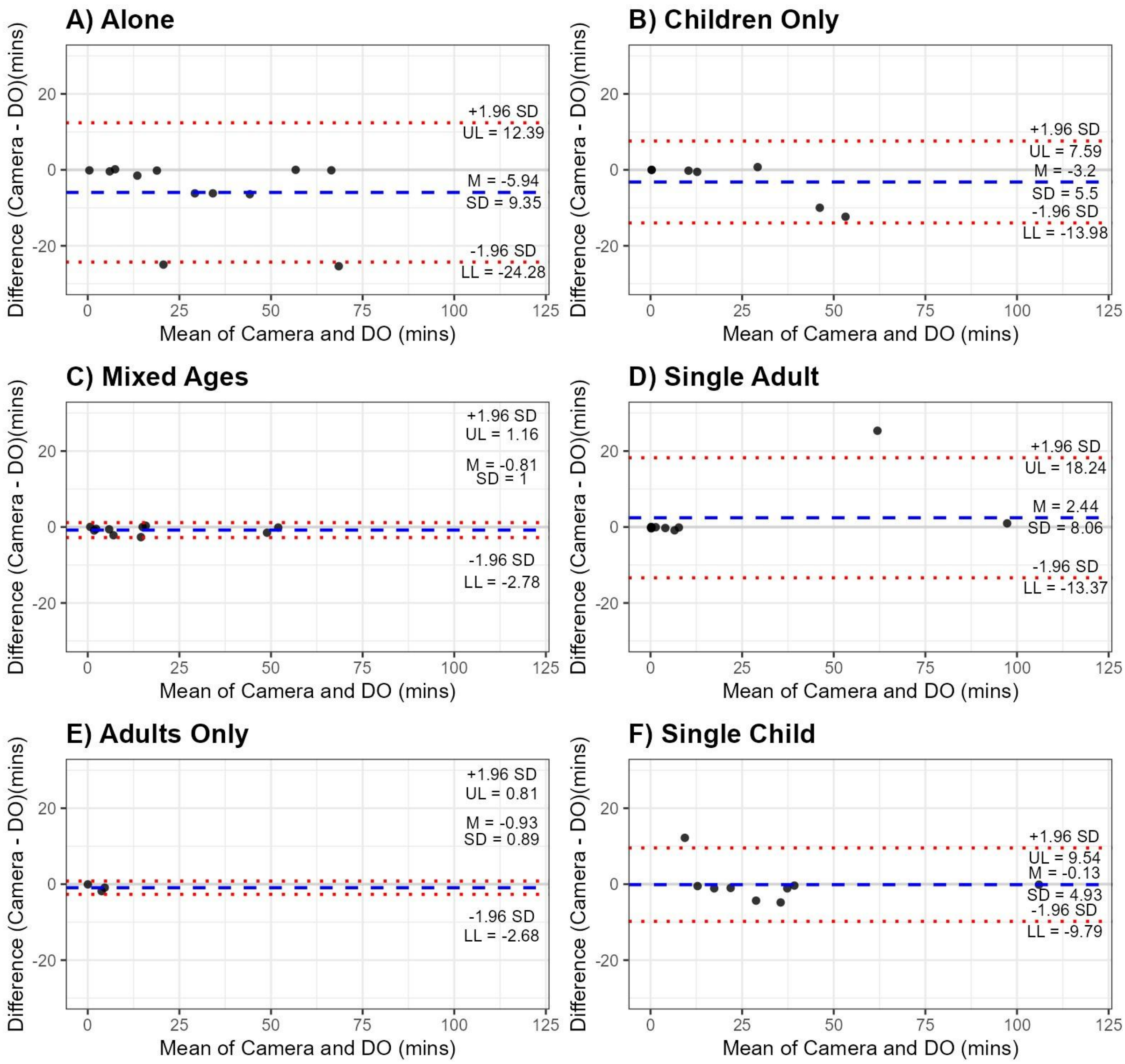
Note. A) Bedroom has seven paired observations. B) Living room has 13 paired observations. C) Other room has four paired observations. D) Kitchen and Dining has nine paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement, DO = Direct Observation.

Social Environment

For the social environment of screen use, I found excellent agreement and very strong correlations between direct observation and camera measurements for the categories alone (ICC = .90, 95% CI [.64, .97] and r(10) = .93, p <.001), children only (ICC = .99, 95% CI [.94, .99] and r(5) = .99, p <.001), mixed ages (ICC = .99, 95% CI [.98, 1] and r(8) = .99, p <.001), single adult (ICC = .97, 95% CI [.89, .99] and r(8) = .97, p <.001), and single child categories (ICC = .98, 95% CI [.94, 1] and r(7) = .99, p <.001). For the adults only category, I found good agreement and very strong correlations between direct observation and camera measurements; however, the results were not significant, ICC = .89, 95% CI [.02, .99] and r(1) = .97, p = 0.154. Further, the categories for alone, children only, mixed ages, single adult, and single child categories all had kappa values above .81 indicating an almost perfect agreement level of agreement between methods for categorical variables, .99 (95% CI [.99, 1]), .94 (95% CI [.93, .95]), .98 (95% CI [.97, .99]), 98 (95% CI [.97, .99]), .97 (95% CI [.96, .98]), respectively. For the adults only category, I found a substantial level of agreement between methods, .66 (95% CI [.65, .67]). I found a small systematic bias between camera and direct observation measurements for the categories alone, child only, adults only, mixed ages, single adult only and single child only, -5.94 minutes (*SD* = 9.35), -3.2 minutes (*SD* = 5.5), -0.93 minutes (*SD* = 0.89), -0.81 minutes (*SD* = 1), -2.44 minutes (*SD* = 8.06), -0.13 minutes (*SD* = 4.93). As seen in Figure 4.4, the single adult only category has the widest limits of agreement.

Figure 4.4

Bland-Altman plot of Camera and Direct Observation measurements of alone, single adult, single child, mixed ages, children only and adults only measurement difference against the mean (both in minutes)



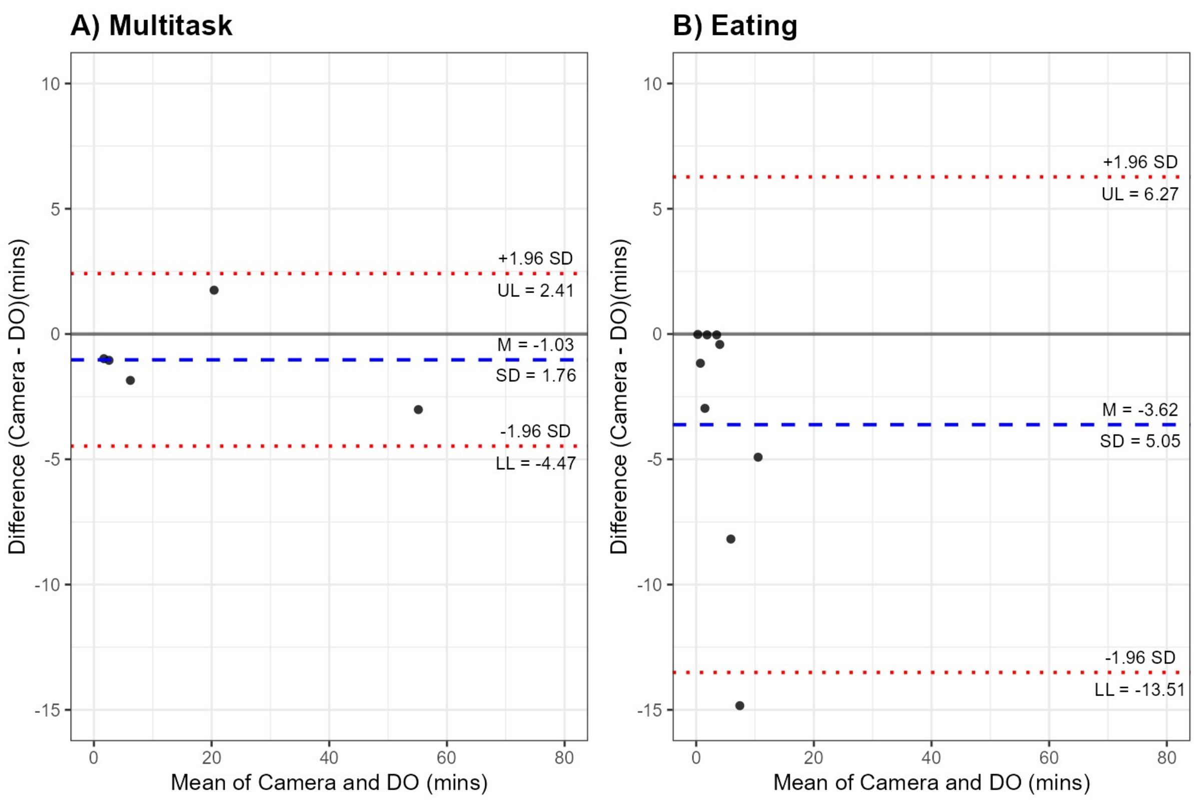
Note. A) Alone has 12 paired observations. B) Children Only has seven paired observations. C) Mixed ages have 10 paired observations. D) Single Adult have 10 paired observations. E) Adults Only have three paired observations. F) Single Child has nine paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement, DO = Direct Observation.

Associated Behaviours

For the associated behaviour categories, I found excellent agreement and very strong correlations between direct observation and camera measurements for the overall associated behaviour and multitasking category, ICC = .93, 95% CI [.79, .97] and r(14) = .94, $p < .001$, and ICC = .99, 95% CI [.97, 1] and r(3) = .99, $p < .001$, respectively. For the eating category, I found poor agreement and weak correlations between direct observation and camera measurements; however, the results were not significant, ICC = .23, 95% CI [-0.25, .71] and r(7) = .37, $p = 0.323$. Further, the overall associated behaviour and multitasking category had kappa values .81 and above indicating an almost perfect agreement level of agreement between methods for categorical variables, .81 (95% CI [.80, .82]) and .94 (95% CI [.93, .94]), respectively. For the eating category, I found a moderate level of agreement between methods, .48 (95% CI [.47, .49]). For the overall associated behaviour category, I found a small systematic error between camera and direct observation measurements of -4.38 minutes (*SD* = 9.22) with wide limits of agreement (upper limit of agreement was 13.7 minutes (+1.96 *SD*) and the lower limit of agreement was -22.46 minutes (-1.96 *SD*). I have presented Bland-Altman plots for the eating and multitask category in Figure 4.5.

Figure 4.5

Bland-Altman plot of Camera and Direct Observation measurements of multitask and eating measurement difference against the mean (both in minutes)



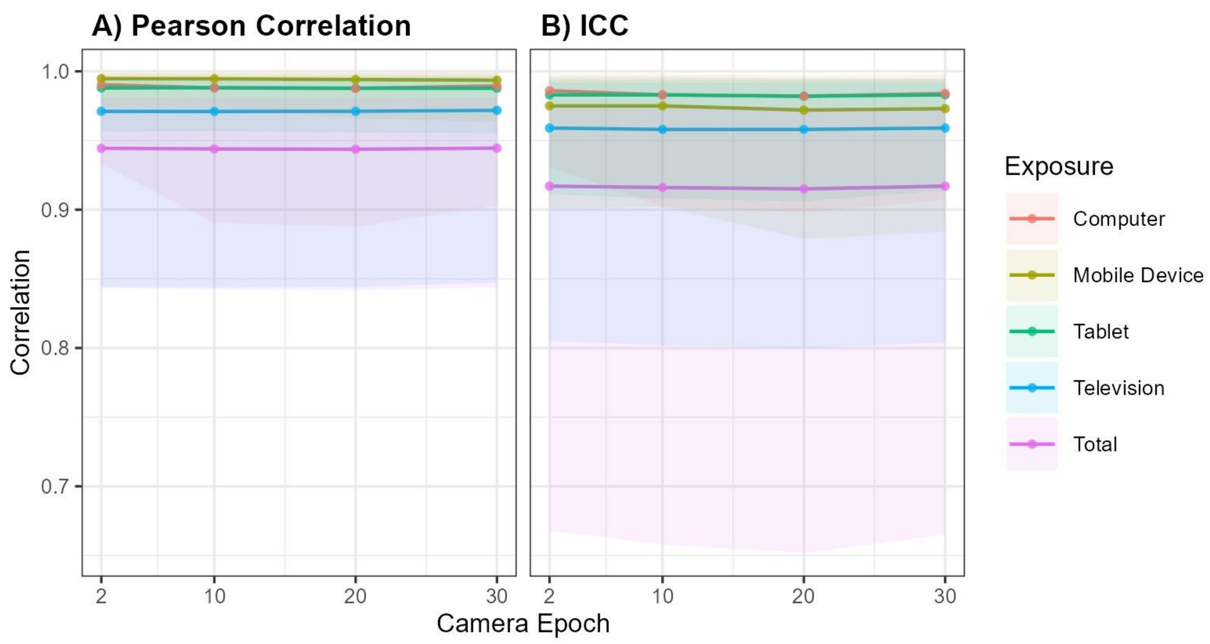
Note. A) Multitask has five paired observations. B) Eating has nine paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement, DO = Direct Observation.

Impact of Camera Epoch on Screen Use Measurements

To address my Research Question 2 of whether camera epoch length impacts camera measurements of screen use, I have reported the results of the total screen exposure and device data that was subset to represent a 2-second, 10-second, 20-second, and 30-second camera epoch. I show the results of the agreement between the two methods for each camera epoch for ICC and correlation results in Figure 4.6. I found excellent agreement and very strong correlations between direct observation and camera measurements of total screen exposure, tablet exposure, television exposure, mobile device exposure and computer exposure for all camera epochs (i.e., 2-second, 10-second, 20-second, and 30-second). I found very small differences for correlation and ICC results between camera epochs. The correlation coefficient and ICC were slightly smaller at a 20-second camera epoch, compared to 2-second, 10-second, and 30-second camera epochs.

Figure 4.6

Pearson correlation (r) and ICC results for total and type of device screen exposure at 2-second, 10-second, 20-second, and 30-second Camera Epochs

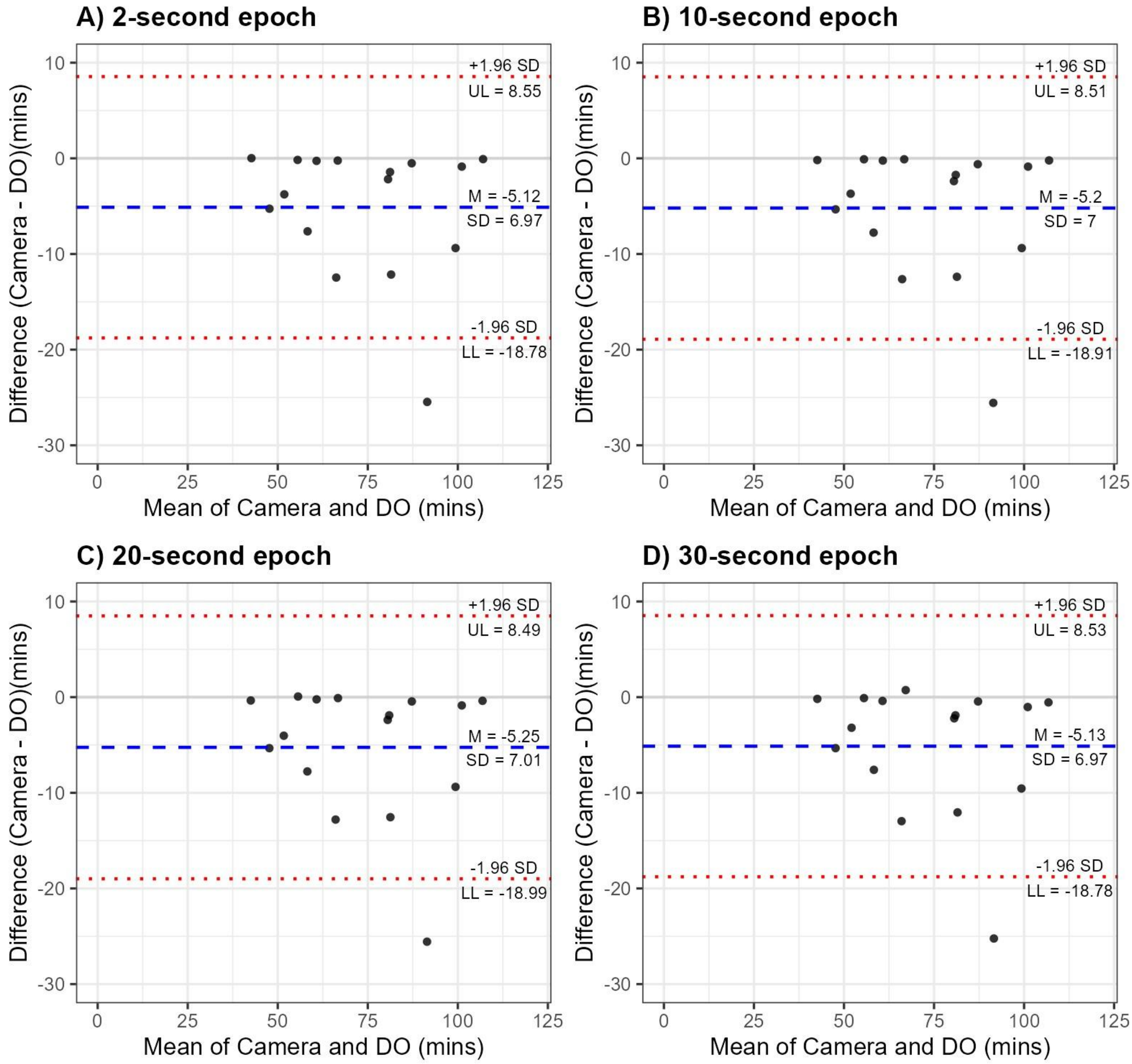


Note. All p-values = ≤.001.

As seen in Figure 4.7, I found a small and consistent negative bias across the epochs for total screen time duration measurements. The mean difference between the camera and direct observation measurements were slightly larger at a 20-second camera epoch, compared to 2-second, 10-second, and 30-second camera epochs.

Figure 4.7

Bland-Altman plot of Camera and Direct Observation measurements of Total Screen Time Duration at 2-second, 10-second, 20-second, and 30-second camera epoch measurement difference against the mean (both in minutes)

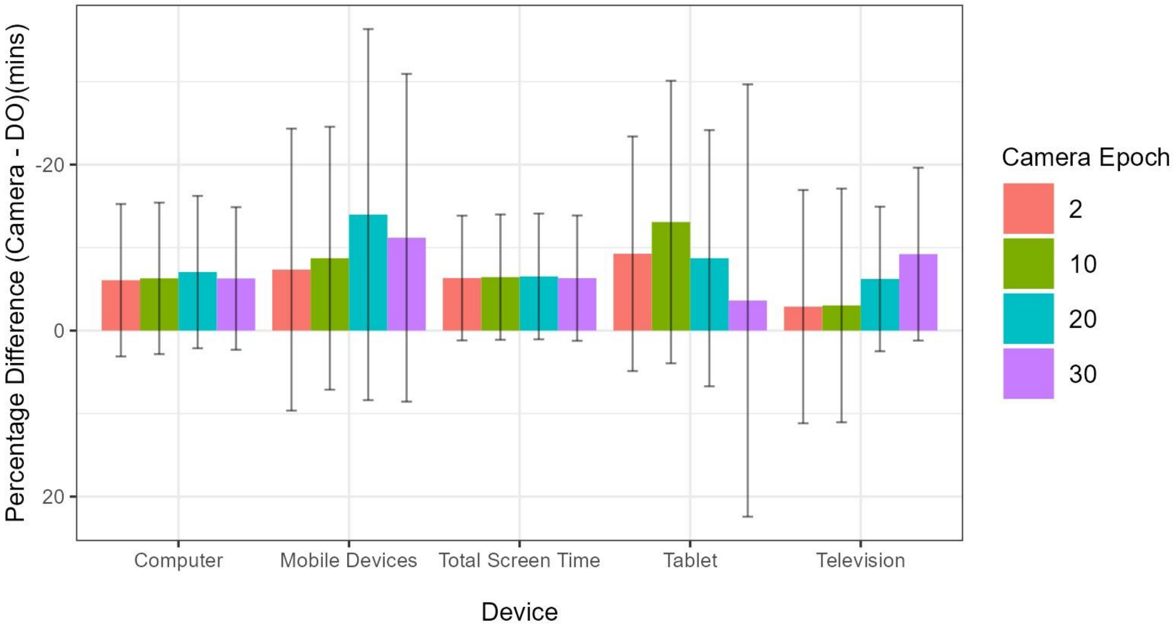


Note. There is one marker for each paired observation (n = 16). Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement, DO = Direct Observation.

I observed similar patterns across television, mobile device, and computer categories but not tablet, where the mean difference between the camera and direct observation measurements was slightly larger at a 20-second camera epoch, compared to 2-second, 10-second, and 30-second camera epochs. For television duration, the mean differences at 10-second and 30-second epochs were slightly smaller than the difference at 20-second epoch, -2.56 minutes (*SD* = 6.53, 95% CI [10.24, -15.36]), and -2.56 minutes (*SD* = 6.46, 95% CI [10.1, -15.23]) compared to -2.59 minutes (*SD* = 6.53, 95% CI [10.2, -15.39]), respectively. Similarly, for mobile duration, the mean differences at 10-second and 30-second epochs were slightly smaller than the mean difference at 20-second epoch, -0.64 minutes (*SD* = 2.41, 95% CI [-5.37, 4.09]), -0.67 minutes (*SD* = 2.46, 95% CI [4.15, -5.5]), compared to -0.72 minutes (*SD* = 2.49, 95% CI [4.17, -5.6]), respectively. For computer duration, the mean differences followed a similar pattern. At 10-second, 20-second, and 30-second intervals were, -0.9 minutes (*SD* = 3.32, 95% CI [5.6, -7.4]), -0.95 minutes (*SD* = 3.36, 95% CI [5.64, -7.54]), and -0.91 minutes (*SD* = 3.2, 95% CI [5.36, -7.18]), respectively. For tablet duration, the mean differences at 10-second and 20-second were slightly larger than the differences at 30-second epochs, -2.71 minutes (*SD* = 4.51, 95% CI [6.13, -11.56]) and -2.75 minutes (*SD* = 4.57, 95% CI [6.21, -11.72]), compared to -2.64 minutes (*SD* = 4.57, 95% CI [6.31, -11.59]), respectively. To further visualise the patterns across the categories observed for each camera epoch length, I have shown the average percentage differences between automated wearable camera and direct observation measurements for each type of device and epoch length in Figure 4.8.

Figure 4.8

Average percentage differences of Camera and Direct Observation measurements for the different types of screen time duration at 2-second, 10-second, 20-second, and 30-second camera epoch.



Discussion

The primary purpose of this study was to evaluate the convergent validity of automated wearable camera measurements for measuring children's screen use with direct observation measurements. The secondary purpose was to examine the impact of the camera epoch on the camera's measurements of screen use. In this study, I used the coding protocol in Chapter 3 to classify the type of device, content, location, associated behaviours, and social environment captured in the imaged-based data.

Automated wearable cameras can accurately measure key aspects of screen use, including aggregated screen time, type of device, and location, when compared to direct observation. Consistent with my hypothesis, overall, there was an excellent agreement and very strong correlation between the automated wearable camera and direct observation measurements for most screen-use behaviours. The results from the study suggest that automated wearable cameras have excellent agreement and very strong convergent validity with direct observation measurements for total screen time duration, device duration and type, and location. Supporting these results, other studies using automated wearable cameras have reliably identified the location (Leask et al., 2015; Signal, Stanley, et al., 2017; Watkins et al., 2018), the type of devices and duration of screen use (Hänggi et al., 2020; Smith et al., 2019). However, to my knowledge, this is the first study to demonstrate the convergent validity of automated wearable camera measurements for screen time duration, type of device, and location.

Automated wearable cameras demonstrate a strong potential for accurately measuring some content and contextual factors of screen use (content type, social environment, and associated behaviours), when compared to direct observation. However, associated behaviours, such as eating, need further refinement. I hypothesised that automated wearable camera and direct observation measurements would have substantial and moderate correlations for measuring the content, social environment, and associated behaviours. The results from the study suggest that automated wearable cameras have excellent agreement and very strong convergent validity with direct observation measurements for content (i.e., classifying content as recreational or educational), social environment (i.e., alone, children only, mixed ages, single adult, and single child categories) and associated behaviours (i.e., overall category and multitask category) within the confines of this study. These findings align with evidence that suggests that automated cameras can reliability identify the social context of health behaviours in adult populations (Gemming, Doherty, et al., 2015; Leask, 2015). However, automated wearable camera's ability to capture eating behaviours remains limited. In this study I found poor agreement and weak correlations for the eating category of associated behaviours; however, the results were not significant. The poor agreement between automated wearable cameras and direct observation measurements for coding eating (i.e., eating while using a screen-based media device) may be due to the increased challenge of capturing the health behaviour due to camera position and difficulty in coding. For instance, Gemming, Rush, et al. (2018) highlighted challenges in capturing eating behaviours which may be missed due to the position of the camera on the body (i.e., lens angle may be affected by posture, foods in bowls or the height of tables). While Cowburn et al. (2016) found that coding images alone (i.e., without participant verification) can present a challenge, especially when searching for short sporadic behaviours such as snacking. Other studies have suggested that automated wearable cameras should be used in combination with image-assisted recall to improve validity and reliability of reporting of dietary behaviour (Gemming et al., 2013; O'Loughlin et al., 2013; Zhou et al., 2019). Thus, while automated wearable cameras show strong potential for accurately measuring content and contextual factors of screen use (content type, social environment, and multitasking behaviours), their ability in capturing associated behaviours such as eating behaviours, remains limited and requires further methodological refinement.

A camera epoch of 20-seconds may introduce more systematic bias compared to camera epoch of 2-seconds, 10-seconds or 30-seconds. However, the differences may be too small to be meaningful. I hypothesised that camera epoch length would affect camera measurements of screen time, where a longer camera epoch would result in lower accuracy than shorter camera epoch lengths. I found that, overall, there is a small, consistent negative bias across all epochs (i.e., 2-second, 10-second, 20-second, and 30-second). These findings suggest automated wearable cameras may underestimate screen use duration by approximately 5 minutes when compared to direct observation measurements, with small variations observed in camera epoch and device-specific exposures. A similar problem was faced in a study comparing automated wearable cameras measurements of travel journey to school with direct observation (Kelly et al., 2012). Kelly (2012) found that automated wearable camera measurements of travel duration may introduce a small systematic bias through the camera epoch as the journey would lose 0 to 15 seconds at the start of the journey and gain 0 to 15 seconds at the end. Further, during data analysis, I found that the camera epoch lengths were sometimes inconsistent. For instance, sometimes the camera randomly took an image at a 4-second epoch before returning to the 2-second epoch length. This may have introduced a small systematic bias into the calculation of the screen time measurements as the screen exposure episode would lose time at the start of the screen exposure and gain time at the end. Further, I found slight differences between camera epoch lengths for correlation, ICC, and Bland Altman plot results. These findings suggest a camera epoch of 20-seconds will have a slightly larger systematic bias compared to a camera epoch of 2-seconds, 10-seconds or 30-seconds for total screen time, television screen time, computer screen time and mobile device screen time. However, the differences may be too small to be meaningful. Despite this, the impact of camera epoch length should be taken into consideration when designing studies, with the choice of epoch length balanced against the research objectives and data processing capabilities.

Automated wearable camera measurements of total screen time may be more appropriate for large-scale studies rather than studies requiring estimates of an individual's total screen time. I found a small negative systematic bias and a wide limit of agreement for between automated wearable camera and direct observation measurements of total screen time duration, television screen time, recreational screen time, living room, eating, and being alone or with a single adult. The wide limits of agreement indicate variability in the discrepancies between the two methods (Mansournia et al., 2021). A potential source of variability may be due to the type of device in each of those contexts. For example, television viewing appears to produce wider limits of agreements compared to other screen-based media devices. The wide limits of agreement may be due to the differences in television viewing behaviours compared to other devices. For example, children tend to watch television lying down, which may block the camera lens. Televisions are also often used as a second device in media multitasking or playing in the background while completing other tasks. These behaviours impact the camera's ability to capture the screen-based media device and increases the difficulty in coding as it requires more subjective input from the coder. The wide limits of agreement suggest large random errors at an individual-level, indicating that automated wearable cameras may be inappropriate for assessing total screen time at an individual-level. The small difference in means between automated wearable camera and direct observation measurements suggest that automated wearable camera measurements of total screen time may be a better measurement at a group-level rather than an individual-level. The finding of large random errors is consistent with other studies using automated wearable cameras with other time-use studies and image-assisted recalls; however, the causes remain unclear (Bulungu et al., 2022; Kelly et al., 2012). Thus, automated wearable camera measurements of total screen time may be more appropriate for large-scale studies rather than studies requiring estimates of an individual's total screen time; however, further research investigating the causes of large random errors is needed.

Limitations and Future Research

In the process of validating a measurement instrument, evidence from different types of validity should be used to assess the degree of validity of the instrument in the specific context and population (de Vet et al., 2011). A limitation of this study was that I only assessed the convergent validity of automated wearable camera measurements in children aged 8-11 years old in a home setting. As such, future research on automated wearable camera measurements of screen use needs to assess different types of validity, such as content validity and criterion validity, to conclude the degree of validity of automated wearable camera measurements. Further, the small sample size in this study limited my ability to investigate the convergent validity of all the categories of the coding protocol in Chapter 3. In this study I was only able to assess the convergent validation of screen time, type of device, location, recreational and social content, and social environment (i.e., who the participant is with). Consequently, the validity of automated wearable cameras measurements of screen use behaviours such as educational content, social interactions, and device attention, is unknown. Thus, future research that assess educational content, social interactions, and device attention using the coding protocol presented in this study should interpreted with caution. Further research is needed to investigate the ability of automated wearable cameras to measure screen use behaviours such as educational content, social interactions (i.e., co-viewing and co-participating), and device attention.

Another limitation was that some of the image coding was unblinded. While most of the image coding was blinded (i.e., conducted by a coder who had not conducted the direct observation), I coded two observations and conducted all direct observations. While this should not affect my adherence to the coding protocol, it may have influenced my decision-making when coding the images. To minimise the chance of bias, I conducted the image coding at least two weeks after the observation session. Additionally, the inter-rater reliability results from the image coding quality check for all dimensions were above .81, indicating an almost perfect level of agreement between coders, suggesting there was minimal difference between the blinded and unblinded coders.

One of the key difficulties in direct observation research is that the results are susceptible to the Hawthorne effect (McCambridge et al., 2014). The observation sessions were structured to simulate the participant's usual screen interactions. The purpose of structuring around the participant's usual screen use was to encourage them to interact with the screen-based media device as if it was a typical day. However, participant's may have changed their behaviour due to their awareness of being observed. Additionally, the direct observation sessions were limited to a 2-hour period, which may not reflect the variability of children's screen use behaviours throughout the day. Screen use patterns can fluctuate throughout the day based on factors such as, the time of day, day of the week, and contextual factors such as parental supervision (Thomas, 2022). As a results, the behaviours observed during the direct observation session may not accurately represent a child's typical screen use. Moreover, while direct observation is often considered as the 'gold standard' measurement of screen use, I could not find a pre-existing observation coding protocol that measured the location, type of device, content, and associated behaviours of screen use. Therefore, I tested and developed my own observation protocol based on the image coding protocol. As my observation protocol has not been validated, it may have introduced measurement error into the results. For this reason, I decided to assess the convergent validity of the measurements rather than the criterion validity. Further, due to how I structured the direct observation coding in the BORIS software, I could not measure media multitasking as a standalone variable (i.e., separate from total screen use or devices). Thus, while the original coding protocol can account for media multitasking, I could not evaluate the convergent validity of automated wearable camera measurements of media multitasking separately in this study. Media multitasking was instead included in the total screen use and device measurements. This limits our ability to analyse the effects of media multitasking separately from total screen use. More than ever before, young people are media multitasking (i.e., switching between media on a single device or using multiple screen-based media devices simultaneously; van der Schuur et al., 2017). A systematic review of 56 studies (correlational and experimental studies) examined the possible consequences of media multitasking within adolescents' cognitive control (i.e., ability to sustain attention), academic performance, and socioemotional function, and found that media multitasking is negatively related to aspects of cognitive control, academic performances, and emotional functioning and sleep (van der Schuur et al., 2015). Given these findings, it is important to measure media multitasking separately from total screen use. Future research should assess the validity of automated wearable camera measurements of media multitasking.

The measurements for the 10-second, 20-second, and 30-second camera epochs should also be interpreted with caution. The images were coded at a 2-second epoch. I then subset the data to represent a 10-second, 20-second, and 30-second camera epoch. Consequently, all images were coded with context cues based on a 2-second interval.

This study only included participants aged 8 to 11 years old. To be included participants had to have access to at least one screen-based media device and have a parent/caregiver available for the direct observation session. Children were excluded if they did not have access to any screen-based media devices at home. Thus, this limits the results of this study to children who have access to screen-based media devices and a supportive family environment. Additionally, the majority (70%) of participants came from families where at least one parent held a higher education qualification, and almost all participants (90%) had a sibling, with 50% having more than one sibling. As a result, the findings of this study can only be generalised to this population. Future studies should consider assessing the validity of the automated wearable camera measurements in younger children and adolescents, as well as in more diverse populations.

Finally, I did not control for potential confounders, such as age, sex, ethnicity, socioeconomic status, parental screen use habits, or children’s screen use preferences, which may influence the results (Stiglic & Viner, 2019). The absence of these covariates may impact the validity of the results, as it limits the ability to determine if the findings are solely a result of the observed behaviour or if they are influenced by other unmeasured factors

(VanderWeele. 2019). The lack of control for confounders may lead to biased estimates, reducing the reliability and generalisability of the findings. Future research should include a wider range of covariates, including age, sex, socioeconomic status, parental screen use habits, or children's screen use preferences, to account for potential confounders. Incorporating these covariates will lead to more robust findings on the validity of automated wearable camera measurements of children's screen use.

Implications in Research and Policy

At present, the most common way to measure children's screen use is through self- and proxy-reported measurement instruments, which are prone bias and inaccuracies due to recall (i.e., recall bias, misclassification bias, and social desirability bias; Hardy et al., 2013). Other device-based measurements of screen use (i.e., smartphone usage applications or television monitors) are limited to single devices and do not take into account context of screen use (Perez et al., 2023). The findings from the current study suggest automated wearable cameras may offer a solution that can accurately measure children's screen use alongside important contextual information (i.e., location, social environment, associated behaviours, content, or device type). This allows researchers, parents and policymakers to gain a more detailed and accurate understanding of screen use behaviours. Thus, automated wearable cameras may not only enhance the robustness of the data collected in the literature, but also provide policymakers with richer insights into the nuanced patterns of children's screen use.

Conclusion

Automated wearable cameras may offer a solution that provides an accurate measurement and is more practical than the current screen use methodology. The findings from this study demonstrated that automated wearable camera measurements of total screen time, type of device, social environment, content, associated behaviours, such as multitasking, and the location of the screen use show strong convergent validity with direct observation measurements. However, I found weak convergent validity for food-related behaviours. The findings indicate automated wearable cameras may underestimate total screen time by approximately 5 minutes when compared to direct observation measurements, with variations observed in camera epoch and device-specific exposures. Moreover, the findings indicate that when compared to direct observation measurements, automated wearable camera measurements of total screen use were accurate at the mean group level but may be imprecise at the individual level. Therefore, automated wearable camera measurements of total screen time may be more appropriate for large-scale studies rather than studies requiring estimates of an individual's total screen time. Given the challenge of accurately measuring children's screen use behaviours, these cameras have the potential for accurately measuring complex screen use behaviours such as the content and context of the behaviour.

Chapter 5: Comparing the Automated Wearable Camera Measurements with Self-Report

Introduction

As I outlined in Chapter 1, the inconsistent findings in screen use research may be due to the way screen use has been measured (Kaye et al., 2020). Currently, most screen use research is based on self- or proxy-reported data, which many researchers concede as suboptimal (Hale & Guan, 2015; Stiglic & Viner, 2019). Questionnaires typically ask participants to estimate their screen use behaviours retrospectively. For example, a standard question in self-reported measurements of screen use includes “How many hours of screen time did you use in a typical day last week?” (Kaye et al., 2020). Self- and proxy-reported measurements are more affordable and accessible for researchers than other screen use measurement instruments (i.e., direct observation or television monitors; Hardy et al., 2013). However, a major concern with these measurement instruments is that they are susceptible to bias and inaccuracies, including social desirability and recall bias (Hardy et al., 2013).

Self-reported estimates of sedentary behaviour are prone to inaccuracies and bias, especially for children (Ainsworth et al., 2012; B. Clark et al., 2011). Child-reported measurements have been shown to be an unreliable and inaccurate approach to screen use measurement due to the limited cognitive capacity and increased recall bias among paediatric populations in research (Atkin, Ekelund, et al., 2013; Saunders et al., 2011). For example, Saunders et al. (2011) found that some children have reported unrealistically high amounts of daily screen time (e.g., 13.5 hours per day) when estimating their screen time. Further, self-reported screen use measurements may be susceptible to recall bias due to the large role screen-based media devices play in children’s lives (Schwarz, 2007). Studies suggest that the accuracy with which individuals can recall daily behaviours becomes increasingly challenging as individuals are expected to report over extended periods (Schwarz, 2007).

Given that screen-based media devices are now ubiquitous in childhood and adolescence, individuals' ability to accurately recall their use may be reduced. Further, screen use behaviours are often performed simultaneously with other behaviours (i.e., watching TV while eating dinner), making it more difficult to accurately recall behaviours from memory (Schwarz, 2007).

Proxy-reported (i.e., parent-reported) measurements are similarly flawed. Studies suggest parent-reported screen use may be prone to misclassification bias due to parents not being aware of their child's actual behaviour (e.g., children may engage with screens in separate rooms or at their friends' homes; Jordan et al., 2007). For instance, when looking at television use, Robinson et al. (2006) found that compared to an electronic television monitor, parents overestimated screen time by four hours per week if there was no television in the child's bedroom and underestimated by three hours per week if there was a television in the bedroom. Radesky et al. (2020) found that 34.8% of parents overestimated and 35.7% underestimated their child's mobile device use compared to applications that tracked children's mobile device use. Further, parent-reported screen use measurements may also be susceptible to social desirability bias. Social desirability bias refers to the tendency of individuals to respond in a manner that they believe is more acceptable or favourable rather than respond in a manner that reflects their true thoughts and behaviours (Grimm, 2010). Children not meeting the current national and organisation screen time guidelines ($\leq$ 2 hours) is often labelled as an undesirable behaviour associated with adverse health and behaviour outcomes (Houghton et al., 2015). Therefore, parents may be more likely to falsely report screen use behaviours that align with the national screen time guidelines. Thus, the most common methods for assessing children's screen use (i.e. self- and parent-reported) are known to be prone to inaccuracies and bias.

Recent technological advances (e.g., television monitors, smartphone tracking applications, and computer software programs) have allowed researchers to determine the accuracy of self- or proxy-reported estimates of screen use by comparing these estimates to device-based measurements (Perez et al., 2023). However, such methods are limited to single device measurements and cannot be applied to participants' total screen use across different devices or platforms (Perez et al., 2023). For example, Verbeij et al. (2021) examined the associations between self-report measurements and smartphone tracked data for specific social media platforms, including Snapchat, Instagram and WhatsApp use in adolescents. They found that adolescents overestimated their use of social media platforms compared to smartphone tracked data. However, Wade et al. (2021) examined the associations between self-report measurements and a smartphone application that tracked participants' smartphone use and found that children underestimated their overall smartphone use. As such, estimates of individual device or platform use may not be reflective of estimates for all screen use (Mahalingham et al., 2023). Further, most self- and proxy-reported measurements of screen use have used aggregated 'total' screen time measurements that measure daily or weekly screen use rather than measurements from specific devices or platforms (Griffioen et al., 2020). Therefore, to establish the accuracy of self- or proxy-reported measurements of screen use, it is important to use measurements that are comparable to ones used in most screen use research.

A crucial step in improving our understanding of the impact of screen use on children's health and behaviour outcomes is accurately measuring the context and duration of screen use across multiple devices (Kaye et al., 2020; Odgers & Jensen, 2020). Automated wearable cameras present an alternative approach that may overcome some of the challenges of other device-based measurements of screen use. In Chapter 3 of this thesis, I demonstrated how I developed and tested a coding protocol for coding images captured by automated wearable cameras to measure the type of device, content, and context of children's screen use. In Chapter 4, I evaluated the validity of a coding protocol for coding images captured by automated wearable cameras. I found that automated wearable camera measurements of screen time duration, type of device, social environment, content, associated behaviours, such as multitasking, and the location show strong convergent validity with direct observation measurements. To address the current literature gaps, this study aimed to examine if screen use measurements from device-based measurements were similar to self- and parent-report measurements. Specifically, the primary aim of the study was to examine if measurements of screen use from automated wearable cameras were similar to a self- and parent-reported measurement of screen use.

Research Questions and Hypotheses

Research Question

Are estimates of screen use from automated wearable cameras (i.e., device-based) measurement of screen use similar to self- and proxy-reported measurements of screen use?

Hypothesis

Given that prior research has found low correlations between device-based measurements and self-report, I hypothesised that there would be a weak correlation ($r = \leq .3$) between automated wearable camera measurements of screen use and child- and proxy-reported measurements for aggregated screen time, educational screen use, recreational screen use, and social screen use. I further hypothesised that there would be a weak correlation ($r = \leq .3$) between automated wearable camera measurements of screen use and child- and proxy-reported measurements for activity specific screen time, including total gaming, total browsing, programme viewing and communication (i.e., messaging).

Methods

Participants and Recruitment

I used data collected at baseline from the KidVision project. The KidVision project is an Australian Catholic University project that uses automated wearable cameras, accelerometers, and parent- and child-reported questionnaires to measure children's screen use in Sydney, Wollongong, and Central Coast in New South Wales, Australia. The KidVision project aims to evaluate the effect of screen exposure on children's educational, developmental and behavioural outcomes over 5 years of development. The study uses a longitudinal design, and collects data at three time points, 12 months apart.

For recruitment, we initially tried to recruit participants through Independent and Catholic primary schools located in Sydney, Wollongong, and Central Coast in New South Wales; however, at the time, schools were not interested in participating due to the disruption of COVID-19. Participants were instead recruited from the community through a recruitment agency, Trialfacts (https://trialfacts.com/). The recruitment agency advertised the project through mailing lists and online advertising. The inclusion criteria were children aged 7-10 years old in Greater Sydney, Central Coast or Wollongong, New South Wales. The exclusion criteria were children diagnosed with a neurodevelopmental disorder with significant problems with vision, hearing, cognition, balance, or movement.

The research team contacted eligible participants. Participants were blinded to the purpose of the study to reduce the likelihood of reactivity from wearing a camera. This was done by stating the study was testing a child's reaction to the world during the recruitment process. This blinding technique has been used in a previous automated wearable camera study (Signal, Smith, et al., 2017). Participants received a $20 supermarket gift card compensation for their time at each time point.

The KidVision project began data collection in September 2022 and is currently ongoing. The present study uses demographic data, automated wearable camera data, and parent- and child-reported questionnaire data collected at baseline from the first 27 participants of the KidVision project.

Ethical Approval

Written consent was obtained from all the participants before the commencement of data collection (see Appendix J for the participant information letter, child information letter, consent form and assent form). The procedures for this study were guided by international ethical guidelines for automated wearable cameras (Kelly et al., 2013), and were approved by the Australian Catholic University Ethics Committee (Approval #2020-142H). See Appendix K for ethics approval.

Measures and Procedures

Data collection for the KidVision project occurred over eight days, with two visits for each participant. The visits were conducted by a research team member in the participants' homes or at the Australia Catholic University's North Sydney or Strathfield campus according to the participant's preferences. At the first visit, a research team member provided information on the study protocol and how to operate the camera. At the second visit (approximately eight days following the first), a research team member processed the images and asked the parent and child to complete the parent- and child-reported screen use questionnaire.

Automated Wearable Cameras

The device-based measurement of screen use was measured using automated wearable cameras. Participants were asked to wear a Brinno TLC130 camera (Brinno, 55mm x 55mm x 28.2mm, 74.5g, 138° field-of-view) on an adjustable chest harness during all waking hours across four randomly selected days (three weekdays, one weekend day). On school days, participants were asked to wear the camera during all waking hours before and after school. The camera was programmed using the Brinno phone application. The camera was set to timelapse mode and took images every 15-20 seconds (no video or audio). The camera timestamp setting was turned on. The camera data was collected using an encrypted SD card to protect participants' and third parties' privacy.

At the first visit, a research team member provided instructions on how to operate and charge the camera. The research team member demonstrated how to wear the camera on the chest harness and adjust the camera position. Further, information booklets were given with detailed instructions on how to operate and charge the camera and where the participants should remove or turn 'off' the camera (e.g., on school grounds or bathrooms). Participants then practised operating the camera until the research team member was confident that the participant could operate the camera. Parents were given information cards about the project to give to third parties (e.g., the public, family or friends) to explain the study if required. Parents were sent a text message each morning and afternoon on camera wear days to remind them to put the camera on their child and to charge the camera at night.

At the second visit, a research team member collected the camera data and ran the data through face-blurring software (Sanders, 2023). The face blurring software converts the timelapse videos to images and uses a machine learning model to detect and blur faces in the images. Participants and their parents were then given the opportunity to review the images and delete any images they did not want the researchers to view, per ethical guidelines (Kelly et al., 2013). The remaining images were stored on a secure network drive. Access to this drive was restricted to members of the research team. This was done in accordance with ethical guidelines to protect participant and third-party privacy (Kelly et al., 2013).

I calculated screen time measurements from the camera data for each participant by multiplying the number of images coded with the specific behaviour or content of interest by the camera epoch (i.e., 20-seconds). I then converted the result from seconds to minutes. For instance, if a screen exposure consisted of 30 images coded as programme viewing, the calculated screen time for programme viewing would be 10 minutes. I then adjusted camera data so that all participants had the same observed lengths, assuming that the observed time would represent their daily screen use behaviours. To do this, I scaled the data linearly such that all participants had 14 hours of observed camera data. I chose to do 14 hours assuming that the participants would sleep on average 10 hours per night (Xiao et al., 2020). I then calculated the participant's mean unadjusted and adjusted screen time for each screen use measurement. Additionally, I conducted a sensitivity analysis to address any discrepancies in the results based on the adjusted hour and found only a slight variance between the results. I have presented the results of this sensitivity analysis in Appendix N.

Child- and Proxy-Reported Screen Use

At the second visit, participants and their parents were asked to complete a short questionnaire separately. Parent- and child-reported screen use was measured using questions from the Screen Based Media Use Scale (SBMUS; Houghton et al., 2015). The SBMUS measures daily screen use and considers newer screen-based media devices and activities. The scale has been tested in children and adolescents aged 8-16 years old and has a test-retest reliability of 0.52 (Houghton et al., 2015).

Child-Reported Screen Use. To measure child-reported screen use, participants were asked to estimate their average daily screen use by type of device and activity. Participants were asked the following ten questions: (1) "On an average day in the past week, how much time do you play single-player video games on a computer, console, phone, or other device (Xbox, PlayStation, iPad, Apple TV)?" (2) "On an average day in the past week, how much time do you play multiplayer video games on a computer, console, phone, or other device (Xbox, PlayStation, iPad, Apple TV) where you can interact with others in the game?" (3)

"On an average day in the past week, how much time do you text on a mobile phone, tablet, iPod, or other electronic device(e.g., GChat, WhatsApp, Kik)?" (4) "On an average day in the past week, how much time do you visit social media apps (e.g., Snapchat, Facebook, Twitter, Instagram, TikTok)?" (5) "On an average day in the past week, how much time do you video chat (Skype, FaceTime, VRchat) that is NOT for school?" (6) "On an average day in the past week, how much time do you spend searching or browsing the internet (e.g., using Google) that is NOT for school?" (7) "On an average day in the past week, how much time do you watch "or stream" movies, videos, or TV shows? (Such as Hulu, Netflix, Amazon, YouTube, Twitch)" (8) "On an average day in the past week, how much time do you do schoolwork or homework on a computer, phone, tablet, or other electronic device?" (9) "On an average day in the past week, how much time do you video chat for school?" and (10) "On an average day in the past week, how much time do you search or browse the internet for school?" For each question, options included a dropdown list of time intervals starting from 0, 15 minutes, 30 minutes, 45 minutes, 1 hour, 1.5 hours, 2 hours, 2.5 hours, 3 hours and then increasing by hourly increments until 24 hours.

Parent-Reported Screen Use. To measure parent-reported screen use, parents were asked to estimate their average daily screen use by type of device and activity. Parents were asked the following four questions: (1) "Thinking of the same weekday last week (Monday to Friday), how many hours and minutes 'child name' spent browsing the Web?" (2) "Thinking of the same weekday last week (Monday to Friday), how many hours and minutes 'child name' spent on TV/Videos/Music?" (3) "Thinking of the same weekday last week (Monday to Friday how many hours and minutes 'child name' spent on social networking and instant messenger?" and (4) "Thinking of the same weekday last week (Monday to Friday), how many hours and minutes 'child name' spent gaming?". Parents used an interactive slide bar for each question that measured screen use in hours and minutes. Parents were asked to complete questions regarding demographic information, including child age, child sex, child school, spoken language at home, ethnicity, the highest level of education attained by the parent, parental occupation, and family factors.

To ensure participants remained blinded to the study purpose, participants and parents were also asked a range of questions related to family factors, quality of life, aggressive behaviour and conduct problems, temperament and persistence, physical activity, reading and homework (see Appendix L for the Child Questionnaire and Appendix M for the Parent Questionnaire). All questionnaire data was collected and stored using REDCap electronic data capture tools hosted at the Australia Catholic University (Harris et al., 2019, 2009).

To address my research question of whether automated wearable camera measurements were similar to self- and proxy-reported measurements of screen use, I converted all parent-reported and child-reported responses from hours to minutes to match the wearable camera data. Further, some questions in the child SBMUS asked children to estimate their screen time spent in specific activities (i.e., single-player video games vs multiplayer video games) or specific content (i.e., screen use related to school work vs non-school; Houghton et al., 2015). Questions in the parent SBMUS asked parents more broad questions (i.e., how many hours and minutes ‘child name’ spent gaming; Houghton et al., 2015) and did not differentiate between screen use related to school work or non-school work. To ensure the parent-reported and child-reported responses were comparable, I aggregated child-reported responses into broader content and activity groups to match the parent-reported questions. For instance, I calculated the child-reported gaming measurement by summing responses from questions one and two (i.e., how much time do you play single-player video games on a computer, console, phone, or other device, and how much time do you play multiplayer video games on a computer, console, phone, or other device). I calculated the child-reported educational screen time measurement by summing questionnaire responses that only measured school-related screen use (i.e., questions eight, nine, and ten). I calculated the child-reported recreational screen time measurement by summing questionnaire responses that only measured screen use that was NOT related to school screen use (i.e., questions one to seven). I calculated child-reported internet browsing by summing questionnaire responses related to school and non-school internet browsing (i.e., questions six and 10). I calculated child-reported screen use by social content by summing questions related to social media and communication application (i.e., questions three, four, and nine). I calculated child and parent-reported aggregated screen time by summing all responses, respectively.

Image Coding Protocol

Image coding was completed between November and December 2023. One coder coded the images using the open-source software Timelapse 2.0 Image Analyser software (v2.3.0.6; Greenberg, 2023). We used an object classification model to reduce the amount of data needing to be coded. The model was trained to determine if a screen-based media device was present in an image by identifying images containing electronic screens (with power on).

The images were coded using the coding protocol presented in Chapter 3. Coders coded each image for a screen exposure. Screen exposure was defined as an event or episode where a person is in the presence of one or more active screen-based media devices, regardless of whether or not the person is consciously attending to the device. Images that did not contain a screen device could still be coded as screen exposure if they occurred in a screen exposure episode and had an image before or after depicting a screen-based media device. For each identified screen exposure, coders were required to code the image for the number of devices, type of devices, content, social environment, location, and associated behaviours. However, only the categories of number of devices, type, and content are relevant to the present study. If the image was codable, coders continued coding the image. If the image was determined as uncodable (i.e., did not contain a screen exposure), the coders coded the images as 'Uncodable' and moved on to the next image. Coders were required to code the image to the lowest level possible. For example, if a coder was confident that the content depicted on the screen-based media device was interactive, but could not tell what interactive media category, they coded the content as 'unclassifiable – interactive'. Table 5.1 presents an overview of the codes used in this study.

Prior to coding the images, coders completed coder training (Sanders, n.d.). Coder training included ethics training, coding process familiarisation, timelapse software familiarisation, and image coding practice and evaluation. Coder's practised image coding on a model set of images. During this practice session coders could ask questions about the image coding process. After the set of images were coded, the coded data was compared to the coded data by another research team member.

Table 5.1

Summary of the Image Codes Included in this Study

Device	Content Type		Content Purpose
	Passive Screen Media	Interactive Screen Media	
Television	Unclassifiable – Passive	Unclassifiable – Interactive	Educational
Laptop Computer	Programme/TV/Video	General Internet	Recreational
Desktop Computer	Reading	Video Game	Social
Digital Signage		Creation	Other
Projector		Communication	Unclassifiable
Mobile Device (Phone, iPod)		Social Media	
Tablet (iPad, eReader)			
Handheld Game Console			
Smartwatch			
Not Elsewhere Classified			
Unclassifiable			

Statistical Analysis

I conducted all analyses using R (v4.3.1; R Core Team, 2023). For the camera measurements, I only included participants who had at least 4 hours of wear time for one day and wore the camera for at least two weekdays. As automated wearable camera research is a relatively new approach to measuring children's health behaviours there are no preexisting recommendations minimum wear time to be considered a valid day. Therefore, I based the wear time criteria on a calculation that participants would have a maximum of eight hours of wear time on a school day (two hours before and six hours after school). I excluded weekend camera data as this was not measured in the child-reported measurement instrument.

To measure the degree of relationship between methods, I calculated the Pearson correlation coefficient (*r*) use using the R-package *Hmisc* (v5.1.1; Harrell, 2023). I interpreted the Pearson correlation coefficient as values less than .1 indicate negligible correlation, values between .1 and .39 as weak correlation, values between .4 and .69 as moderation correlation, values between .7 and .89 as strong correlation, and values between .9 and 1 as very strong correlation (Schober et al., 2018). Further, I used Bland-Altman plots to assess the overall agreement between the methods and examine any potential bias between the methods. I created the Bland-Altman plot using the R-package *ggplot2* (Wickham, 2016).

Missing values in the child-reported and parent-reported measurements were ignored when calculating the mean values and correlations for each screen use behaviour. I chose this approach to prevent the distortion in the results as data imputation of missing values is inappropriate for method comparison studies. For all tests, the alpha level was set at .05.

Results

Participant Characteristics and Camera Data

From 27 participants, 122,231 images were captured across the four-day camera observation period. I excluded 75,156 images that were non-wear time data (e.g., the camera was accidentally left on overnight), 10,548 images that were captured on weekend days, and 8,842 images from 9 participants who did not meet the wear time criteria (i.e., worn for less than two days or had less than three hours per day). I then excluded one participant from child-reported measurement analyses due to reporting unrealistic screen time hours (i.e., > 36 hours of aggregated screen time in a single day). The final study sample consisted of 27,685 images from 17 participants.

The final study sample included participants aged 7 to 11 years old (M = 8.5, SD = 1.3). Over half (53%) of the participants were male (Table 5.2). The average number of days the children wore the camera was 2.6 days (SD = 0.6), and the average daily wear time was 11.2 hours (SD =1.8, ranged between 7.7 to 12.8 hours). All parents who completed the parent-reported questionnaire identified as the study child's biological mother.

Table 5.2

Participant Characteristics and Descriptive Data

Variable	*n*	%	*M*	*SD*
Sex				
Males	9	53		
Females	8	47		
Indigenous				
Non-Indigenous	17	100		
Aboriginal or Torres Strait Islander	0	0		
Spoken language at home				
English	12	71		
Other Language	5	29		
Parent's highest completed level in high school [a]				
Year 12 or equivalent	17	100		
Year 11 or below	0	0		
Did not attend	0	0		

Variable	*n*	%	*M*	*SD*
Parent's highest qualification completed [a]				
Postgraduate degree	6	35		
Graduate diploma/Graduate certificate	2	12		
Bachelor's degree (with or without honours)	5	29		
Advanced diploma	2	12		
Certificate 3 or 4 (including trade certificate)	0	0		
Certificate 1 or 2	1	6		
Other non-school qualification	0	0		
Have not completed educational qualification	1	6		
Camera wear time per day (hr)			11.2	1.8
Number of days camera was worn			2.6	0.6

[a] Only includes the participating parent.

To address my research question of whether estimates of screen use from automated wearable cameras (i.e., device-based) measurement are similar to self- and proxy-reported measurements of screen use, I have reported the results of aggregated screen time measurements, screen use by content measurements and screen use by activities measurements separately. Further, I found strong correlations between adjusted and unadjusted camera measurements for all screen use measurements. Therefore, I have only included the Bland-Altman plots using adjusted camera measurements throughout this chapter. The Bland-Altman plots using the unadjusted camera measurements are in Appendix O.

Aggregated Screen Time

I have presented the descriptive statistics and Pearson correlation coefficients for aggregated screen time for unadjusted and adjusted camera measurements and parent- and child-reported measurements in Table 5.3. I found a weak correlation between parent-reported and camera unadjusted and adjusted camera measurements of aggregated daily screen time; however, the results were not significant. Parents' estimates of aggregated screen time were typically higher than the adjusted camera measurements of aggregated screen time (*M* = 89.37 minutes, *SD* = 182.39; shown in Figure 5.1).

I found a weak correlation between child-reported and camera unadjusted camera measurements of aggregated daily screen time and a weak negative correlation between child-reported and adjusted camera measurements of aggregated daily screen time; however, neither result was significant. Children's estimates of aggregated screen time were typically higher than the adjusted camera measurements of aggregated screen time (*M* = 191.2 minutes, *SD* = 185.5).

I found a moderate correlation between child-reported and parent-reported screen time measurements of aggregated daily screen time. Children's estimates of aggregated screen time were typically higher than parent's estimates of aggregated screen time (*M* = 137.4 minutes, *SD* = 141.75).

Table 5.3

Descriptive Statistics and Pearson Correlations for Aggregated Screen Time for Unadjusted and Adjusted Camera, Parent-Report and Child-Report Measurements

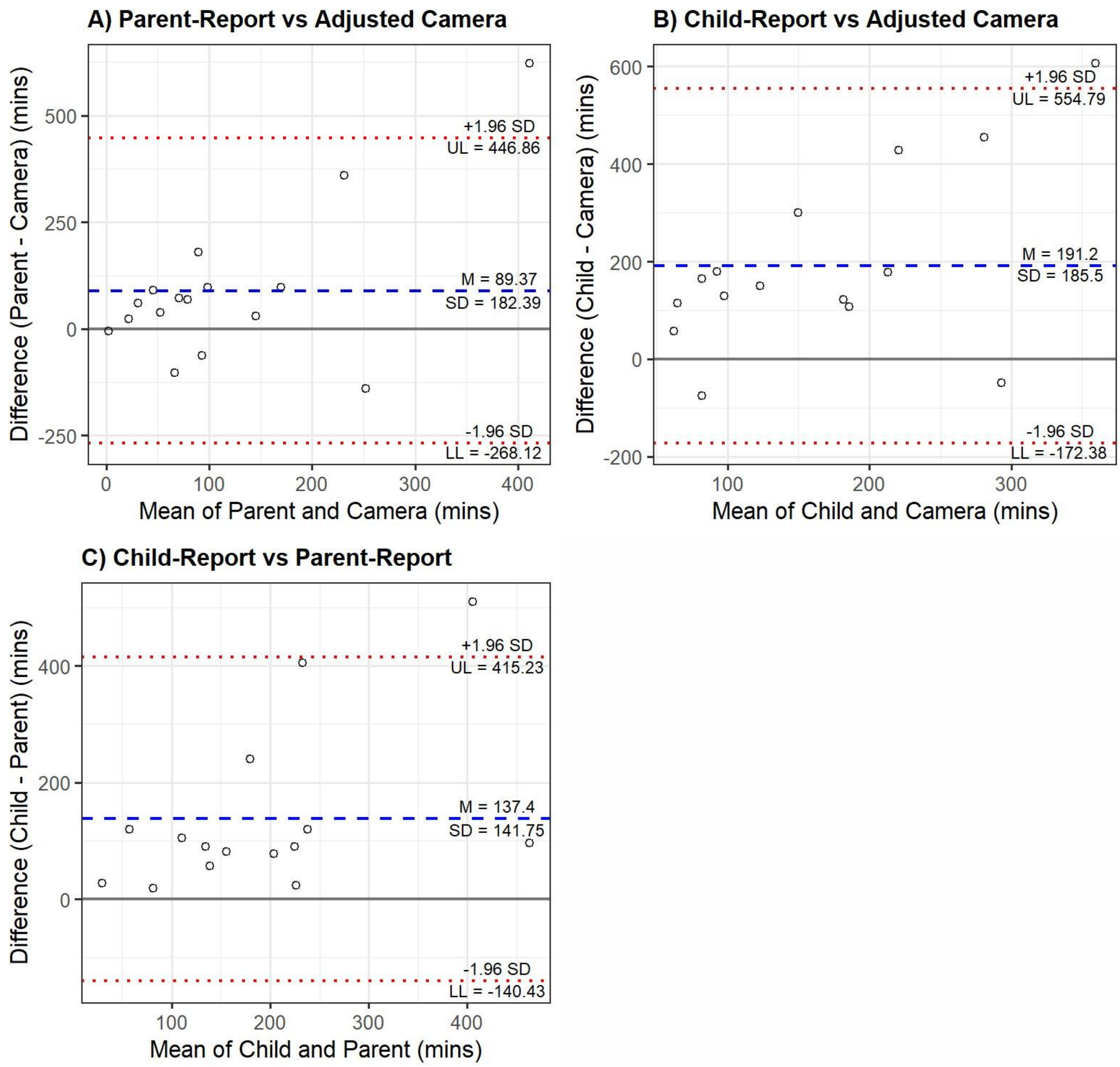
Variable	*n*	*M*	*SD*	1	2	3	4
1. Unadjusted camera screen time	15	52.6	49.1	—			
2. Adjusted camera screen time [a]	15	75	80.8	$.96^{***}$	—		
3. Parent-report screen time	15	160.9	179.3	.22	.22	—	
4. Child-report screen time	15	261	164.9	0.05	-.00	$.52^{*}$	—

[a] Adjusted camera screen time scaled based on 14-hours of camera wear time per participant

$^{*}p < .05$. $^{***}p < .001$.

Figure 5.1

Bland-Altman plot of Aggregated Screen Time Adjusted Camera, Parent-Report and Child-Report Measurement Differences



Note. Adjusted camera screen time scaled based on 14-hours of camera wear time per participant. Plot A has 16 paired observations. Plot B has 16 paired observations. Plot C has 15 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

Screen Use by Content

Screen use by content was separated into three categories: Recreational, educational and social. Recreational and educational screen use was only measured in the child-reported and camera measurements. Social screen use was measured in the child-reported, parent-reported and camera measurements. I have presented the Pearson correlation coefficients for recreational and educational screen use for unadjusted and adjusted camera and parent-reported measurements, and social screen use for unadjusted and adjusted camera, parent-, and child-reported measurements in Table 5.4.

Recreational

I found a weak correlation between child-reported measurements of recreational screen time and unadjusted and adjusted camera measurements of recreational screen time, but the result was not significant. Children's estimates of recreational screen time were typically higher than adjusted camera measurements of recreational screen time. I found a large systematic bias of 146.73 minutes (*SD* = 121.74) and a wide limit of agreement for child-reported measurements of recreational screen time compared to adjusted camera measurements of recreational screen time.

Educational

I found a weak, negative correlation between child-reported measurements of educational screen time and unadjusted and adjusted camera measurements of educational screen time; however, the results were not significant. As shown in Figure 5.2, children's estimates of educational screen time were typically higher than the adjusted camera measurements of educational screen time. I found a systematic bias of 62.5 minutes (*SD* = 107.14) and a wide limit of agreement for child-reported measurements of educational screen time compared to the adjusted camera measurements of educational screen time.

Table 5.4

Pearson Correlations for Screen Use by Content for Unadjusted and Adjusted Camera, Parent-Report and Child-Report Measurements

Variable	1	2	3	4	5	6	7	8	9	10
Recreational screen use										
1. Unadjusted camera recreational content	—									
2. Adjusted camera recreational content [a]	.97***	—								
3. Child-reported recreational content	.34	.32	—							
Educational screen use										
4. Unadjusted camera educational content	-.08	-.07	.09	—						
5. Adjusted camera educational content [a]	-.08	-.07	.09	1***	—					
6. Child-reported educational content	-.19	-.21	.05	-.02	-.02	—				
Social screen use										
7. Unadjusted camera social content	.35	.18	.05	.09	.09	.04	—			
8. Adjusted camera social content [a]	.35	.18	.05	.15	.15	.03	1***	—		
9. Parent-reported social content	.21	.23	.11	.92***	.92***	-.19	.21	.26	—	
10. Child-reported social content	-.22	-.23	.1	.67**	.67**	-.04	.13	.17	.66**	—

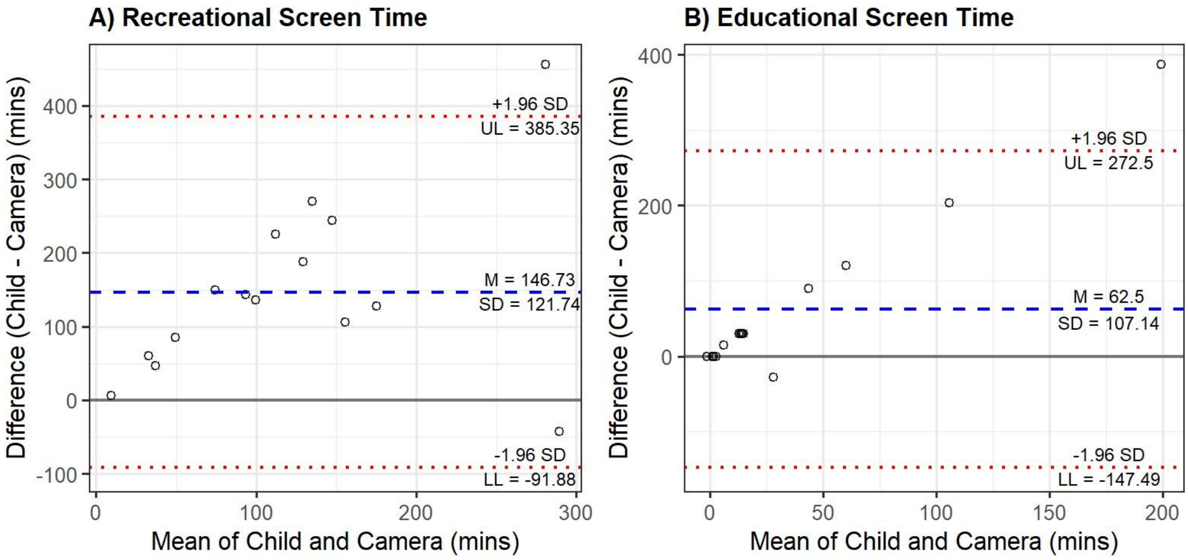
Note. N = 16 participants

[a] Adjusted camera screen time scaled based on 14-hours of camera wear time per participant

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Figure 5.2

Bland-Altman plot of Recreational and Educational Screen Use Adjusted Camera and Child-Report Measurement Differences



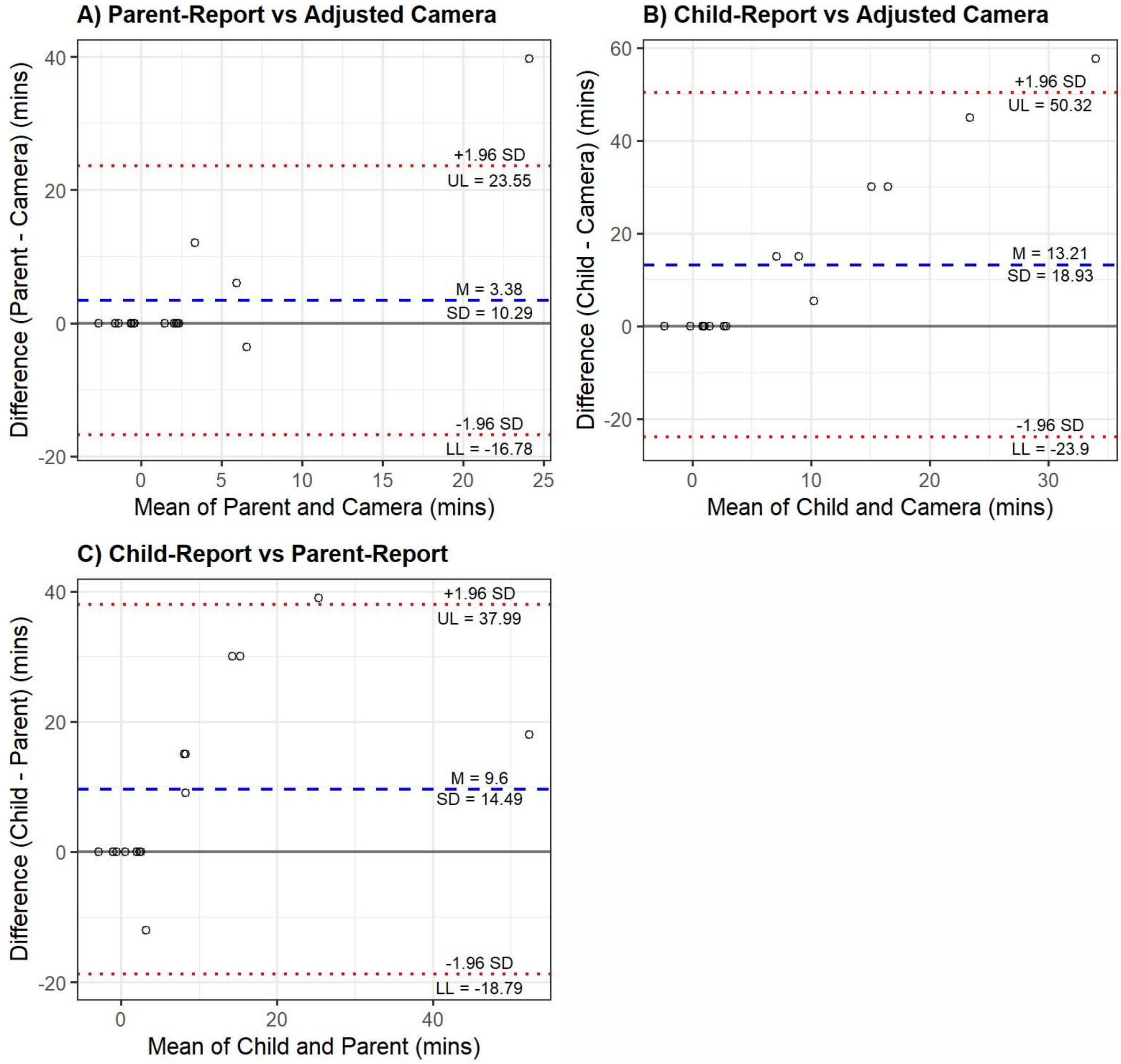
Note. Adjusted camera screen time scaled based on 14-hours of camera wear time per participant. Plot A has 16 paired observations. Plot B has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

Social

I found a weak correlation between parent-reported measurements of social screen time and unadjusted and adjusted camera measurements of social screen time, but the results were not significant. As shown in Figure 5.3, I found a small systematic bias of 3.38 minutes (SD =10.29), with an upper limit of agreement of 23.55 minutes (+1.96 SD) and lower limit of agreement of -16.78 minutes (-1.96 SD) between parent-reported and camera measurements of social screen time. Further, I found a very strong correlation between parent-reported social screen time and unadjusted and adjusted camera educational screen use.

Figure 5.3

Bland-Altman plot of Social Screen Use Adjusted Camera, Parent-Report and Child-Report Measurement Differences



Note. Adjusted camera screen time scaled based on 14-hours of camera wear time per participant. Plot A has 17 paired observations. Plot B has 16 paired observations. Plot C has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

I found a weak correlation between child-reported measurements of social screen time unadjusted and adjusted camera measurements of social screen time; however, the results were not significant. Children's estimates of social screen time were typically higher than the adjusted camera measurements of social screen time (M = 13.21 minutes, SD = 18.93). Further, I found a moderate correlation between child-reported measurements of social screen time and unadjusted and adjusted camera measurements of educational screen time.

I found a moderate correlation between child-reported and parent-reported measurements of social screen use. Children's estimates of social screen time were typically higher than the parents' estimates (M = 9.6 minutes, SD = 14.49).

Screen Use by Activities

Screen time activities were separated into four categories: Gaming, internet browsing, programme viewing and communication. Gaming, internet browsing, and programme viewing was measured in the child-reported, parent-reported and camera measurements. Communication-related screen time was only measured in the child-reported and camera measurements. I have presented the Pearson correlation coefficients for all screen time activity measurements between methods in Table 5.5.

Table 5.5

Pearson Correlations for Screen Use by Activities for Unadjusted and Adjusted Camera, Parent-Report and Child-Report Measurements

Variables	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Gaming															
1. Unadjusted camera gaming	—														
2. Adjusted camera gaming	1***	—													
3. Parent-reported gaming	.18	.2	—												
4. Child-reported gaming	.04	.06	.61*	—											
Internet browsing															
5. Unadjusted camera internet browsing	.12	.11	.09	.1	—										
6. Adjusted camera internet browsing	.12	.11	.1	.11	1***	—									
7. Parent-reported internet browsing	-.09	-.09	.01	-.04	.14	.14	—								
8. Child-reported internet browsing	.79***	.79***	-.1	-.15	.12	.12	-.12	—							
Programme viewing															
9. Unadjusted camera programme viewing	-.03	.01	.45	.54*	.16	.16	.26	-.19	—						
10. Adjusted camera programme viewing	-.06	-.02	.49	.51	.06	.06	.28	-.18	.97***	—					
11. Parent-reported programme viewing	-.19	-.2	.31	.83***	-.08	-.07	-.19	-.25	.16	.13	—				
12. Child-reported programme viewing	.06	.06	.15	.61*	-.25	-.24	-.26	.08	.15	.13	.8***	—			
Communication															
13. Unadjusted camera communication	-.08	-.08	.08	-.02	.89***	.89***	.16	.11	-.02	-.03	-.1	-.3	—		
14. Adjusted camera communication	-.08	-.08	.08	-.02	.9***	.89***	.16	.11	-.02	-.03	-.1	-.3	1***	—	
15. Child-reported communication	.29	.28	-.04	.07	.5	.5	-.31	.51*	-.34	-.33	.11	.18	.58*	.58*	—

Note. N = 16 participants

[a] Adjusted camera screen time scaled based on 14-hours of camera wear time per participant

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$

Gaming

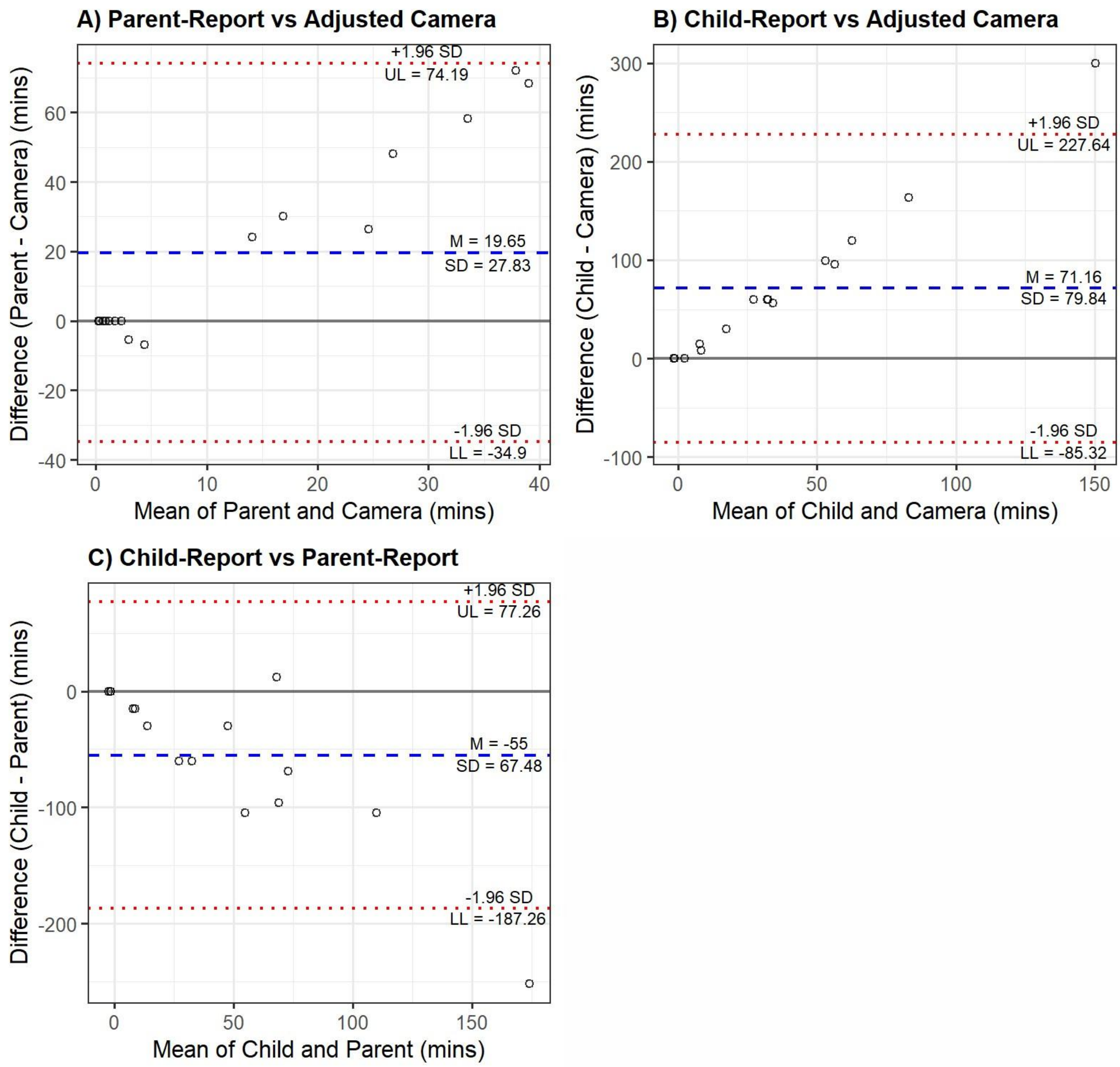
For parent-reported measurements of gaming, I found a weak correlation between unadjusted and adjusted camera measurements of gaming; however, the results were not significant. As shown in Figure 5.4, parent's estimates of gaming were typically higher than the adjusted camera measurements of gaming (M = 19.65 minutes, SD = 27.83).

I found a weak correlation between child-reported measurements of gaming and unadjusted and adjusted camera measurements of gaming; however, the results were not significant. Children's estimates of gaming were typically higher than the adjusted camera measurements of gaming (M = 71.16 minutes, SD = 79.85). Further, I found a moderate correlation between unadjusted camera programme viewing and child-reported measurements of gaming, a strong correlation with parent-reported programme viewing and child-reported measurements of gaming, and a moderate correlation with child-reported programme viewing and child-reported measurements of gaming.

I found a moderate correlation between child-reported and parent-reported measurements of gaming. Children's estimates were typically lower than parents' estimates (M = -55 minutes, SD = 67.48).

Figure 5.4

Bland-Altman plot of Gaming Adjusted Camera, Parent-Report and Child-Report Measurement Differences



Note. Adjusted camera screen time scaled based on 14-hours of camera wear time per participant. Plot A has 16 paired observations. Plot B has 16 paired observations. Plot C has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

Internet Browsing

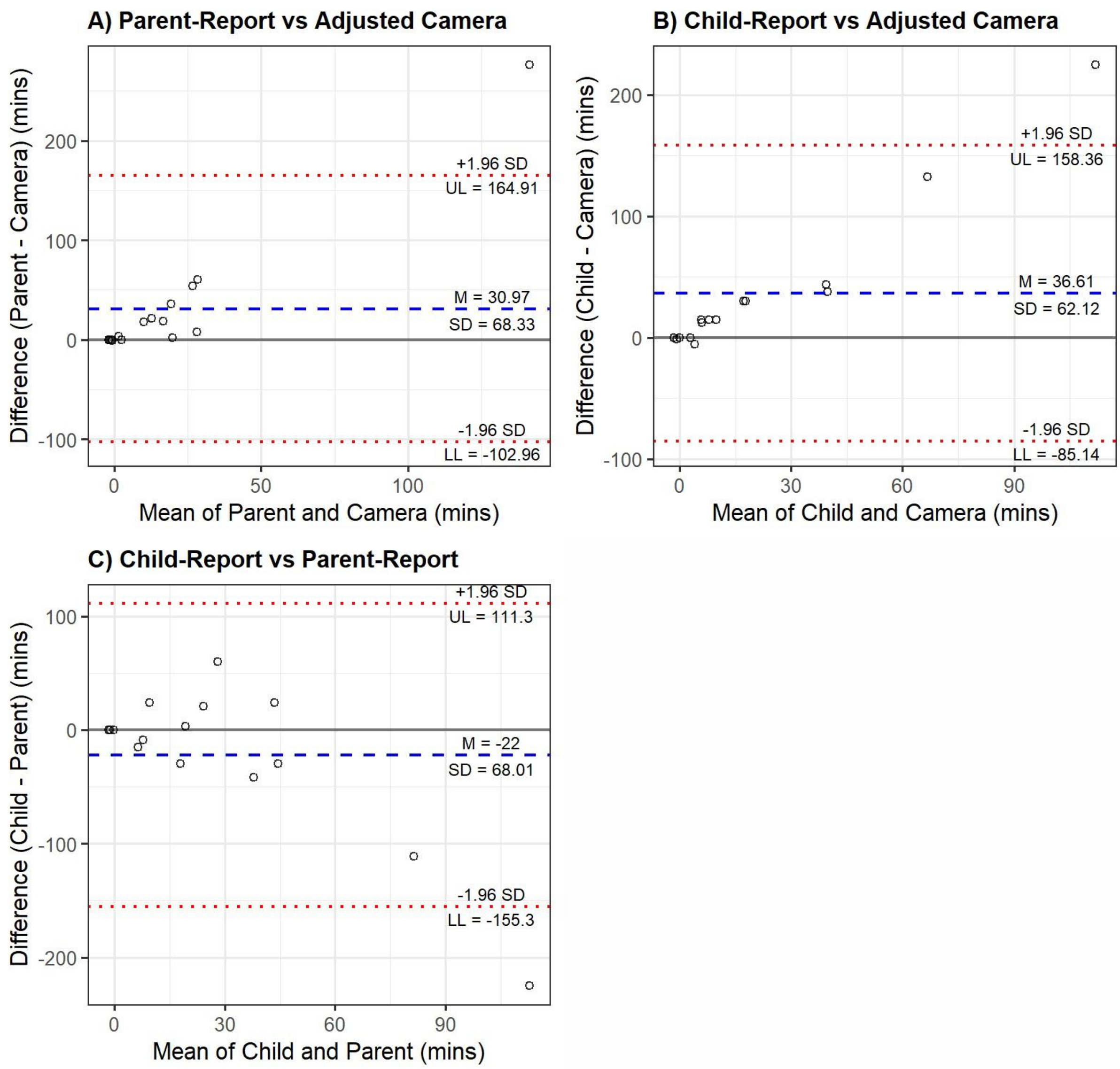
For parent-reported measurements of internet browsing, I found a weak correlation between unadjusted and adjusted camera measurements, but the results were not significant. Parent's estimates of internet browsing were typically higher than the adjusted camera measurements (M = 30.97 minutes, SD = 68.33).

For child-reported measurements of internet browsing, I found a weak correlation between unadjusted camera measurements and a weak correlation between adjusted camera measurements; however, the results were not significant. As shown in Figure 5.5, children's estimates of internet browsing were typically higher than the adjusted camera measurements (M = 36.61 minutes, SD = 62.12).

I found a weak negative correlation between child- and parent-reported internet browsing; however, the result was not significant. Children's estimates of internet browsing were typically lower than parent's estimates (M = -22, SD = 68.01).

Figure 5.5

Bland-Altman plot of Internet Browsing Adjusted Camera, Parent-Report and Child-Report Measurement Differences



Note. Adjusted camera screen time scaled based on 14-hours of camera wear time per participant. Plot A has 17 paired observations. Plot B has 16 paired observations. Plot C has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

Programme Viewing

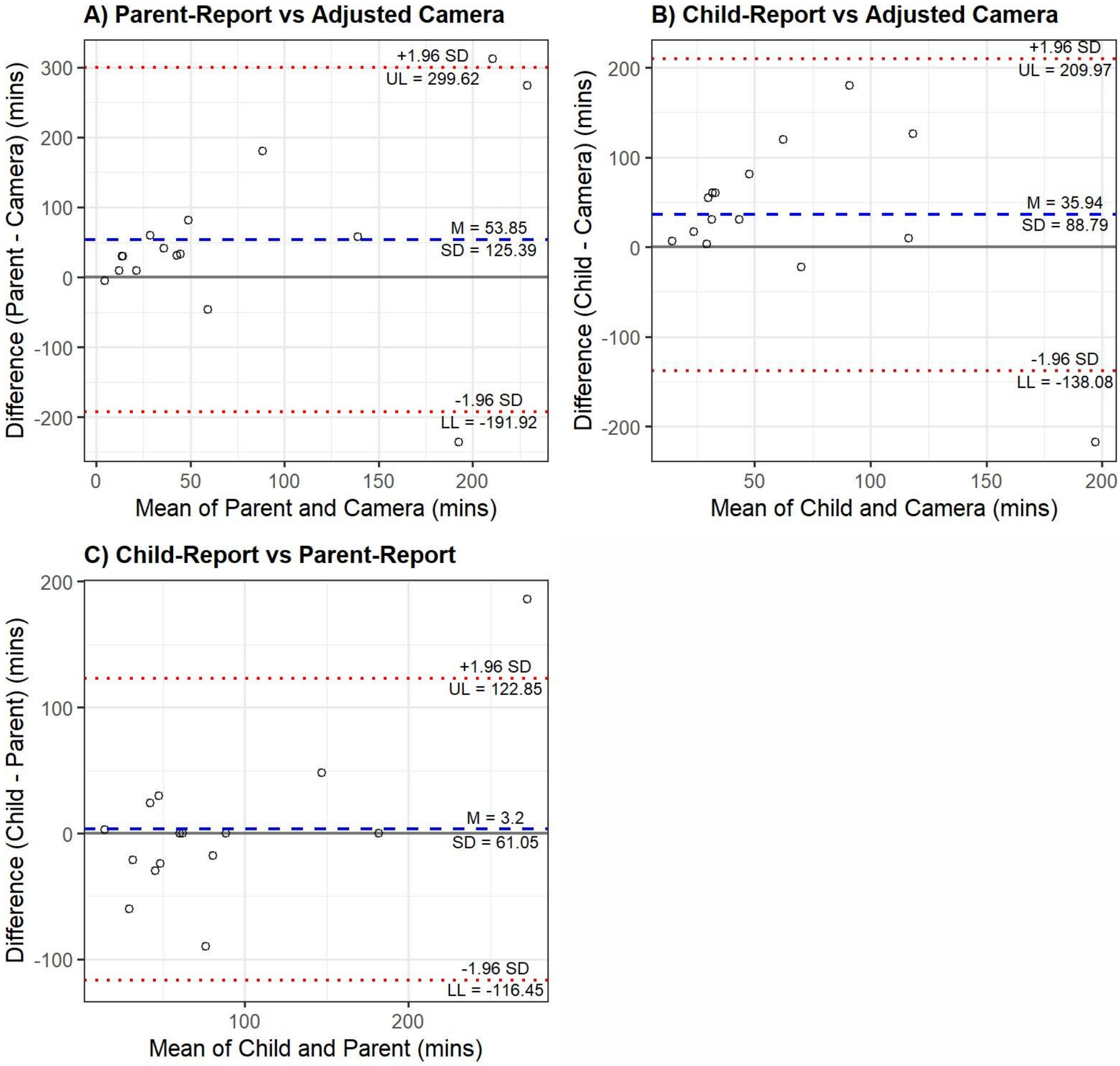
For parent-reported measurements of programme viewing, I found a weak correlation between unadjusted and adjusted camera measurements of programme viewing, but the results were not significant. Parents' estimates of programme viewing were typically lower than the adjusted camera measurements of programme viewing. I found a systematic bias of 53.85 minutes (SD = 125.39) and a wide limit of agreement (the upper limit of agreement was 299.62 minutes (+1.96 SD) and the lower limit of agreement was -191.92 minutes (-1.96 SD).

I found a weak correlation between child-reported measurements of programme viewing and unadjusted and adjusted camera measurements, but the results were not significant. As shown in Figure 5.6, children's estimates of programme viewing were typically lower than the adjusted camera measurements (M = 35.94 minutes, SD = 88.79).

I found a strong correlation between child- and parent-reported programme viewing. I found a systematic bias of 3.2 minutes (SD = 61.05) and a wide limit of agreement for child-reported measurements of programme viewing compared to parent-reported measurements (the upper limit of agreement was 122.85 minutes (+1.96 SD) and the lower limit of agreement was -116.45 minutes (-1.96 SD).

Figure 5.6

Bland-Altman plot of Programme Viewing Adjusted Camera, Parent-Report and Child-Report Measurement Differences



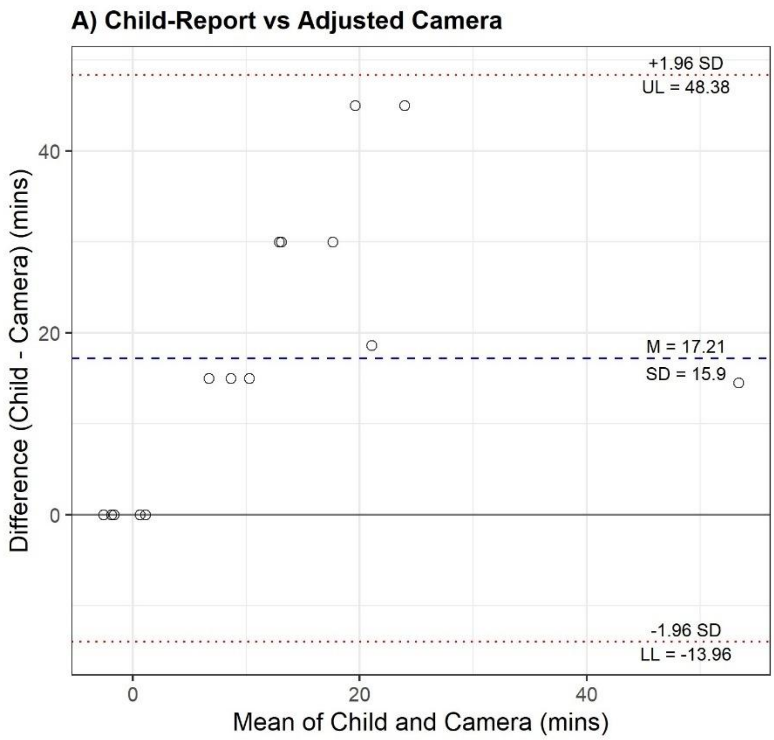
Note. Adjusted camera screen time scaled based on 14-hours of camera wear time per participant. Plot A has 17 paired observations. Plot B has 16 paired observations. Plot C has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

Communication

I found a moderate correlation between children-reported communication and unadjusted and adjusted camera communication measurements. Further, child-reported communication was moderately correlated with chid-reported internet browsing. Children's estimates of communication were typically lower than adjusted camera communication measurements. As shown in Figure 5.7, I found a systematic bias of 17.21 (SD = 15.9) between child-reported and adjusted camera measurements of communication.

Figure 5.7

Bland-Altman plot of Communication Adjusted Camera, Parent-Report and Child-Report Measurement Differences



Note. Adjusted camera screen time scaled based on 14-hours of camera wear time per participant. Plot has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement

Discussion

This study aimed to examine whether estimates of screen use from automated wearable cameras were similar to self- and parent-report measurements. To do this, I used baseline data collected as part of the KidVision project to compare children's screen use as measured by automated wearable cameras, a parent-reported questionnaire and a child-reported questionnaire.

Consistent with my hypothesis, I found that overall, there was a weak correlation between automated wearable camera measurements and child- and parent-reported measurements for aggregated screen time and screen use content (i.e., educational screen use, recreational screen use and social screen use). However, these results were not statistically significant. These findings indicate limited evidence of a linear relationship between the parent- and child-reported measurements and camera measurements for aggregated screen time, educational screen use, recreational screen use, and social screen use. This aligns with evidence that suggests self- and proxy-reported measurement instruments are susceptible to various forms of bias and inaccuracies due to recall bias, misclassification bias, and social desirability bias (Hardy et al., 2013). The low correlations found in this study are also consistent with the notion that self-reported data often underestimates or overestimates actual behaviour, especially in contexts where the measured behaviours are multifaceted or sporadic (Ainsworth et al., 2012). Moreover, these findings align with a systematic review that found self-reported measurements typically have lower correlations to direct observation than device-based measurements (i.e., electronic TV monitors and smartphone applications) or media diaries (Perez et al., 2023). However, these findings are in contrast with some studies investigating the accuracy of self-reported mobile device and social screen use. A study by Wade et al. (2021), found a low correlation between parent-reported screen time on mobile devices and a phone tracking application in preteens aged 11 to 12 years, but found a moderate correlation between child-reported screen time on mobile devices and a phone tracking application. Further, my findings for social screen time are in contrast with a study examining the criterion validity of self-reported measurements of Facebook use against a computer monitor software among adults, which found a moderate correlation between self-report time spent on Facebook and computer monitoring software, and a strong correlation between self-report time spent on Twitter and computer monitoring software, $r = .59$ ($p < .001$) and $r = .87$ ($p < .001$), respectively (Junco, 2013). However, these studies were limited to single device and platform measurements, which some researchers suggest may not accurately reflect estimates for broader screen use (Mahalingham et al., 2023).

I hypothesised that there would be a weak correlation between automated wearable camera measurements of screen use and child- and proxy-reported measurements for screen use activities, including gaming, internet browsing, programme viewing and communication (i.e., messaging). I found a weak correlation between automated wearable camera measurements and child- and parent-reported measurements for gaming, internet browsing and programme viewing; however, these results were not significant. These results indicate limited evidence of a linear relationship between the parent- and child-reported measurements of gaming, internet browsing and programme viewing and camera measurements. These findings are consistent with previous studies on the accuracy of proxy- and self-reported measurements of screen use that show that self-reported measurements of gaming, internet use and programme viewing have low correlations with device-based measurements such as phone tracking applications and television monitors (Borzekowski & Robinson, 1999; Perez et al., 2023). For instance, a study investigating internet use among 690 adults found low correlations ($r = .29$, $p < 0.01$) between self-reported internet use on an average day and internet use tracking software (Araujo et al., 2017). These findings emphasise the potential biases in self-reported data, particularly in contexts where health behaviours may be under- or over-reported due to classification or social desirability bias (Hardy et al., 2013). However, my findings contrast with a study that found a moderate correlation between self-report time spent on games and phone tracking applications in young adults (Lee et al., 2021). Another study found a moderate correlation between self-reported television viewing and a Television monitor in adults (Otten et al., 2010). This inconsistency may be attributed to the differences in the age groups studies, as the accuracy of self-reported data may vary across different developments stages (Sallis & Saelens, 2000). Moreover, I found a moderate correlation between child-reported and camera measurements of communication. This finding does not align with previous studies showing that adolescent and adult self-reported measurements of communication (i.e., messaging) correlate poorly with phone tracking applications (Lee et al., 2021). One reason to explain my inconsistent findings is that it may be due to different screen use habits amongst different age groups. For instance, previous studies often only investigate the accuracy of self-reported measurements in adolescent and adult populations rather than child populations (Perez et al., 2023). Children may spend more time engaging in different screen use behaviours than adolescents and adults. For instance, children may be more likely to engage in more recreational and less communication screen use than adolescents or adults, making it easier for parents and children to recall different behaviours (Schwarz & Oyserman, 2001).

In addition, this study also explored the correlation between child- and parent-reported measurements of screen use. I found a moderate correlation between child- and parent-reported measurements for aggregated screen time, social screen use, and gaming. I found a strong correlation between child- and parent-reported measurements of programme viewing. However, I found a weak negative correlation between child- and parent-reported internet browsing, but the result was not significant. These findings suggest that the child- and parent-reported measurements may be linearly associated for aggregated screen time, social screen use, gaming, and programme viewing, but not internet browsing. This is in contrast with a study of 1,124 adolescents and 1,002 parents found that parent-reported internet use was significantly correlated with adolescent self-reported internet use (Liau et al., 2008). This discrepancy may be due to differences in the age groups studied, as well as variations in how internet browsing was reported by parents and children. Moreover, a study investigating the concordance of parent and child reports of children's screen use among children aged 9 to 11 years (*N* = 14) found that child and parent reports had similar time estimates for gaming, computer use and television viewing, which were generally within 10 minutes of each other (Wood et al., 2019).

Children and parents were more likely to overestimate than underestimate children's screen use for all screen use measurements compared to the camera measurements. The findings from the Bland-Altman Plot analyses indicated that parent estimates were more similar to camera measurements when estimating aggregated screen time (parent over-reported by 1.5 hours compared to child over-reported by 3.2 hours), social screen use (parent over-reported by three minutes compared to child over-reported by 13 minutes), gaming (parent over-reported by 20 minutes compared to child over-reported by 1.2 hours) and browsing (parent over-reported by 31 minutes compared to child over-reported by 37 minutes). Child estimates were more similar to camera measurements when estimating programme viewing (child over-reported by 36 minutes compared to parent over-reported by 54 minutes). This contrasts with a previous study examining of children's television viewing measurements, which found that mother's general estimates were more accurate than children's (Borzekowski & Robinson, 1999). Moreover, children were more accurate at estimating their social screen use than recreational or educational screen use. Compared to the camera measurements, children over-reported their social screen use by 13 minutes, their educational screen use by 63 minutes, and their recreational screen use by 2.4 hours. Overall, these findings are consistent with previous research that has found that low levels of measurement often correspond with tendencies of over-reporting for internet use, social media use and mobile phone use (Araujo et al., 2017; Junco, 2013; Vanden Abeele et al., 2013).

The accuracy of parent- and child-reported measurements may vary between individuals. While my findings indicate that parent- and child-reported measurements may not be accurate when measuring screen use compared to a device-based measurement, it is important to note the high level of variance shown in the Bland-Altman plots. The wide limits of agreement indicate that the extent of over-reporting may not be consistent between individuals. This is consistent with a systematic review, which found high levels of heterogeneity in effect sizes for correlations and ratio of means between device-based and self-reported screen use (Parry et al., 2021). Some studies have suggested that the use of self-report may result in attenuated effect sizes, which has resulted in a reduced magnitude of observed effects (Jones-Jang et al., 2020). Others have suggested that the inaccuracy of self-reported screen use measurements may be due to systematic error (Parry et al., 2021). The inaccuracy of self-reported screen use may also be attributed to biases, such as social desirability or recall bias (Parry et al., 2021). Screen use behaviours are susceptible to biases due to their frequent and habitual nature (Parry et al., 2021). For example, respondents may underreport screen time due to negative social connotations associated with excess screen use.

Moreover, previous research has found that the accuracy of self-reported screen use may depend on how much the respondent uses the screen media or device. For instance, a study found that higher levels of internet use were associated with higher levels of under-reporting and lower levels of internet use were associated with higher levels of over-reporting (Araujo et al., 2017). Another study found that the extent of inaccuracy between self-reported mobile phone use estimates and Apple's Screen Time application tracked screen use was associated with respondent's level of well-being and amount of use (Sewall et al., 2020). Further investigation is needed to examine unexplained variance in child- and parent-reported screen use measurements.

Limitations and Future Research

This was the first study to use automated wearable cameras to assess the similarity of parent- and child-reported screen use measurements. Automated wearable cameras present an alternative approach that may overcome some of the challenges of other device-based measurements of screen use. That said, the automated wearable camera measurements may not be free of measurement error. The automated wearable camera measurements may have introduced a measurement error due to the camera epoch. Although I determined in Chapter 4 that a camera epoch of 15 to 20 seconds overestimated children's screen use by -5.15 (SD =7.01) compared to direct observation, this may have introduced a small measurement error into the study. Additionally, the automated wearable camera measurements in the free-living setting may have missed some screen exposures. This may be due to non-wear time (i.e., the participant forgot to put on the camera) or the Hawthorne effect (i.e., participants changed their habits when wearing the camera). While my study in Chapter 4 showed that automated wearable camera measurement can accurately measure screen use while wearing the camera, it does not account for potential missed screen exposures due to non-wear time or behaviour change while wearing the camera. Automated wearable camera research is a relatively new approach to measuring children's health behaviours. As a result, there are no preexisting recommendations for minimum wear time needed to be considered a valid day. Therefore, I based the wear time criteria on a calculation that participants would have a maximum of eight hours of wear time on a school day (two hours before and six hours after school). Further, like many device-based measurements, participants' camera wear time differed by the number of days worn and hours worn per day. To ensure that the screen use represented an average day across all participants, I adjusted camera data so that all participants had the same observed lengths assuming that the observed time would represent their daily screen use behaviours. To address any discrepancies in the results based on the adjusted hours, I conducted a sensitivity analysis and found only slight variances between the results (Appendix N).

Moreover, the self-reported measurements also had several limitations. The child- and parent-reported measurements primarily focused on the amount of screen time and type of activity. Thus, the self-reported measurements were limited to broad categories of screen use without accounting for contextual factors such as the specific content and variations in screen use patterns throughout the day. Additionally, to ensure comparability with the automated wearable camera data, I converted all child-reported responses from hours to minutes. Child-reported questionnaire include questions regarding specific screen activities and content types (i.e., single-play and multiplayer video games, and school-related screen use and non-school related screen use), whereas the parent-reported questionnaire were limited to broad categories that did not differentiate between specific screen activities. To reconcile these differences, I aggregated the child-reported responses into a broader content and screen activity groups to match the parent-reported responses. For instance, time spent gaming was calculated by summing child-reported responses related to single-player and multiplayer video games. Similarly, child-reported response for educational and recreational screen time were aggregated based on specific questions related to school-related and non-school related screen use, respectively. In doing so, this may have contributed to under-estimating or over-estimating the child-reported responses.

Another limitation of this study was that the findings were based on a small sample (17 participants). In this study, the sample size of participants was determined by the amount of available data collected at baseline from the KidVision project. At the time of writing, the KidVision project was in the process of collecting baseline data. Thus, at the time of this study data was available for only 27 participants. A power analysis indicated that a sample size of 84 participants would be required to detect a low correlation of .3 with 80% power (Arifin, 2023). This small sample size may have influenced the variation in the estimates due to sampling error and may limit the detection of significant correlations between the measurement instruments (Hopkins et al., 2009). Consequently, while the results of this study provide preliminary insights, they should be interpreted with caution. I recommend that future research increase the sample size to improve statistical power. Larger sample sizes will help detect more accurate correlations and provide more robust findings.

Although the goal was to examine the similarity between estimates of a device-based measurement and parent- and child-reported measurements, the generalisability of the findings is limited to the children aged 7 to 11 years old. Additionally, the sample consisted predominately of children from higher socioeconomic status backgrounds, as parents were mostly mothers who had completed higher education (88%). Families who have a higher socioeconomic status may be more likely to have access to screen-based medica devices and may not reflect the experiences of families from diverse socio-economic backgrounds. Therefore, the results should be interpreted with caution when generalising to the wider population. As such, future research should examine the similarity between device-based and child- and parent-reported measurements of screen use in different demographics.

Implications in Research and Policy

Currently, much of the research on children's screen use relies on self- or proxy-reported measurements (Ainsworth et al., 2012; Clark et al., 2011). The findings from this current study suggest that both parent- and child-reported measurements may be inaccurate when compared to automated wearable cameras. Specifically, the study found that parent-and child-reported measurements of screen use tend to overestimate children's screen use in comparison to automated wearable camera measurements. These findings align with previous research indicating that self-reported data often show low correlations with device-based measurements (i.e., direct observation, television monitors, and smartphone applications). The study's findings raise concerns reliability and validity of child- and parent-reported measurements, which have frequently informed our understanding of the impact of screen use on children's health and behaviour. Given that much of the existing literature and policy recommendations are based on self-reported data, the findings of this study suggest that research and policy conclusions drawn from such data should be approached with caution. By highlighting the limitations of self-reported measurements, this study demonstrates the need for more accurate and objective measurement tools, to better inform health policies and interventions aimed at managing children's screen use. Implementing more reliable measurement methods could lead to more informed and evidence-based interventions that will improve the effectiveness of public health policies to improve children's health and well-being.

Conclusion

The findings from this study suggest that parent- and child-reported measurements of screen use tend to be inaccurate compared to automated wearable camera measurements. My findings indicated a weak, non-significant correlation between nearly all screen use measurements (i.e., aggregated screen time, social, recreation and educational screen use, programme viewing, internet browsing, and gaming). Only child-reported and camera measurements of communication were moderately correlated. The findings indicated that children and parents were more likely to overestimate than underestimate children's screen use for all screen use measurements. Moreover, the findings suggest that the extent of over-reporting may not be consistent between individuals or screen use activities. As such, future research should examine potential causes of the systematic discrepancies in child- and parent-reported screen use measurements. Finally, the findings of this study raise concerns about the accuracy of child- and parent-reported measurements of screen use. Given the widespread use of child- and parent-reported measurements of screen use in the academic literature, such research should be interpreted with caution.

Chapter 6: General Discussion and Conclusion

Introduction

A crucial step in improving our understanding of the impact of screen use on children's health and behaviour outcomes is accurately measuring the duration, context and content of children's screen use (Kaye et al., 2020; Odgers & Jensen, 2020). The primary aim of this thesis was to examine the use of automated wearable cameras to establish a more accurate measure of children's screen use. The secondary aim was to examine whether estimates of screen use from automated wearable camera measurements were similar to self-report measurements.

In order to accomplish my primary aim, the first study in this thesis (Chapter 2) was a systematic review that included a narrative synthesis to identify health behaviours automated wearable cameras have measured, examine the evidence on the validity and reliability of automated wearable camera measurements, and describe the common methodological procedures. In this review, I synthesized data from 23 studies and found evidence that suggested automated wearable cameras may provide a reliable method for measuring specific health behaviours, including children's screen use. However, I found limited evidence of the validity of automated wearable camera measurements. This gap in the literature indicated that the overall body of evidence of automated wearable camera research may be lacking, and additional research on the validation of automated wearable cameras should be conducted.

To address this limitation, I developed a coding protocol (Chapter 3) for automated wearable camera measurements to assess screen-based behaviours within children and evaluated its convergent validity (Chapter 4) among children aged 8-11 years old in a home setting. In Chapter 4, I also examined the impact of the camera epoch on the camera's measurements of screen use. To evaluate the convergent validity, I compared automated wearable camera measurements with data captured by direct observation. I chose direct observation as it provides the most comprehensive, reliable and valid measurement of screen use to date (Perez et al., 2023). From this study, I concluded that automated wearable cameras had a strong convergent validity with direct observation for measuring screen use duration, type of device, social environment, content, associated behaviours, such as multitasking, and location. However, I found poor validity for food-related behaviours. I also found that automated wearable cameras may underestimate screen use duration by approximately 5 minutes when compared to direct observation, with small variations observed in camera epoch (i.e., 2-second, 10-second, 20-second, and 30-second) and device exposures (i.e., television, tablet, mobile device, computer).

In my final study (Chapter 5), I aimed to examine whether estimates of screen use from my validated automated wearable camera measurements were similar to a self- and parent-reported measurements of screen use. To do this, I used baseline data from the KidVision project, which used automated wearable cameras and parent- and child-reported questionnaires to measure children's screen use. From this study, I concluded that parent- and child-reported measurements of screen use tend to be inaccurate compared to automated wearable camera measurements. The findings also indicated that children and parents were more likely to overestimate than underestimate children's screen use for all screen use measurements; however, the extent of overestimating may not be consistent between individuals or screen use activities.

In this sixth and final chapter, I have summarised my key findings from this thesis and discussed them within the broader literature. I discussed the accuracy of automated wearable camera measurements of screen use, the accuracy of self-reported measurements of screen use compared to automated wearable cameras, and the limitations of automated wearable cameras as a measurement instrument. In addition, I discussed the strengths and limitations of this thesis and provided recommendations for future research.

Summary and Interpretation of Key Findings

Automated Wearable Cameras may be an Accurate Measurement of Screen Use

The findings of the included studies demonstrated that automated wearable cameras may provide an accurate measurement of children's screen use. The systematic review (Chapter 2) and validation study (Chapter 4) demonstrated that automated wearable cameras may accurately measure children's screen use. In Chapter 2, I concluded that automated wearable cameras are a feasible and reliable method for measuring the type of device and location of screen use behaviours. However, I found no evidence of the validity of automated wearable cameras for measuring children's screen use. While a measurement instrument can have high reliability, it does not equate to the measurement instrument being valid (Kimberlin & Winterstein, 2008). Therefore, it is possible for automated wearable camera measurements to be reliable but have poor validity. Given the importance of improving the quality of research on children's screen use, it is important to consider the validity of automated wearable camera measurements. To evaluate the validity of automated wearable cameras in measuring children's screen use, I conducted the validation study presented in Chapter 4.

In Chapter 4, I found strong convergent validity for automated wearable camera measurements with direct observation for screen time duration, type of device, social environment, content, associated behaviours, such as multitasking, and location, but weak convergent validity for food-related behaviours. Additionally, nearly all of the included studies measured other contextual factors alongside the health behaviours, most commonly the location of the health behaviour and whether the participant was alone or with other people. Previous studies using automated wearable cameras have reliably identified the location of behaviours (Leask et al., 2015; Signal, Stanley, et al., 2017; Watkins et al., 2018), the social context of behaviours (Gemming, Doherty, et al., 2015), and the type of devices and duration of screen-based behaviours (Hänggi et al., 2020; Smith et al., 2019). These findings are also consistent with my findings from Chapter 2, where I found mixed evidence on the reliability of automated camera measurements of food-related behaviours. The inclusion of contextual factors is consistent with the literature's recognition of the importance of ecological validity in health behaviour research (Shiffman et al., 2008). Two studies reported that dietitians often misidentified foods, suggesting that verification interviews should be used alongside automated wearable camera images when assessing portion sizes and calorie metrics (Beltran et al., 2016, 2018). The authors con concluded it was challenging to measure eating behaviours using automated wearable cameras alone due to the infrequent short duration of eating behaviours (Beltran et al., 2016, 2018). Other studies have also suggested that automated wearable cameras should be used in combination with image-assisted recall to improve the validity and reliability of reporting of dietary behaviour (Gemming et al., 2013; O'Loughlin et al., 2013; Zhou et al., 2019). This approach aligns with best practices in dietary measurement, which advocate for the combination of multiple methods to mitigate the limitations of any single approach (Gemming, Utter, et al., 2015). Moreover, to my knowledge, Chapter 4 was the first study to demonstrate the convergent validity of automated wearable camera measurements with direct observation for screen time duration, device duration and type, and location.

Despite this, automated wearable camera measurements of screen use may not be free of measurement error. In the comparison study presented in Chapter 5, I found a high level of variance in the Bland-Altman plots. The high level of variance indicates discrepancies between the automated wearable camera and self-reported measurements of screen use (Mansournia et al., 2021). While the high level of variance may be due to systematic discrepancies in parent- and child-reported measurements, the variance could also be due to measurement errors in the automated wearable camera measurements. There are several sources that may have introduced measurement errors into the screen use measurements.

These sources include unobserved screen use behaviours, participants changing their behaviour due to the camera's presence, and the impact of the camera epoch on screen time calculations. I have discussed each of these sources below.

Automated wearable cameras may have missed screen use behaviours. My findings in Chapter 4 may not represent the accuracy of automated wearable camera measurements in a free-living setting. In Chapter 4, I compared the automated wearable camera measurements with measurements from direct observation in a semi-structured observation session in participants' homes. I structured the direct observation sessions around the participant's usual screen use interactions to encourage them to interact with the screen-based media devices as if it was a typical day. However, this study did not take into account wear time adherence, which remains a common issue for device-based measurements in free-living settings (Hardy et al., 2013). Thus, the results of this study demonstrated that automated wearable cameras could accurately measure children performing screen use behaviours in a home setting but does not indicate the accuracy of the measurements when used in a free-living setting. Consequently, in Chapter 5, the automated wearable camera measurements may have missed screen use behaviours due to the participants not wearing the camera. To increase adherence to the study protocol in Chapter 5, text-messages were sent to parents in the morning to remind participants to wear the camera. While this technique has been shown to help increase adherence in health care services, participant's may still have forgotten to wear the camera (Schwebel & Larimer, 2018).

Participants may have changed their behaviour due to the Hawthorne effect. Participants may have chosen not to wear the camera while engaging in screen use (i.e., social desirability bias) or changed their behaviour while wearing the camera (i.e., Hawthorne effect). Previous studies using automated wearable cameras have noted that the camera's presence may change the participant's behaviour via the Hawthorne effect (Barr et al., 2015).

To overcome this, in Chapter 5, we blinded the participants from the main study aim. This blinding technique has been used in a previous automated wearable camera study (Signal, Smith, et al., 2017). While this technique may help mitigate the effect to some extent, participants may still have changed their behaviour because of the presence of the automated wearable camera. However, qualitative evidence has shown that individuals wearing the camera often forget about the camera's presence, remembering it only sporadically (Wilson et al., 2016).

Finally, the camera epoch may have contributed to under-reporting the automated wearable camera screen use measurements. In Chapter 4, I found that automated wearable cameras may underestimate screen use duration by approximately 5 minutes compared to direct observation, with variations observed in camera epoch and type of device. I found that a camera epoch of 20 seconds had a slightly larger systematic bias compared to a camera epoch of 2 seconds, 10 seconds or 30 seconds for total screen exposure, television exposure, computer exposure and mobile device exposure. However, the difference may be too small to be meaningful. Overall, the camera epoch may introduce a small systematic bias into the calculation of the screen time estimates as the screen time estimate would lose epochs throughout the screen exposure episode. Supporting this finding, other studies have found that automated wearable camera measurements may introduce a small systematic bias through the camera epoch as the measurement would lose time at the start and gain time at the end of the behaviour (Kelly et al., 2012). These measurement errors may have resulted in the automated wearable cameras under-reporting children's actual screen use.

Overall, the body of research in this thesis concludes that automated wearable cameras may provide a more accurate measurement of children's screen use, compared to self-reported measurements, but the measurements may not be completely free of measurement error and bias. Additional research needs to be conducted on the validation of automated wearable camera measurements to further explore the extent of measurement errors in free-living settings. I have addressed this recommendation in more details below.

Self-reported Measurements May be an Inaccurate Measurement of Screen Use

Despite the potential limitations explored above, the body of research in this thesis supported the existing literature that suggests parent- and child-reported measurements of screen use may be inaccurate compared to device-based measurements. In Chapter 5, I found weak correlations between automated wearable camera and parent- and child-reported measurements of aggregated screen time, screen use content, including educational screen use, recreational screen use and social screen use, and screen use activities, including gaming, internet browsing and programme viewing; however, the results were not significant. Some studies have found moderate correlations between self-reported measurements and phone tracking applications for social screen time and mobile device use (Junco, 2013; Wade et al., 2021). However, these studies were limited to single device and platform measurements, which some researchers suggest may not accurately reflect estimates for broader screen use (Mahalingham et al., 2023). In general, studies have found that self-reported measurements such as questionnaires typically have lower correlations to direct observation than device-based measurements (i.e., electronic TV monitors and smartphone applications) or media diaries (Perez et al., 2023).

In addition, the findings from this thesis suggest children and parents were more likely to overestimate than underestimate children's screen use for all screen use measurements compared to the automated wearable camera measurements. The Bland-Altman plot analyses in Chapter 5 indicated that children and parents were more likely to overestimate than underestimate children's screen use for all screen use measurements compared to the automated wearable camera measurements. In this study, I found that parent's estimates were more similar to automated wearable camera measurements for aggregated screen time, social screen use, gaming, and browsing. Children's estimates were more similar to automated wearable camera measurements for programme viewing. This contrasts with a previous study examining measurements of children's television viewing, which found that mothers' general estimates were more accurate than children's estimates (Borzekowski & Robinson, 1999). However, overall, these findings are consistent with previous research that has found that low levels of measurement often correspond with tendencies of over-reporting for internet use, social media use and mobile phone use (Junco, 2013; Vanden Abeele et al., 2013). This tendency to over-estimate children's screen use could lead to misinformed policymaking and intervention strategies. Public health guidelines that are based on inflated estimates of screen time may not accurately reflect '*true*' levels of screen use or its impact on children's health outcomes. This may contribute to overly restrictive recommendations on screen use which may not address harmful patterns of screen use and restrict the potential benefits of screen use. Therefore, adopting more accurate measurement instruments are crucial to ensure that public health recommendations effectively address screen use behaviours and their impacts on children's health outcomes.

An important finding from this thesis was that the extent of these inaccuracies in self-reported measurements of screen use may vary between individuals. That is, there may be individual factors that impact the accuracy of self-reported measurements. As I outlined earlier, in Chapter 5, I found a high level of variance in the Bland-Altman plots, indicating that the extent of over-reporting may differ between individuals. I hypothesised earlier that this could be due to measurement error in the automated wearable cameras; however, evidence that suggests the high level of variance and inaccuracy of self-reported screen use measurements may be due to systematic error (Parry et al., 2021). Parent- and self-reported screen use has shown low correlations with device-based measurements and direct observation (Perez et al., 2023). Further, some systematic reviews have found high levels of variance in effect sizes for correlations and ratio of means between device-based data and self-reported screen use (Parry et al., 2021). Some studies have suggested that the use of self-report may result in attenuated effect sizes, which has resulted in a reduced magnitude of observed effects (Jones-Jang et al., 2020). However, most studies suggest that the inaccuracy of self-reported screen use measurements may be due to systematic error (Parry et al., 2021). Previous studies have found that the accuracy of self-reported screen use may depend on individual factors such as how much the respondent uses the screen media, the type of device, and the respondent's well-being. For instance, a study found that higher levels of internet use were associated with higher levels of under-reporting and lower levels of internet use were associated with higher levels of over-reporting (Araujo et al., 2017). Another study found that the extent of inaccuracy between self-reported mobile phone use estimates and phone application tracked screen use was associated with respondents' level of well-being and amount of use (Sewall et al., 2020). While my findings provide support to existing literature that suggests there may be individual factors that impact the accuracy of self-reported measurements, there is limited evidence to identify the individual factors. Additional research is needed to identify potential factors and examine the likely systematic discrepancies in child- and parent- reported measurements of screen use. For instance, future studies should explore how individual differences in screen use patterns, device types and an individuals' well-being may affect self-report accuracy. Investigating how individual differences may affect self-report accuracy may provide deeper insights into the systematic errors observed in screen use measurements.

In Chapter 5, I also found evidence suggesting that the child- and parent-reported measurements may be more susceptible to recall bias than social desirability bias. In the literature, child-reported measurements have been shown to be an unreliable and inaccurate approach to screen use measurement due to the limited cognitive capacity and increased recall bias among paediatric populations in research (Atkin, Ekelund, et al., 2013; Saunders et al., 2011). Further, screen use behaviours are susceptible to biases due to their frequent and habitual nature (Parry et al., 2021). For example, respondents may underreport screen time due to negative social connotations associated with excess screen use or be unaware of how much time they are actually spending on screen-based media devices. In Chapter 5, I found that children were more likely to overestimate screen use than parents. One participant reported unrealistic screen use estimates aggregating to over 36 hours of daily screen time. Saunders et al (2011) have reported similar cases where children have reported unrealistically high amounts of daily screen time (e.g., 13.5 hours per day) when estimating their screen time. These findings further highlight concerns about the accuracy of child-reported measurements of screen use. Moreover, in Chapter 5, I found that children and parents were more likely to overestimate than underestimate children's screen use for all screen use measurements compared to the automated wearable camera measurements. This finding does not align with the theory that parents and children may be more likely to falsely report screen use behaviours to align with the national screen time guidelines. Instead, my findings are consistent with previous research that has found that low levels of measurement often correspond with tendencies of over-reporting for internet use, social media use and mobile phone use (Araujo et al., 2017; Junco, 2013; Vanden Abeele et al., 2013).

Overall, the body of research in this thesis contributes to the literature suggesting that parent- and child-reported measurements of screen use may be inaccurate when compared to device-based measurements. The findings also indicate that parent and child-reported screen use may be similar when estimating children's total screen time, social screen use, gaming, and programme viewing, but not internet browsing. This suggests that both groups may be subject to similar biases or limitations when estimating children's screen use. The similarity in inaccuracy raises important questions about the reliability of self-reported data in screen use research, regardless of whether it is reported by parents or children. Further, this thesis's findings support existing literature that suggests that the extent of these inaccuracies may vary between individuals. Additional research should be conducted to identify potential factors that may impact the accuracy of individual reports of screen use. Moreover, the findings of this thesis raise concerns about the accuracy of child- and parent-reported measurements of screen use used in the current literature, especially for child-reported measurements.

Limitations of Automated Wearable Cameras

Beyond the measurement of screen use, automated wearable cameras as a measurement instrument have some limitations. The first limitation is that image coding is time and resource intensive. This was first demonstrated in my findings in Chapter 2, where I found that simple coding protocols averaged around 30 minutes per participant (Kelly et al., 2012), while more complex coding protocols ranged from 40 minutes to 9 hours (Beltran et al., 2016; Zhou et al., 2019). I then encountered this challenge in my testing and Chapters 4 and 5. To address this issue, I collapsed specific subcategories into broader categories to minimise the amount of labour required to code each image and used an object classification model in Chapter 5 to reduce the volume of data needing to be coded. Other studies have used similar strategies to reduce the amount of image coding. For instance, studies on adult populations have developed data algorithms and object classification methodology to reduce the amount of image coding (Doherty et al., 2011; Rosenberg et al., 2017). However, while these techniques did reduce the amount of image coding, image coding was still time and resource intensive (i.e., it took over 26 hours to code 27 participants in Chapter 5). For this reason, automated wearable cameras may not be feasible in a study with a large sample size with the current methodology.

Automated wearable camera measurement instruments may only be feasible in certain population groups. The studies in this thesis have shown that some individuals are willing to wear the camera during all waking hours. However, not everyone was willing to participate. For instance, in Chapter 2, I found it was primarily non-participants during recruitment who expressed concerns about wearing the camera. Moreover, in Chapter 5, all included parents were mothers, and most had completed a postgraduate degree. It is likely that the use of automated wearable cameras to measure health behaviours may only be feasible in certain population groups. This may limit the generalisability of findings in automated wearable camera studies.

Finally, automated wearable cameras may be susceptible to the Hawthorne effect. Participants may change their behaviour when wearing the automated wearable camera. To overcome this, some studies have blinded the participants from the main study aim (Barr et al., 2015; Signal, Smith, et al., 2017). While this technique may help mitigate the effect to some extent, participants may still change their behaviour because of the camera's presence. For these reasons, automated wearable cameras may be best used as a substitute for direct observation to validate other measurement instruments that can be used across larger populations.

Strengths

This thesis had several strengths that support the findings of the included studies. In Chapter 2, I collated the current evidence of studies that used automated wearable cameras to measure health behaviours in youth. This was the first systematic review to provide an overview of the use of automated wearable cameras to measure health behaviours in young people. I conducted the systematic review using best practice, including prospective registration of the study protocol on PROSPERO (#CRD42021213532) and reported the findings according to the Preferred Reporting Items for Systematic Reviews and Meta-

Analyses (PRISMA) statement (Page et al., 2021). Further, I used an ‘objective approach’ developed by Hausner et al. (2015) to generate my search terms. I chose to do this to ensure the search for the studies was comprehensive, minimise subjectivity, and maximise sensitivity in the search strategy. Using this approach, I identified 23 studies that represented several different countries and study populations. These techniques ensured that the review was comprehensive and helped minimise potential areas of bias.

Another key strength of this thesis was that I developed the image coding protocol (Chapter 3) based on a step-by-step guideline for developing and modifying behaviour coding protocols in paediatric populations (Chorney et al., 2015). I followed this guideline to ensure the image coding protocol was created systematically to increase inter-rater reliability and validity, ensure that all relevant codes were identified and decrease the chance of bias occurring through the development phase. In this process, I used and adapted existing coding protocols to increase the consistency of coding frameworks in automated wearable camera research, which was a gap in the literature I identified in Chapter 2. Further, I tested the image coding protocol in a rigorous iterative cycle to refine and optimise the protocol’s functionality. This minimised the amount of labour required to code each image, which is an issue in automated wearable camera research. This method ensured that the coding protocol was comprehensive, functional, and helped reduce the chance of bias occurring in the coding protocol.

A final strength of this thesis was that I compared automated wearable camera measurements to a commonly used screen use measurement instrument in Chapter 5 (SBMUS; Houghton et al., 2015). In Chapter 5, I measured parent- and child-reported screen use using questions from the SBMUS (Houghton et al., 2015) The SBMUS measures daily screen use and includes newer screen-based media devices and activities. The SBMUS has been tested in children and adolescents aged 8-16 years old and has a test-retest reliability of

0.52 (Houghton et al., 2015). I compared the automated wearable camera measurements to a commonly used screen use measurement instrument to ensure the findings from the study could be applied to the general screen use literature.

Limitations

This thesis also has some limitations. As discussed earlier, the automated wearable camera measurements may not be a *true* representation of screen use behaviours. While I evaluated the validity of automated wearable camera measurements in a home setting, the findings from this study may not be a true representation of automated wearable camera measurements in a free-living setting. Further, the automated wearable camera measurements may have missed screen use behaviours due to non-adherence to the study protocol. For instance, participants may have forgotten to wear the camera or forgotten to switch the camera on after switching the camera off due to privacy reasons (e.g., going to the bathroom). Wear time adherence is a common issue for device-based measurements in free-living settings (Hardy et al., 2013). To overcome this in Chapter 5, text messages were sent to parents in the morning to remind participants to wear the camera. Text message reminders have increased adherence to health care services (Schwebel & Larimer, 2018). However, there is still a possibility that participants did not adhere to the study protocol. As such, future research studies could prepare strategies to mitigate non-adherence to the study protocol.

Another limitation was that automated wearable camera measurements may be susceptible to the Hawthorne effect. That is, participants may have changed their behaviour while wearing the camera. To overcome this, in Chapter 5, participants were blinded to the main study aim. This blinding technique has been used in previous automated wearable camera studies (Barr et al., 2015; Signal, Smith, et al., 2017). This technique may help mitigate the effect to some extent; however, some participants may have changed their behaviour due to the presence of the automated wearable camera.

An additional limitation was that I only assessed the convergent validity of the automated wearable camera measurements. In the process of validating a measurement instrument, evidence from different types of validity should be used to assess the degree of validity of the instrument in the specific context and population (de Vet et al., 2011). As such, future research on automated wearable camera measurements of screen use needs to assess different types of validity, such as content validity and criterion validity, to conclude the degree of validity of automated wearable camera measurements.

Further, Chapters 4 and 5 had relatively small sample sizes. The small sample sizes may have influenced estimates through sampling error (Hopkins et al., 2009). Contributing to this, the findings of this thesis may not be generalisable to all children and parents from other cultures or countries. While I tried to include participants from diverse backgrounds, participants in Chapter 5 were primarily mothers who had completed a postgraduate degree (35%). Therefore, the results should be interpreted with caution when generalising to the wider population. Future studies should examine the similarity between automated wearable camera measurements and child- and parent-reported measurements of screen use in different demographics.

Implications for Policy, Practice and Research

Overall, this thesis aimed to examine the use of automated wearable cameras to establish a more accurate measure of children's screen use and examine whether estimates of screen use from automated wearable cameras were similar to self-reported measurements. The findings from Chapters 2 and 4 demonstrated that automated wearable cameras can accurately measure the device, content and context of children's screen use behaviours. At present, most studies have used aggregated 'total' screen use measurements, that do not take into account the different types of devices, content and contexts of screen use that may impact the effects of screen use on children's health and behaviour outcomes (Odgers &

Jensen, 2020; Sanders et al., 2019). This methodology limits our understanding of how the different types of devices, content, and contexts may affect children's outcomes differently. Thus limiting our understanding of the true impact of screen use on children's health and behaviour outcomes. Given the challenge of accurately measuring children's screen use behaviours, the findings from this thesis demonstrated the potential of automated wearables cameras as a measurement instrument to accurately measure complex screen use behaviours, such as the content and context of the behaviour.

Further, the findings from Chapters 2 and 4 demonstrated that automated wearable cameras may reliably measure other health behaviours and identified important factors researchers should consider when using automated wearable cameras among youth. These factors included ethical and privacy considerations, image quality, camera placement, battery life and time for manual image coding. These findings will assist researchers' decision-making when using automated wearable cameras to measure health behaviours in youth and improve the methodological standards around automated wearable camera measurements.

Currently, much of the research is based on self- or proxy-reported measurements of screen use (Ainsworth et al., 2012; Clark et al., 2011). The findings from this thesis (specifically Chapter 5) suggest that parent- and child-reported measurements of screen use may be inaccurate when compared to device-based measurements. Moreover, the findings suggest that parent- and child-reported measurements tend to overestimate than underestimate children's screen use compared to automated wearable camera measurements. These findings are consistent with previous studies that have shown that self-reported measurements such as questionnaires typically have lower correlations to direct observation than device-based measurements (i.e., electronic TV monitors and smartphone applications) or media diaries (Perez et al., 2023). Thus, the findings of this thesis raise concerns about the accuracy of child- and parent-reported measurements of screen use. Given the widespread use of child-and parent-reported measurements of screen use in the academic literature, much of our understanding of the impact of screen use on children's health and behaviour outcomes has been informed from studies relying solely on self-reported measurements of screen use. The findings from this thesis suggest that research or policy recommendations from studies solely relying on self-reported measurements of screen use should be interpreted with caution.

Recommendations for Future Research

Given the challenge of accurately measuring young people's health behaviours, automated wearable cameras have the potential for measuring multiple health behaviours; however, additional research is needed to strengthen the methodology. As I stated in Chapter 2, there is limited research examining the validity of automated wearable camera measurements. To address this gap in the literature, in Chapter 4, I evaluated the convergent validity of automated wearable camera measurements of screen behaviours among children in a home setting. The findings of this study demonstrated that automated wearable cameras can accurately measure children performing screen use behaviours; however, these findings may not be a true representation of the use of automated wearable cameras in a free-living setting. Further, in Chapter 4, I only assessed the convergent validity of automated wearable camera measurements in children. As such, additional research also needs to be conducted on the different types of validity, such as content validity and criterion validity, in different contexts (i.e., in public settings), and with different populations (i.e., younger children and adolescents) to conclude on the degree of validity of automated wearable camera measurements. To assess the criterion validity of automated wearable camera measurements in a free-living setting, future studies should consider validating automated wearable camera measurements with direct observation (i.e., the gold standard measurement). These studies could observe participants daily while they wear the automated wearable camera. To assess the content validity of the coding protocols, future studies could ask experts in the field to review the coding protocol. Studies could also ask a small sample from the target population to review the coding protocol to gain qualitative feedback on the accuracy of categories in the coding protocol based on the target population's perspective. In doing so, researchers could assess different types of validity to conclude the overall degree of validity of automated wearable camera measurements for measuring certain health behaviours.

Although studies in this thesis aligned with evidence of systematic discrepancies for parent- and child-reported screen use measurements, our understanding of the underlying mechanisms remains limited. In Chapter 5, I found evidence that the extent of over-reporting screen use in self-reported measurements may be different between individuals. Previous studies have found that the accuracy of self-reported measurements of screen use may depend on the respondent's level of well-being and amount of use (Araujo et al., 2017; Sewall et al., 2020). However, further evidence is needed to examine the extent of these systematic discrepancies in self-reported measurements (Parry et al., 2021). To address this gap in the literature, future studies could investigate individual factors that may be contributing to the variability in self-reported screen use measurements. While statistical approaches may not account for all sources of error in self-reported measurements of screen use if future research can identify the individual factors that account for the discrepancies, statistical approaches can be modelled and used to account for measurement error in self-reported measurements (Parry et al., 2021). Thus, future studies can contribute to the refinement of screen use measurement instruments and develop strategies to minimize potential bias and improve the overall accuracy of self-reported screen use data. Doing so will improve the robustness of self-reported measurements in screen use research and contribute to improving our understanding of the impact of screen use on children's health and behaviour outcomes.

Significance of the Thesis

- Study 1 filled the gap in the literature by systematically synthesising the current evidence of the use of automated wearable cameras to measure health behaviours in child and adolescent populations. The gaps identified within this study are addressed in Study 2.
- Study 2 was the first study to investigate the convergent validity of automated wearable camera measurements for assessing children's screen-based behaviours compared to direct observation.
- Study 1 and Study 2 will assist researchers' decision-making when using automated wearable cameras to measure health behaviours in youth and improve the methodological standards around automated wearable camera measurements.
- Study 3 was the first study to compare automated wearable camera measurements (device-based) of screen use with a self- and proxy-reported measure of screen use.
- The results of all three studies will provide evidence and contribute to the growing body of literature on the use of automated wearable cameras in research.

Conclusion

As outlined in detail in this thesis, the influence of screen-based media devices on children's health outcomes remains controversial, with the current research showing inconsistent findings (Stiglic & Viner, 2019). Much of the inconsistency between studies may be due to the way screen use has been measured, as the field has largely relied on unvalidated self- or parent-reported measurements (Kaye et al., 2020; Stiglic & Viner, 2019). The aim of this thesis was to examine the use of automated wearable cameras to establish a more accurate measure of children's screen use and then examine whether estimates of screen use from automated wearable camera measurements were similar to self-report measurements. In Chapter 2, I synthesised the current academic literature and found that automated wearable cameras may provide a reliable method for measuring specific health behaviours; however, there was limited evidence on the validity of automated wearable camera measurements. In Chapter 4, I addressed this gap by conducting a validation study. I found that automated wearable camera measurements of screen use duration, type of device, social environment, content, associated behaviours, such as multitasking, and location of the screen use had strong convergent validity with direct observation measurements but weak convergent validity for food-related behaviours. In Chapter 5, I investigated whether estimates of screen use from the automated wearable camera measurements were similar to a self-report measurements of screen use. I found that parent- and child-reported measurements of screen use were inaccurate compared to automated wearable camera measurements, with children and parents being more likely to overestimate than underestimate children’s screen use for all screen use measurements. I also found evidence that suggests that there may be systematic discrepancies in child- and parent-reported screen use measurements.

Overall, the findings of this thesis raise concerns about the accuracy of child- and parent-reported measurements of screen use. Given the widespread use of child- and parent-reported measurements of screen use in the academic literature, such research should be interpreted with caution. Additionally, given the challenge of accurately measuring children’s screen use behaviours, automated wearable cameras have the potential for accurately measuring complex screen use behaviours such as the content and context of the behaviour. The studies presented in this thesis have contributed both substantively and methodologically to development of automated wearable camera research literature and screen use literature.

References

Ainsworth, B. E., Caspersen, C. J., Matthews, C. E., Mâsse, L. C., Baranowski, T., & Zhu, W. (2012). Recommendations to improve the accuracy of estimates of physical activity derived from self report. *Journal of Physical Activity & Health*, *9 Suppl 1*, S76-84. https://www.ncbi.nlm.nih.gov/pubmed/22287451

Allé, M. C., Manning, L., Potheegadoo, J., Coutelle, R., Danion, J.-M., & Berna, F. (2017). Wearable cameras are useful tools to investigate and remediate autobiographical memory impairment: A systematic PRISMA review. *Neuropsychology Review*, *27*(1), 81–99. https://doi.org/10.1007/s11065-016-9337-x

Anderson, D., Field, D., Collins, P., Lorch, E., & Nathan, J. (1985). Estimates of young children's time with television: A methodological comparison of parent reports with time-lapse video home observation. *Child Development*, *56*(5), 1345–1357. https://doi.org/10.1111/j.1467-8624.1985.tb00202.x

Andriyani, F.D., Biddle, S. J. H., Priambadha, A. A., Thomas, G. & Cocker, K. (2022). Physical activity and sedentary behaviour of female adolescents in Indonesia: A multi-method study on duration, pattern and context, *Journal of Exercise Science & Fitness, 20*(2), 128-139. https://doi.org/10.1016/j.jesf.2022.02.002

Araujo, T., Wonneberger, A., Neijens, P., & de Vreese, C. (2017). How much time do you spend online? Understanding and improving the accuracy of self-reported measures of internet use. *Communication Methods and Measures*, *11*(3), 173–190. https://doi.org/10.1080/19312458.2017.1317337

Arifin, W. N. (2023). *Sample size calculator (web)*. http://wnarifin.github.io

Atkin, A. J., Corder, K., & van Sluijs, E. M. F. (2013). Bedroom media, sedentary time and screen-time in children: A longitudinal analysis. *The International Journal of*

Behavioral Nutrition and Physical Activity, *10*, 137. https://doi.org/10.1186/1479-5868-10-137

Atkin, A. J., Ekelund, U., Møller, N. C., Froberg, K., Sardinha, L. B., Andersen, L. B., & Brage, S. (2013). Sedentary time in children: Influence of accelerometer processing on health relations. *Medicine and Science in Sports and Exercise*, *45*(6), 1097–1104. https://doi.org/10.1249/MSS.0b013e318282190e

Australian Catholic University. (n.d.). *Research Code of Conduct*. ACU. Retrieved August 18, 2020, from https://policies.acu.edu.au/research/general_policies/research_code_of_conduct

Barr, M., Signal, L., Jenkin, G., & Smith, M. (2015). Capturing exposures: Using automated cameras to document environmental determinants of obesity. *Health Promotion International*, *30*(1), 56–63. https://doi.org/10.1093/heapro/dau089

Bartel, K. A., Gradisar, M., & Williamson, P. (2015). Protective and risk factors for adolescent sleep: A meta-analytic review. *Sleep Medicine Reviews*, *21*, 72–85. https://doi.org/10.1016/j.smrv.2014.08.002

Bauman, A., Bittman, M., & Gershuny, J. (2019). A short history of time use research; implications for public health. *BMC Public Health*, *19*(2), 607.

Bélanger, R. E., Akre, C., Berchtold, A., & Michaud, P.-A. (2011). A U-shaped association between intensity of Internet use and adolescent health. *Pediatrics*, *127*(2), e330-5. https://doi.org/10.1542/peds.2010-1235

Bell, B. M., Alam, R., Alshurafa, N., Thomaz, E., Mondol, A. S., de la Haye, K., Stankovic, J. A., Lach, J., & Spruijt-Metz, D. (2020). Automatic, wearable-based, in-field eating detection approaches for public health research: A scoping review. *NPJ Digital Medicine*, *3*, 38. https://doi.org/10.1038/s41746-020-0246-2

Bell, W., Colaiezzi, B. A., Prata, C. S., & Coates, J. C. (2017). Scaling up dietary data for decision-making in low-income countries: New technological frontiers. *Advances in Nutrition*, *8*(6), 916–932. https://doi.org/10.3945/an.116.014308

Beltran, A., Dadabhoy, H., Chen, T. A., Lin, C., Jia, W., Baranowski, J., Yan, G., Sun, M., & Baranowski, T. (2016). Adapting the eButton to the abilities of children for diet assessment. *Proceedings of Measuring Behavior 2016*, *2016*, 72–81. https://www.ncbi.nlm.nih.gov/pubmed/31742257

Beltran, A., Dadabhoy, H., Ryan, C., Dholakia, R., Jia, W., Baranowski, J., Sun, M., & Baranowski, T. (2018). Dietary assessment with a wearable camera among children: Feasibility and intercoder reliability. *Journal of the Academy of Nutrition and Dietetics*, *118*(11), 2144–2153. https://doi.org/10.1016/j.jand.2018.05.013

Biddle, S. J. H., Bengoechea García, E., Pedisic, Z., Bennie, J., Vergeer, I., & Wiesner, G. (2017). Screen time, other sedentary behaviours, and obesity risk in adults: A review of reviews. *Current Obesity Reports, 6,* 134-147. https://doi.org/10.1007/s13679-017-0256-9

Biddle, S. J. H., Bengoechea García, E., & Wiesner, G. (2017). Sedentary behaviour and adiposity in youth: A systematic review of reviews and analysis of causality. *The International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 43. https://doi.org/10.1186/s12966-017-0497-8

Biswas, D., Su, H., Wang, C., Blankenship, J., & Stevanovic, A. (2017). An automatic car counting system using overfeat framework. *Sensors*, *17*(7), 1535. https://doi.org/10.3390/s17071535

Bittman, M., Rutherford, L., Brown, J., & Unsworth, L. (2011). Digital natives? New and old media and children's outcomes. *Australian Journal of Education in Chemistry*, *55*(2), 161–175. https://doi.org/10.1177/000494411105500206

Bland, J. M., & Altman, D. G. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*, *1*(8476), 307–310. https://www.ncbi.nlm.nih.gov/pubmed/2868172

Bland, J. M., & Altman, D. G. (1995). Comparing methods of measurement: Why plotting difference against standard method is misleading. *The Lancet*, *346*(8982), 1085–1087. https://doi.org/10.1016/s0140-6736(95)91748-9

Boak, A., Hamilton, H. A., Adlaf, E. M., Henderson, J. L., & Mann, R. E. (2020). *The mental health and well-being of Ontario students, 1991-2019: Detailed findings from the Ontario Student Drug Use and Health Survey (OSDUHS)*. Toronto, ON: Centre for Addication and Mental health.

Bochkovskiy, A., Wang, C.-Y., & Liao, H.-Y. M. (2020, April 23). *YOLOv4: Optimal speed and accuracy of object detection*. https://doi.org/10.48550/arXiv.2004.10934

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica*, *129*(3), 387–398. https://doi.org/10.1016/j.actpsy.2008.09.005

Borca, G., Bina, M., Keller, P. S., Gilbert, L. R., & Begotti, T. (2015). Internet use and developmental tasks: Adolescents' point of view. Computers in Human Behavior, 52, 49-58. https://doi.org/10.1016/j.chb.2015.05.029

Borzekowski, D. L. G., & Robinson, T. N. (1999). Viewing the viewers: Ten video cases of children's television viewing behaviors. *Journal of Broadcasting & Electronic Media*, *43*(4), 506–528. https://doi.org/10.1080/08838159909364507

Boushey, C. J., Spoden, M., Zhu, F. M., Delp, E. J., & Kerr, D. A. (2017). New mobile methods for dietary assessment: Review of image-assisted and image-based dietary assessment methods. *The Proceedings of the Nutrition Society*, *76*(3), 283–294. https://doi.org/10.1017/S0029665116002913

Brown, J. D., & Bobkowski, P. S. (2011). Older and newer media: Patterns of use and effects on adolescents' health and well-being. *Journal of Research on Adolescence*, *21*(1), 95–113. https://doi.org/10.1111/j.1532-7795.2010.00717.x

Bulungu, A. L., Palla, L., Priebe, J., Forsythe, L., Katic, P., Varley, G., Galinda, B. D., Sarah, N., Namboose, J., Wellard, K., & Ferguson, E. (2022). Validation of automated wearable camera-based image-assisted recall method and the 24-h recall method for assessing women's time allocation in nutritionally vulnerable population: The case of rural Uganda. *Nutrients*, *14*(9), 1833. https://doi.org/10.3390/nu14091833

Byrne, R., Terranova, C. O., & Trost, S. G. (2021). Measurement of screen time among young children aged 0-6 years: A systematic review. *Obesity Reviews*, *22*(8), e13260. https://doi.org/10.1111/obr.13260

Cain, M., Leonard, J., Gabrieli, J., & Finn, A. (2016). Media multitasking in adolescence. *Psychonomic Bulletin & Review*, *23*(6), 1932–1941. https://doi.org/10.3758/s13423-016-1036-3

Cain, N., & Gradisar, M. (2010). Electronic media use and sleep in school-aged children and adolescents: A review. *Sleep Medicine*, *11*(8), 735–742. https://doi.org/10.1016/j.sleep.2010.02.006

Calamaro, C. J., Mason, T. B. A., & Ratcliffe, S. J. (2009). Adolescents living the 24/7 lifestyle: Effects of caffeine and technology on sleep duration and daytime functioning. *Pediatrics*, *123*(6), e1005-10. https://doi.org/10.1542/peds.2008-3641

Cambridge University Press. (n.d.-a). Bedroom. In *Cambridge Dictionary*. Retrieved October 9, 2019, from https://dictionary.cambridge.org/dictionary/english/bedroom

Cambridge University Press. (n.d.-b). Desktop Computer. In *Cambridge Business English Dictionary*. Retrieved October 23, 2019, from https://dictionary.cambridge.org/dictionary/english/desktop-computer

Cambridge University Press. (n.d.-c). Home. In *Cambridge Dictionary*. Retrieved October 9, 2019, from https://dictionary.cambridge.org/dictionary/english/home

Cambridge University Press. (n.d.-d). Interactive Whiteboard. In *Cambridge Advanced Learner's Dictionary & Thesaurus*. Retrieved October 9, 2019, from https://dictionary.cambridge.org/dictionary/english/interactive-whiteboard?q=Interactive+Whiteboard+

Cambridge University Press. (n.d.-e). Kitchen. In *Cambridge Advanced Learner's Dictionary*. Retrieved October 9, 2019, from https://dictionary.cambridge.org/dictionary/english/kitchen

Cambridge University Press. (n.d.-f). Living Room. In *Cambridge Advanced Learner's Dictionary & Thesaurus*. Retrieved October 9, 2019, from https://dictionary.cambridge.org/dictionary/english/living-room

Cambridge University Press. (n.d.-g). Mobile Device. In *Cambridge Advanced Learner's Dictionary & Thesaurus*. Retrieved October 9, 2019, from https://dictionary.cambridge.org/dictionary/english/mobile-device

Cambridge University Press. (n.d.-h). Projector. In *Cambridge Advanced Learner's Dictionary & Thesaurus*. Retrieved October 9, 2019, from https://dictionary.cambridge.org/dictionary/english/projector

Cambridge University Press. (n.d.-i). Smartwatch. In *Cambridge Advanced Learner's Dictionary & Thesaurus*. Retrieved September 10, 2019, from https://dictionary.cambridge.org/dictionary/english/smartwatch

Cambridge University Press. (n.d.-j). Social Media. In *Cambridge Advanced Learner's Dictionary & Thesaurus*. Retrieved October 9, 2019, from https://dictionary.cambridge.org/dictionary/english/social-media

Cambridge University Press. (n.d.-k). Tablet. In *Cambridge Advanced Learner's Dictionary & Thesaurus*. Retrieved October 8, 2019, from https://dictionary.cambridge.org/dictionary/english/tablet

Cambridge University Press. (n.d.-l). Television. In *Cambridge Advanced Learner's Dictionary & Thesaurus*. Retrieved October 12, 2019, from https://dictionary.cambridge.org/

Carson, V., Hunter, S., Kuzik, N., Gray, C. E., Poitras, V. J., Chaput, J.-P., Saunders, T. J., Katzmarzyk, P. T., Okely, A. D., Connor Gorber, S., Kho, M. E., Sampson, M., Lee, H., & Tremblay, M. S. (2016). Systematic review of sedentary behaviour and health indicators in school-aged children and youth: An update. *Applied Physiology, Nutrition, and Metabolism*, *41*(6), S240-65. https://doi.org/10.1139/apnm-2015-0630

Charmaramen, L., Cobas, S., Weed, J., Gu, Q., Kiel, E., Chin, H., Gramajo, A. & Mueller, M. K. (2022). From regulating emotions to less lonely screen time: Parents' Qualitative perspectives of the benefits and challenges of adolescent pet companionship, *Behavioral Sciences, 12*(5), 143. https://doi.org/10.3390/bs12050143

Chen, Q., Song, Z., Dong, J., Huang, Z., Hua, Y., & Yan, S. (2015). Contextualizing object detection and classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *37*(1), 13–27. https://doi.org/10.1109/TPAMI.2014.2343217

Chinchanachokchai, S., Duff, B. R. L., & Sar, S. (2015). The effect of multitasking on time perception, enjoyment, and ad evaluation. *Computers in Human Behavior*, *45*, 185–191. https://doi.org/10.1016/j.chb.2014.11.087

Chorney, J. M., McMurtry, C. M., Chambers, C. T., & Bakeman, R. (2015). Developing and modifying behavioral coding schemes in pediatric psychology: A practical guide. *Journal of Pediatric Psychology*, *40*(1), 154–164. https://doi.org/10.1093/jpepsy/jsu099

Clark, B., Healy, G., Winkler, E., Gardiner, P., Sugiyama, T., Dunstan, D., Matthews, C., & Owen, N. (2011). Relationship of television time with accelerometer-derived sedentary time: NHANES. *Medicine and Science in Sports and Exercise*, *43*(5), 822–828. https://doi.org/10.1249/MSS.0b013e3182019510

Clark, J., Glasziou, P., Del Mar, C., Bannach-Brown, A., Stehlik, P., & Scott, A. M. (2020). A full systematic review was completed in 2 weeks using automation tools: A case study. *Journal of Clinical Epidemiology*, *121*, 81–90. https://doi.org/10.1016/j.jclinepi.2020.01.008

Costigan, S. A., Barnett, L., Plotnikoff, R. C., & Lubans, D. R. (2013). The health indicators associated with screen-based sedentary behavior among adolescent girls: A systematic review. *The Journal of Adolescent Health*, *52*(4), 382–392. https://doi.org/10.1016/j.jadohealth.2012.07.018

Cowburn, G., Matthews, A., Doherty, A., Hamilton, A., Kelly, P., Williams, J., Foster, C., & Nelson, M. (2016). Exploring the opportunities for food and drink purchasing and consumption by teenagers during their journeys between home and school: A feasibility study using a novel method. *Public Health Nutrition*, *19*(1), 93–103. https://doi.org/10.1017/S1368980015000889

Coyne, S. M., Padilla-Walker, L. M., & Holmgren, H. G. (2018). A six-year longitudinal study of texting trajectories during adolescence. *Child Development,* 89(1), 58-65. https://doi.org/10.1111/cdev.12823

de Vet, H., Terwee, C., Mokkink, L., & Knol, D. (2011). *Measurement in Medicine: A Practical Guide*. Cambridge University Press.

Doherty, A., Caprani, N., Conaire, C., Kalnikaite, V., Gurrin, C., Smeaton, A., & O'Connor, N. (2011). Passively recognising human activities through lifelogging. *Computers in Human Behavior*, *27*(5), 1948–1958. https://doi.org/10.1016/j.chb.2011.05.002

Doherty, A., Hodges, S., King, A., Smeaton, A., Berry, E., Moulin, C., Lindley, S., Kelly, P., & Foster, C. (2013). Wearable cameras in health: The state of the art and future possibilities. *American Journal of Preventive Medicine*, *44*(3), 320–323. https://doi.org/10.1016/j.amepre.2012.11.008

Doherty, A., Kelly, P., Kerr, J., Marshall, S., Oliver, M., Badland, H., & Foster, C. (2012). Use of wearable cameras to assess population physical activity behaviours: An observational study. *The Lancet*, *380*, S35. https://doi.org/10.1016/S0140-6736(13)60391-8

Doherty, A., Kelly, P., Kerr, J., Marshall, S., Oliver, M., Badland, H., Hamilton, A., & Foster, C. (2013). Using wearable cameras to categorise type and context of accelerometer-identified episodes of physical activity. *The International Journal of Behavioral Nutrition and Physical Activity*, *10*, 22. https://doi.org/10.1186/1479-5868-10-22

Ellis, D. A., Davidson, B. I., Shaw, H., & Geyer, K. (2019). Do smartphone usage scales predict behavior? *International Journal of Human-Computer Studies*, *130*, 86–92. https://doi.org/10.1016/j.ijhcs.2019.05.004

Everson, B., Mackintosh, K. A., McNarry, M. A., Todd, C., & Stratton, G. (2019). Can wearable cameras be used to validate school-aged children’s lifestyle behaviours? *Children*, *6*(2). https://doi.org/10.3390/children6020020

Falbe, J., Davison, K. K., Franckle, R. L., Ganter, C., Gortmaker, S. L., Smith, L., Land, T., & Taveras, E. M. (2015). Sleep duration, restfulness, and screens in the sleep environment. *Pediatrics*, *135*(2), e367-75. https://doi.org/10.1542/peds.2014-2306

Ferguson, C. J. (2017). Everything in moderation: Moderate use of screens unassociated with child behavior problems. *The Psychiatric Quarterly*, *88*(4), 797–805. https://doi.org/10.1007/s11126-016-9486-3

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*(5), 378–382. https://doi.org/10.1037/h0031619

Floyd, F. J., Baucom, D. H., Godfrey, J. J., & Palmer, C. (1998). *Observational Methods*. *3*, 1–21. https://doi.org/10.1016/B0080-4270(73)00223-6

Freeman, N., Gage, R., Chambers, T., Blaschke, P., Cook, H., Stanley, J., Pearson, A., Smith, M., Barr, M., & Signal, L. (2020). Where do the children play? An objective analysis of children's use of green space. *Health Promotion International*, *36*(3), 846–853. https://doi.org/10.1093/heapro/daaa106

Gage, R., Barr, M., Stanley, J., Reeder, A., Mackay, C., Smith, M., Chambers, T., Leung, W., & Signal, L. (2018). Sun protection and shade availability in New Zealand's outdoor recreation spaces. *The New Zealand Medical Journal*, *131*(1484), 30–37.

Gage, R., Girling-Butcher, M., Joe, E., Smith, M., Ni Mhurchu, C., McKerchar, C., Puloka, V., McLean, R., & Signal, L. (2020). The frequency and context of snacking among children: An objective analysis using wearable cameras. *Nutrients*, *13*(1). https://doi.org/10.3390/nu13010103

Gage, R., Leung, W., Stanley, J., Reeder, A., Barr, M., Chambers, T., Smith, M., & Signal, L. (2017). Clothing protection from ultraviolet radiation: A new method for assessment. *Photochemistry and Photobiology*, *93*(6), 1513–1518. https://doi.org/10.1111/php.12803

Gage, R., Leung, W., Stanley, J., Reeder, A., Mackay, C., Chambers, T., Smith, M., Barr, M., & Signal, L. (2019). Studying third-parties and environments: New Zealand sun-safety research. *Health Promotion International*, *34*(3), 440–446. https://doi.org/10.1093/heapro/dax094

Gage, R., Leung, W., Stanley, J., Reeder, A., Mackay, C., Smith, M., Barr, M., Chambers, T., & Signal, L. (2018). Sun protection among New Zealand primary school children. *Health Education & Behavior*, *45*(5), 800–807.

Gagnier, J., Lai, J., Mokkink, L., & Terwee, C. (2021). COSMIN reporting guideline for studies on measurement properties of patient-reported outcome measures. *Quality of Life Research*, *30*, 2197–2218. https://doi.org/10.1007/s11136-021-02822-4

Gamer, M., Lemon, J., Fellows, I., & Singh, P. (2019). *irr: Various Coefficients of Interrater Reliability and Agreement* (0.84.1) [R-package]. https://CRAN.R-project.org/package=irr

Gemming, L., Doherty, A., Kelly, P., Utter, J., & Ni Mhurchu, C. (2013). Feasibility of a SenseCam-assisted 24-h recall to reduce under-reporting of energy intake. *European Journal of Clinical Nutrition*, *67*(10), 1095–1099. https://doi.org/10.1038/ejcn.2013.156

Gemming, L., Doherty, A., Utter, J., Shields, E., & Ni Mhurchu, C. (2015). The use of a wearable camera to capture and categorise the environmental and social context of self-identified eating episodes. *Appetite*, *92*, 118–125. https://doi.org/10.1016/j.appet.2015.05.019

Gemming, L., Utter, J., & Ni Mhurchu, C. (2015). Image-assisted dietary assessment: A systematic review of the evidence. *Journal of the Academy of Nutrition and Dietetics*, *115*(1), 64–77.

Gentile, D. A., Anderson, C. A., Yukawa, S., Ihori, N., Saleem, M., Ming, L. K., Shibuya, A., Liau, A. K., Khoo, A., Bushman, B. J., Huesmann, L. R., & Sakamoto, A. (2009). The effects of prosocial video games on prosocial behaviors: International evidemce form correlational, longitudinal and experimental studies. *Personality and Social Psychology Bulletin, 35*(6), 753-763. https://doi.org/10.1177/0146167209333045

Gochman, D. S. (1997). Health behavior research: Definitions and diversity. *Handbook of Health Behavior Research*, *1*, 5–34.

Granich, J., Rosenberg, M., Knuiman, M. W., & Timperio, A. (2011). Individual, social, and physical environment factors associated with electronic media use among children: Sedentary behavior at home. *Journal of Physical Activity & Health*, *8*(5), 613–625.

Greenberg, S. (2023). *Timelapse: An image analyser for camera traps* (2.3.0.6) [Computer software]. https://saul.cpsc.ucalgary.ca/timelapse/pmwiki.php?n=Main.HomePage

Greitemeyer, T., & Mügge, D. O. (2014). Video games do affect social outcomes: A meta-analytic review of the effects of violent and prosocial video game play. *Personality & Social Psychology Bulletin*, *40*(5), 578–589.

Griffioen, N., van Rooij, M., Lichtwarck-Aschoff, A., & Granic, I. (2020). Toward improved methods in social media research. *Technology, Mind, and Behavior*, *1*(1). https://doi.org/10.1037/tmb0000005

Grimm, P. (2010). *Social Desirability Bias in Encyclopedia of Marketing*. Wiley.

Haidet, K. K., Tate, J., Divirgilio-Thomas, D., Kolanowski, A., & Happ, M. B. (2009). Methods to improve reliability of video-recorded behavioral data. *Research in Nursing & Health*, *32*(4), 465–474. https://doi.org/10.1002/nur.20334

Hale, L., & Guan, S. (2015). Screen time and sleep among school-aged children and adolescents: A systematic literature review. *Sleep Medicine Reviews*, *21*, 50–58.

Hänggi, J., Spinnler, S., Christodoulides, E., Gramespacher, E., Taube, W., & Doherty, A. (2020). Sedentary behavior in children by wearable cameras: Development of an annotation protocol. *American Journal of Preventive Medicine*, *56*(6), 880–886.

Hardy, L. L., Hills, A. P., Timperio, A., Cliff, D., Lubans, D., Morgan, P. J., Taylor, B. J., & Brown, H. (2013). A hitchhiker's guide to assessing sedentary behaviour among young people: Deciding what method to use. *Journal of Science and Medicine in Sport*, *16*(1), 28–35.

Harrell, F. E., Jr. (2023). *Harrell Miscellaneous* (5.1-1) [R-package]. https://hbiostat.org/R/Hmisc/

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., McLeod, L., Delacqua, G., Delacqua, F., Kirby, J., Duda, S. N., & REDCap Consortium. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*(103208), 103208. https://doi.org/10.1016/j.jbi.2019.103208

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*, *42*(2), 377–381.

Hausner, E., Guddat, C., Hermanns, T., Lampert, U., & Waffenschmidt, S. (2015). Development of search strategies for systematic reviews: Validation showed the noninferiority of the objective approach. *Journal of Clinical Epidemiology*, *68*(2), 191–199. https://doi.org/10.1016/j.jclinepi.2014.09.016

Hausner, E., Guddat, C., Hermanns, T., Lampert, U., & Waffenschmidt, S. (2016). Prospective comparison of search strategies for systematic reviews: An objective approach yielded higher sensitivity than a conceptual one. *Journal of Clinical Epidemiology*, *77*, 118–124. https://doi.org/10.1016/j.jclinepi.2016.05.002

Hausner, E., Waffenschmidt, S., Kaiser, T., & Simon, M. (2012). Routine development of objectively derived search strategies. *Systematic Reviews*, *1*, 19.

Hopkins, W. G., Marshall, S. W., Batterham, A. M., & Hanin, J. (2009). Progressive statistics for studies in sports medicine and exercise science. *Medicine and Science in Sports and Exercise*, *41*(1), 3–13. https://doi.org/10.1249/MSS.0b013e31818cb278

Houghton, S., Hunter, S. C., Rosenberg, M., Wood, L., Zadow, C., Martin, K., & Shilton, T. (2015). Virtually impossible: Limiting Australian children and adolescents daily screen based media use. *BMC Public Health*, *15*, 5. https://doi.org/10.1186/1471-2458-15-5

Idris, G., Smith, C., Galland, B., Taylor, R., Robertson, C. J. & Farella, M. (2021). Home-based monitoring of eating in adolescents: A pilot study, *Nutrients*, *13*(12), 4354. https://doi.org/10.3390/nu13124354

Jepson, R. G., Harris, F. M., Platt, S., Tannahill, C. (2012). The effectiveness of interventions to change six health behaviours: a review of reviews. *BMC Public Health, 10,* 538. https://doi.org/10.1186/1471-2458-10-538

Jobarteh, M. L., McMroy, M. A., Lo, B., Konstantinos, K., Triantafyllidis, J. Q., Griffin, J. P., Sazonov, E., Sun, M., Jia, W., Baranowski, T., Anderson, A. K., Maitland, K. & Frost, G. (2023). Evaluation and acceptability, functionalirt, and validity of a passive image-based dietary intake assessment method in adults and children of Ghanian and Kenyan origin living in London, UK. *Nutrients*, 15, 4075. https://doi.org/10.3390/nu15184075

Jones-Jang, S. M., Heo, Y.-J., McKeever, R., Kim, J.-H., Moscowitz, L., & Moscowitz, D. (2020). Good news! Communication findings may be underestimated: Comparing effect sizes with self-reported and logged smartphone use data. *Journal of Computer-Mediated Communication: JCMC*, *25*(5), 346–363. https://doi.org/10.1093/jcmc/zmaa009

Jordan, A., Trentacoste, N., Henderson, V., Manganello, J., & Fishbein, M. (2007). Measuring the time teens spend with media: Challenges and opportunities. *Media Psychology*, *9*(1), 19–41. https://doi.org/10.1080/15213260709336801

Junco, R. (2013). Comparing actual and self-reported measures of Facebook use. *Computers in Human Behavior*, *29*(3), 626–631. https://doi.org/10.1016/j.chb.2012.11.007

Kamar, M., Evans, C., & Hugh-Jones, S. (2019). Factors influencing british adolescents' intake of whole grains: A pilot feasibility study using sensecam assisted interviews. *Nutrients*, *11*(11), 2620. https://doi.org/10.3390/nu11112620

Kaye, L. K., Orben, A., Ellis, D. A., Hunter, S. C., & Houghton, S. (2020). The conceptual and methodological mayhem of "screen time." *International Journal of Environmental Research and Public Health*, *17*(10), 3661. https://doi.org/10.3390/ijerph17103661

Kelly, P., Doherty, A., Berry, E., Hodges, S., Batterham, A., & Foster, C. (2011). Can we use digital life-log images to investigate active and sedentary travel behaviour? Results from a pilot study. *The International Journal of Behavioral Nutrition and Physical Activity*, *8*, 44.

Kelly, P., Doherty, A., Hamilton, A., Matthews, A., Batterham, A., Nelson, M., Foster, C., & Cowburn, G. (2012). Evaluating the feasibility of measuring travel to school using a wearable camera. *American Journal of Preventive Medicine*, *43*(5), 546–550. https://doi.org/10.1016/j.amepre.2012.07.027

Kelly, P., Doherty, A., Mizdrak, A., Marshall, S., Kerr, J., Legge, A., Godbole, S., Badland, H., Oliver, M., & Foster, C. (2014). High group level validity but high random error of a self-report travel diary, as assessed by wearable cameras. *Journal of Transport & Health*, *1*(3), 190–201. https://doi.org/10.1016/j.jth.2014.04.003

Kelly, P., Marshall, S., Badland, H., Kerr, J., Oliver, M., Doherty, A., & Foster, C. (2013). An ethical framework for automated, wearable cameras in health behavior research. *American Journal of Preventive Medicine*, *44*(3), 314–319. https://doi.org/10.1016/j.amepre.2012.11.006

Kerr, J., Marshall, S. J., Godbole, S., Chen, J., Legge, A., Doherty, A. R., Kelly, P., Oliver, M., Badland, H. M., & Foster, C. (2013). Using the SenseCam to improve classifications of sedentary behavior in free-living settings. *American Journal of Preventive Medicine*, *44*(3), 290–296. https://doi.org/10.1016/j.amepre.2012.11.004

Kimberlin, C. L., & Winterstein, A. G. (2008). Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy*, *65*(23), 2276–2284. https://doi.org/10.2146/ajhp070364

Kirkorian, H. L., & Anderson, D. R. (2008). Learning from educational media. In S. L. Calvert & B. J. Wilson (Eds.), *The handbook of children, media, and development* (pp. 188–213). Blackwell Publishing Ltd.

Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, *15*(2), 155–163. https://doi.org/10.1016/j.jcm.2016.02.012

Krizhevsky, A., Sutskever, I., & Hinton, G. (2015). ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84–90. https://doi.org/10.1145/3065386

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Latomme, J., Van Stappen, V., Cardon, G., Morgan, P., Lateva, M., Chakarova, N., Kivelä, J., Lindström, J., Androutsos, O., González-Gil, E., & Others. (2018). The association between children's and parents' co-tv viewing and their total screen time in six european countries: Cross-sectional data from the Feel4diabetes study. *International Journal of Environmental Research and Public Health*, *15*(11), 2599. https://doi.org/10.3390/ijerph15112599

Leask, C. F., Harvey, J. A., Skelton, D. A., & Chastin, S. F. (2015). Exploring the context of sedentary behaviour in older adults (what, where, why, when and with whom). *European Review of Aging and Physical Activity*, *12*, 4. https://doi.org/10.1186/s11556-015-0146-7

Lee, P. H., Tse, A. C. Y., Wu, C. S. T., Mak, Y. W., & Lee, U. (2021). Validation of self-reported smartphone usage against objectively-measured smartphone usage in Hong Kong Chinese adolescents and young adults. *Psychiatry Investigation*, *18*(2), 95–100. https://doi.org/10.30773/pi.2020.0197

Liau, A. K., Khoo, A., & Ang, P. H. (2008). Parental awareness and monitoring of adolescent internet use. *Current Psychology*, *27*, 217-233. https://doi.org/10.1007/s12144-008-9038-6

Liljequist, D., Elfving, B., & Skavberg Roaldsen, K. (2019). Intraclass correlation - A discussion and demonstration of basic features. *PloS One*, *14*(7), e0219854. https://doi.org/10.1371/journal.pone.0219854

Liu, J., Riesch, S., Tien, J., Lipman, T., Pinto-Martin, J., & O'Sullivan, A. (2022). Screen media overuse and associated physical, cognitive, and emotional/behavioral outcomes in children and adolescents: an integrative review. *Journal of Pediatric Health Care, 36*(2), 99-109. https://doi.org/10.1016/j.pedhc.2021.06.003

Liu, J., Tang, W., Chen, G., Lu, Y., Feng, C., & Tu, X. M. (2016). Correlation and agreement: Overview and clarification of competing concepts and measures. *Shanghai Archives of Psychiatry*, *28*(2), 115–120. https://doi.org/10.11919/j.issn.1002-0829.216045

Lowe, B. (2017). *Assessing the nature and extent of children's screen time* [Doctoral dissertation, University of Otago]. https://ourarchive.otago.ac.nz/handle/10523/7560

Lubans, D. R., Hesketh, K., Cliff, D. P., Barnett, L. M., Salmon, J., Dollman, J., Morgan, P. J., Hills, A. P., & Hardy, L. L. (2011). A systematic review of the validity and reliability of sedentary behaviour measures used with children and adolescents. *Obesity Reviews*, *12*(10), 781–799. https://doi.org/10.1111/j.1467-789X.2011.00896.x

Lund, L., Sølvhøj, I. N., Danielsen, D., & Andersen, S. (2021). Electronic media use and sleep in children and adolescents in western countries: A systematic review. *BMC Public Health*, *21*(1), 1598. https://doi.org/10.1186/s12889-021-11640-9

Lyden, K., Petruski, N., Staudenmayer, J., & Freedson, P. (2014). Direct observation is a valid criterion for estimating physical activity and sedentary behavior. *Journal of Physical Activity & Health*, *11*(4), 860–863. https://doi.org/10.1123/jpah.2012-0290

Maddison, R., Cartledge, S., Rogerson, M., Goedhart, N. S., Ragbir Singh, T., Neil, C., Phung, D., & Ball, K. (2019). Usefulness of wearable cameras as a tool to enhance chronic disease self-management: Scoping review. *JMIR MHealth and UHealth*, *7*(1), e10371. https://doi.org/10.2196/10371

Mahalingham, T., McEvoy, P. M., & Clarke, P. J. F. (2023). Assessing the validity of self-report social media use: Evidence of no relationship with objective smartphone use. *Computers in Human Behavior*, *140*, 107–567. https://doi.org/10.1016/j.chb.2022.107567

Mansournia, M. A., Waters, R., Nazemipour, M., Bland, M., & Altman, D. G. (2021). Bland-Altman methods for comparing methods of measurement and response to criticisms. *Global Epidemiology*, *3*, 100045. https://doi.org/10.1016/j.gloepi.2020.100045

McCambridge, J., Witton, J., & Elbourne, D. R. (2014). Systematic review of the Hawthorne effect: New concepts are needed to study research participation effects. *Journal of Clinical Epidemiology*, *67*(3), 267–277. https://doi.org/10.1016/j.jclinepi.2013.08.015

McHugh, M. L. (2012). Interrater reliability: The kappa statistic. *Biochemia Medica*, *22*(3), 276–282. https://www.ncbi.nlm.nih.gov/pubmed/23092060

McKerchar, C., Smith, M., Gage, R., Williman, J., Abel, G., Lacey, C., Ni Mhurchu, C., & Signal, L. (2020). Kids in a candy store: An objective analysis of children's interactions with food in convenience stores. *Nutrients*, *12*(7), 1–14. https://doi.org/10.3390/nu12072143

McKerchar, C., Smith, M., Stanley, J., Barr, M., Chambers, T., Abel, G., Lacey, C., Gage, R., Ni Mhurchu, C., & Signal, L. (2020). Food store environment examination – FoodSee: A new method to study the food store environment using wearable cameras. *Global Health Promotion*, *27*(3), 73–81.

McVeigh, J., Smith, A., Howie, E., & Straker, L. (2016). Trajectories of television watching from childhood to early adulthood and their association with body composition and mental health outcomes in young adults. *PloS one*, 11(4), e0152879. https://doi.org/10.1371/journal.pone.0152879

Menditto, A., Patriarca, M. & Magnusson, B. (2007). Understanding the meaning of acuracy, trueness and precision, *Accreditation and Quality Assurance, 12*, 45-47. https://doi.org/10.1007/s00769-006-0191-z

Mendoza, J. A., McLeod, J., Chen, T.-A., Nicklas, T. A., & Baranowski, T. (2013). Convergent validity of preschool children's television viewing measures among low-income Latino families: A cross-sectional study. *Childhood Obesity* , *9*(1), 29–34. https://doi.org/10.1089/chi.2012.0116

Mok, T. M., Cornish, F., & Tarr, J. (2015). Too much information: Visual research ethics in the age of wearable cameras. *Integrative Psychological & Behavioral Science*, *49*(2), 309–322. https://doi.org/10.1007/s12124-014-9289-8

Mokkink, L., Boers, M., van der Vleuten, C., Bouter, L., Alonso, J., Patrick, D., de Vet, H., & Terwee, C. (2020). COSMIN Risk of Bias tool to assess the quality of studies on reliability or measurement error of outcome measurement instruments: A Delphi study. *BMC Medical Research Methodology*, *20*(1), 293. https://doi.org/10.1186/s12874-020-01179-5

Mokkink, L., de Vet, H., Prinsen, C., Patrick, D., Alonso, J., Bouter, L., & Terwee, C. (2018). COSMIN risk of bias checklist for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*(5), 1171–1179. https://doi.org/10.1007/s11136-017-1765-4

Mokkink, L., Terwee, C., Knol, D., Stratford, P., Alonso, J., Patrick, D., Bouter, L., & de Vet, H. (2010). The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: A clarification of its content. *BMC Medical Research Methodology*, *10*, 22. https://doi.org/10.1186/1471-2288-10-22

Mokkink, L., Terwee, C., Patrick, D., Alonso, J., Stratford, P., Knol, D., Bouter, L., & de Vet, H. (2010). The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *Journal of Clinical Epidemiology*, *63*(7), 737–745. https://doi.org/10.1016/j.jclinepi.2010.02.006

Mokkink, L., Terwee, C., Stratford, P., Alonso, J., Patrick, D., Riphagen, I., Knol, D., Bouter, L., & de Vet, H. (2009). Evaluation of the methodological quality of systematic reviews of health status measurement instruments. *Quality of Life Research*, *18*(3), 313–333. https://doi.org/10.1007/s11136-009-9451-9

Mullan, K. (2018). Technology and children's screen-based activities in the UK: The story of the millennium so far. *Child Indicators Research*, *11*(6), 1781–1800. https://doi.org/10.1007/s12187-017-9509-0

Muppalla, S. K., Vuppalapati, S., Pulliahgaru, A. R., & Sreenivasulu, H. (2023). Effects of excessive screen time on child development: an updated review and strategies for management. *Cureus*, *15*(6). https://doi.org/10.7759/cureus.40608

Narayanan, A., Stewart, T., & Mackay, L. (2020). A dual-accelerometer system for detecting human movement in a free-living environment. *Medicine and Science in Sports and Exercise*, *52*(1), 252–258. https://doi.org/10.1249/MSS.0000000000002107

National Health and Medical Research Council, Australian Research Council, & Universities Australia. (2018). *National statement on ethical conduct in human research*. www.nhmrc.gov.au/guidelines/publications/e72

Nguyen, D. H., Marcu, G., Hayes, G. R., Truong, K. N., Scott, J., Langheinrich, M., & Roduner, C. (2009). Encountering SenseCam: Personal recording technologies in everyday life. *In Proceedings of the 11th International Conference on Ubiquious Computing (UbiComp '09).* Association for Computing Machinery. https://doi.org/10.1145/1620545.1620571

Odgers, C. L., & Jensen, M. R. (2020). Annual Research Review: Adolescent mental health in the digital age: Facts, fears, and future directions. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *61*(3), 336–348. https://doi.org/10.1111/jcpp.13190

O'Loughlin, G., Cullen, S. J., McGoldrick, A., O'Connor, S., Blain, R., O'Malley, S., & Warrington, G. D. (2013). Using a wearable camera to increase the accuracy of dietary analysis. *American Journal of Preventive Medicine*, *44*(3), 297–301. https://doi.org/10.1016/j.amepre.2012.11.007

Orben, A., & Przybylski, A. K. (2019). The association between adolescent well-being and digital technology use. *Nature Human Behaviour*, *3*(2), 173–182. https://doi.org/10.1038/s41562-018-0506-1

Otten, J. J., Littenberg, B., & Harvey-Berino, J. R. (2010). Relationship between self-report and an objective measure of television-viewing time in adults. *Obesity*, *18*(6), 1273–1275. https://doi.org/10.1038/oby.2009.371

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *International Journal of Surgery*, *372*, 105906. https://doi.org/10.1016/j.ijsu.2021.105906

Park, J. H., Moon, J. H., Kim, H. J., Kong, M., H. & Oh, Y. H. (2020). Sedentary lifestyle: Overview of updated evidence of potential health risks. (2020). *Korean Journal of Family Medicine, 41*(6), 365-373. https:// doi.org/10.4082/kjfm.20.0165

Parry, D. A., Davidson, B. I., Sewall, C. J. R., Fisher, J. T., Mieczkowski, H., & Quintana, D. S. (2021). A systematic review and meta-analysis of discrepancies between logged and self-reported digital media use. *Nature Human Behaviour*, *5*(11), 1535–1547. https://doi.org/10.1038/s41562-021-01117-5

Pearson, A. L., Bottomley, R., Chambers, T., Thornton, L., Stanley, J., Smith, M., Barr, M., & Signal, L. (2017). Measuring blue space visibility and ‘blue recreation’ in the everyday lives of children in a capital city. *International Journal of Environmental Research and Public Health*, *14*(6). https://doi.org/10.3390/ijerph14060563

Perez, O., Garza, T., Hindera, O., Beltran, A., Musaad, S., Dibbs, T., Singh, A., Chug, S., Sisson, A., Kumar Vadathya, A., Baranowski, T., & O’Connor, T. (2023). Validated assessment tools for screen media use: A systematic review. *PloS One*, *18*(4), e0283714. https://doi.org/10.1371/journal.pone.0283714

Prinsen, C., Mokkink, L., Bouter, L., Alonso, J., Patrick, D., de Vet, H., & Terwee, C. (2018). COSMIN guideline for systematic reviews of patient-reported outcome measures. *Quality of Life Research*, *27*(5), 1147–1157. https://doi.org/10.1007/s11136-018-1798-3

Przybylski, A. K., Orben, A., & Weinstein, N. (2020). How much is too much? Examining the relationship between digital screen engagement and psychosocial functioning in a confirmatory cohort study. *Journal of the American Academy of Child and Adolescent Psychiatry*, *59*(9), 1080–1088. https://doi.org/10.1016/j.jaac.2019.06.017

Przybylski, A. K., & Weinstein, N. (2017). A large-scale test of the goldilocks hypothesis: Quantifying the relations between digital-screen use and the mental well-being of adolescents. *Psychological Science*, *28*(2), 204–215. https://doi.org/10.1177/0956797616678438

R Core Team. (2023). *R: A Language and Environment for Statistical Computing* (4.3.1) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Raber, M., Patterson, M., Baranowski, T., Jia, W., & Sun, M. (2018). Utility of eButton images for identifying food preparation behaviors and meal-related tasks in adolescents. *Nutrition Journal*, *17*(1), 1–1. https://doi.org/10.1186/s12937-018-0341-2

Radesky, J. S., Weeks, H. M., Ball, R., Schaller, A., Yeo, S., Durnez, J., Tamayo-Rios, M., Epstein, M., Kirkorian, H., Coyne, S., & Barr, R. (2020). Young children’s use of smartphones and tablets. *Pediatrics*, *146*(1). https://doi.org/10.1542/peds.2019-3518

Reeves, B., Robinson, T., & Ram, N. (2020). Time for the Human Screenome Project. *Nature*, *577*(7790), 314–317. https://doi.org/10.1038/d41586-020-00032-5

Rideout, V. (2015). *The common sense census: Media use by tweens and teens*. https://apo.org.au/node/58360

Robinson, A., Hulme-Moir, S., Puloka, V., Smith, M., Stanley, J., & Signal, L. (2017). Housing as a determinant of Tongan children's health: Innovative methodology using wearable cameras. *International Journal of Environmental Research and Public Health*, *14*(10), 1170. https://doi.org/10.3390/ijerph14101170

Robinson, J. L., Winiewicz, D. D., Fuerch, J. H., Roemmich, J. N., & Epstein, L. H. (2006). Relationship between parental estimate and an objective measure of child television watching. *The International Journal of Behavioral Nutrition and Physical Activity*, *3*, 43. https://doi.org/10.1186/1479-5868-3-43

Rosen, L. D., Lim, A. F., Felt, J., Carrier, L. M., Cheever, N. A., Lara-Ruiz, J. M., Mendoza, J. S., & Rokkum, J. (2014). Media and technology use predicts ill-being among children, preteens and teenagers independent of the negative health impacts of exercise and eating habits. *Computers in human behavior, 35*, 364-375. https://doi.org/10.1016/j.chb.2014.01.036

Rosenberg, D., Godbole, S., Ellis, K., Di, C., Lacroix, A., Natarajan, L., & Kerr, J. (2017). Classifiers for accelerometer-measured behaviors in older women. *Medicine and Science in Sports and Exercise*, *49*(3), 610–616. https://doi.org/10.1249/MSS.0000000000001121

Rouse, M. (2014, May 30). *Digital Signage*. Techopedia. https://www.techopedia.com/definition/16302/digital-signage

Rouse, M. (2017, May 19). *Wearable Device*. Techopedia. https://www.techopedia.com/definition/31206/wearable-device

Sallis, J. F., & Saelens, B. E. (2000). Assessment of physical activity by self-report: Status, limitations, and future directions. *Research Quarterly for Exercise and Sport, 71*(2), 1-14. https://doi.org/10.1080/02701367.2000.11082780

Sanders, T. (n.d.). *Square Eyes Documentation*. Retrieved January 26, 2024, from https://motivation-and-behaviour.github.io/SquareEyes/coding_manual/

Sanders, T. (2023). *Faceblurring*. Australian Catholic University. https://github.com/Motivation-and-Behaviour/faceblurring

Sanders, T., Parker, P., Del Pozo-Cruz, B., Noetel, M., & Lonsdale, C. (2019). Type of screen time moderates effects on outcomes in 4013 children: Evidence from the Longitudinal Study of Australian Children. *The International Journal of Behavioral Nutrition and Physical Activity*, *16*(1), 117. https://doi.org/10.1186/s12966-019-0881-7

Santos, R. M. S., Ventura, S. D. A., Nogueira, Y. J. D. A., Mendes, C. G., Paula, J. J. D., 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*, 1-21. https://doi.org/10.1007/s41347-024-00398-7

Saunders, T. J., Prince, S. A., & Tremblay, M. S. (2011). Clustering of children's activity behaviour: The use of self-report versus direct measures. *The International Journal of Behavioral Nutrition and Physical Activity*, *8*, 48. https://doi.org/10.1186/1479-5868-8-48

Schober, P., Boer, C., & Schwarte, L. A. (2018). Correlation coefficients: Appropriate use and interpretation. *Anesthesia and Analgesia*, *126*(5), 1763–1768. https://doi.org/10.1213/ANE.0000000000002864

Scholtes, V. A., Terwee, C. B., & Poolman, R. W. (2011). What makes a measurement instrument valid and reliable? *Injury*, *42*(3), 236–240. https://doi.org/10.1016/j.injury.2010.11.042

Schwarz, N. (2007). Retrospective and concurrent self-reports: The rationale for real-time data capture. In *The Science of Real-Time Data Capture* (pp. 11–26). Oxford University PressNew York, NY. https://doi.org/10.1093/oso/9780195178715.003.0002

Schwarz, N., & Oyserman, D. (2001). Asking questions about behavior: Cognition, communication, and questionnaire construction. *The American Journal of Evaluation*, *22*(2), 127–160. https://doi.org/10.1177/109821400102200202

Schwebel, F. J., & Larimer, M. E. (2018). Using text message reminders in health care services: A narrative literature review. *Internet Interventions*, *13*, 82–104. https://doi.org/10.1016/j.invent.2018.06.002

Selfhout, M. H. W., Branje, S. J. T., Delsing, M., ter Bogt, T. F. M., Meeus, W. H. J. (2009). Different types of Internet use, depression, and social anxiety: The role of perceived friendship quality. *Journal of Adolescence*, *32*(3), 819-833. https://doi.org/10.1016/j.adolescence.2008.10.011

Sewall, C. J. R., Bear, T. M., Merranko, J., & Rosen, D. (2020). How psychosocial well-being and usage amount predict inaccuracies in retrospective estimates of digital technology use. *Mobile Media & Communication*, *8*(3), 379–399. https://doi.org/10.1177/2050157920902830

Shiffman, S., Stone, A. A., & Hufford, M. R. (2008). Ecological momentary assessment. *Annual Review of Clinical Psychology*, *4*, 1–32. https://www.ncbi.nlm.nih.gov/pubmed/18509902

Short, S. E. & Mollborn, S. (2015). Social determinants and health behaviors: Conceptual frames and empirical advances, *Current Opinion in Psychology, 5*, 78-84. https://doi.org/10.1016/j.copsyc.2015.05.002

Sigman, A. (2012). Time for a view on screen time. *Archives of Disease in Childhood*, *97*(11), 935–942. https://doi.org/10.1136/archdischild-2012-302196

Signal, L., Smith, M., Barr, M., Stanley, J., Chambers, T. J., Zhou, J., Duane, A., Jenkin, G., Pearson, A., Gurrin, C., Smeaton, A., Hoek, J., & Ni Mhurchu, C. (2017). Kids'Cam: An objective methodology to study the world in which children live. *American Journal of Preventive Medicine*, *53*(3), e89–e95. https://doi.org/10.1016/j.amepre.2017.02.016

Signal, L., Stanley, J., Smith, M., Barr, M., Chambers, T., Zhou, J., Duane, A., Gurrin, C., Smeaton, A., McKerchar, C., Pearson, A., Hoek, J., Jenkin, G., & Ni Mhurchu, C. (2017). Children's everyday exposure to food marketing: An objective analysis using wearable cameras. *The International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 137. https://doi.org/10.1186/s12966-017-0570-3

Signorell, A. (2023). *DescTools: Tools for Descriptive Statistics* (0.99.49) [R-package]. https://cran.r-project.org/package=DescTools

Silva, A. R., Pinho, M. S., Macedo, L., & Moulin, C. J. A. (2018). A critical review of the effects of wearable cameras on memory. *Neuropsychological Rehabilitation*, *28*(1), 117–141. https://doi.org/10.1080/09602011.2015.1128450

Sirard, J. R., & Pate, R. R. (2001). Physical activity assessment in children and adolescents. *Sports Medicine*, *31*(6), 439–454. https://doi.org/10.2165/00007256-200131060-00004

Smith, C., Galland, B. C., de Bruin, W. E., & Taylor, R. W. (2019). Feasibility of automated cameras to measure screen use in adolescents. *American Journal of Preventive Medicine*, *57*(3), 417–424. https://doi.org/10.1016/j.amepre.2019.04.012

Smith, M., Stanley, J., Signal, L., Barr, M., Chambers, T., Balina, A., Ni Mhurchu, C., & Wilson, N. (2019). Children's healthy and unhealthy beverage availability, purchase and consumption: A wearable camera study. *Appetite*, *133*, 240–251. https://doi.org/10.1016/j.appet.2018.11.004

Steckler, A., & McLeroy, K. R. (2008). The importance of external validity. *American Journal of Public Health*, *98*(1), 9–10. https://doi.org/10.2105/AJPH.2007.126847

Stiglic, N., & Viner, R. M. (2019). Effects of screentime on the health and well-being of children and adolescents: A systematic review of reviews. *BMJ Open*, *9*(1), e023191. https://doi.org/10.1136/bmjopen-2018-023191

Suchert, V., Hanewinkel, R., & Isensee, B. (2015). Sedentary behavior and indicators of mental health in school-aged children and adolescents: A systematic review. *Preventive Medicine*, *76*, 48–57. https://doi.org/10.1016/j.ypmed.2015.03.026

Sullivan, G. M. (2011). A primer on the validity of assessment instruments. *Journal of Graduate Medical Education*, *3*(2), 119–120. https://doi.org/10.4300/JGME-D-11-00075.1

Sweetser, P., Johnson, D., Ozdowska, A., & Wyeth, P. (2012). Active versus passive screen time for young children. *Australasian Journal of Early Childhood*, *37*(4), 94–98. https://doi.org/10.1177/183693911203700413

Tech Encyclopedia Index. (n.d.). PC Mag. Retrieved August 28, 2019, from https://www.pcmag.com/encyclopedia/term/53848/video-game-console

Terwee, C., Mokkink, L., Knol, D., Ostelo, R., Bouter, L., & de Vet, H. (2012). Rating the methodological quality in systematic reviews of studies on measurement properties: A scoring system for the COSMIN checklist. *Quality of Life Research*, *21*(4), 651–657. https://doi.org/10.1007/s11136-011-9960-1

Terwee, C., Prinsen, C., Chiarotto, A., Westerman, M., Patrick, D., Alonso, J., Bouter, L., de Vet, H., & Mokkink, L. (2018). COSMIN methodology for evaluating the content validity of patient-reported outcome measures: A Delphi study. *Quality of Life Research*, *27*(5), 1159–1170. https://doi.org/10.1007/s11136-018-1829-0

The Department of Health. (2019). *Australia's Physical Activity and Sedentary Behaviour Guidelines and the Australian 24-Hour Movement Guidelines*. https://www1.health.gov.au/internet/main/publishing.nsf/content/health-pubhlth-strateg-phys-act-guidelines

Thomas, G., Bennie, J. A., Crocker, K., Andriyani, F. D., Booker, B. & Biddle, S. J. H. (2022). Using wearable cameras to categorize the type and context of screen-based behaviours among adolescents: Observational study, *JMIRPediatrics and Parenting, 5*(1), e28208. https://doi.org/10.2196/28208

Thomas, G., Bennie, J. A., De Cocker, K., Castro, O., & Biddle, S. J. H. (2019). A descriptive epidemiology of screen-based devices by children and adolescents: A scoping review of 130 surveillance studies since 2000. *Child Indicators Research*, *13*, 935–950. https://doi.org/10.1007/s12187-019-09663-1

Tremblay, M. S., Aubert, S., Barnes, J. D., Saunders, T. J., Carson, V., Latimer-Cheung, A. E., Chastin, S. F. M., Altenburg, T. M., Chinapaw, M. J. M., & Others. (2017). Sedentary behavior research network (SBRN) - Terminology consensus project process and outcome. *The International Journal of Behavioral Nutrition and Physical Activity*, *14*, 75. https://doi.org/10.1186/s12966-017-0525-8

Tremblay, M. S., Leblanc, A. G., Janssen, I., Kho, M. E., Hicks, A., Murumets, K., Colley, R. C., & Duggan, M. (2011). Canadian sedentary behaviour guidelines for children and youth. *Applied Physiology, Nutrition, and Metabolism*, *36*(1), 59–64; 65–71. https://doi.org/10.1139/H11-012

Tremblay, M. S., LeBlanc, A. G., Kho, M. E., Saunders, T. J., Larouche, R., Colley, R. C., Goldfield, G., & Connor Gorber, S. (2011). Systematic review of sedentary behaviour and health indicators in school-aged children and youth. *The International Journal of*

Behavioral Nutrition and Physical Activity, *8*, 98. https://doi.org/10.1186/1479-5868-8-98

Twenge, J. M., Hisler, G. C., & Krizan, Z. (2019). Associations between screen time and sleep duration are primarily driven by portable electronic devices: Evidence from a population-based study of U.S. children ages 0-17. *Sleep Medicine*, *56*, 211–218. https://doi.org/10.1016/j.sleep.2018.11.009

van der Schuur, W. A., Baumgartner, S. E., Sumter, S. R., & Valkenburg, P. M. (2018). Media multitasking and sleep problems: A longitudinal study among adolescents. *Computers in Human Behavior*, *81*, 316–324. https://doi.org/10.1016/j.chb.2017.12.024

van der Schuur, W. A., Baumgartner, S. E., Sumter, S. R., & Valkenburg, P. M. (2015). The consequences of media multitasking for youth: A review. *Computers in Human Behavior, 53*, 204-215. https://doi.org/10.1016/j.chb.2015.06.035

VanderWeele, T. J. (2019). Principles of confounder selection. *European Journal of Epidemiology*, *34*, 211-219. https://doi.org/10.1007/s10654-019-00494-6

Vanden Abeele, M., Beullens, K., & Roe, K. (2013). Measuring mobile phone use: Gender, age and real usage level in relation to the accuracy and validity of self-reported mobile phone use. *Mobile Media & Communication*, *1*(2), 213–236. https://doi.org/10.1177/2050157913477095

Veatupu, L., Puloka, V., Smith, M., McKerchar, C., & Signal, L. (2019). Me’akai in Tonga: Exploring the nature and context of the food Tongan children eat in ha’apai using wearable cameras. *International Journal of Environmental Research and Public Health*, *16*(10), 1681. https://doi.org/10.3390/ijerph16101681

Verbeij, T., Pouwels, J. L., Beyens, I., & Valkenburg, P. M. (2021). The accuracy and validity of self-reported social media use measures among adolescents. *Computers in Human Behavior Reports*, *3*, 100090. https://doi.org/10.1016/j.chbr.2021.100090

Wade, N., Ortigara, J. M., Sullivan, R. M., Tomko, R. L., Breslin, F. J., Baker, F. C., Fuemmeler, B. F., Delrahim Howlett, K., Lisdahl, K. M., Marshall, A. T., Mason, M. J., Neale, M. C., Squeglia, L. M., Wolff-Hughes, D. L., Tapert, S. F., Bagot, K. S., & ABCD Novel Technologies Workgroup. (2021). Passive sensing of preteens' smartphone use: An adolescent brain cognitive development (ABCD) cohort. *JMIR Mental Health*, *8*. https://doi.org/10.2196/29426

Watkins, L., Aitken, R., Gage, R., Smith, M., Chambers, T., Barr, M., Stanley, J., & Signal, L. (2018). Capturing the commercial world of children: The feasibility of wearable cameras to assess marketing exposure. *The Journal of Consumer Affairs*, *53*(4), 1396–1420. https://doi.org/10.1111/joca.12234

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. https://ggplot2.tidyverse.org

Willis, J. (2007). *Foundations of qualitative research: Interpretive and critical approaches*. SAGE Publications Inc. https://doi.org/10.4135/9781452230108

Wilson, G., Jones, D., Schofield, P., & Martin, D. J. (2016). Experiences of using a wearable camera to record activity, participation and health-related behaviours: Qualitative reflections of using the Sensecam. *Digital Health*, *2*, 2055207616682628. https://doi.org/10.1177/2055207616682628

Wood, C. T., Skinner, A. C., Brown, J. D., Brown, C. L., Howard, J. B., Steiner, M. J., Perrin, A. J., Levine, C., Ravanbakht, S. N., & Perrin, E. M. (2019). Concordance of child and parent reports of children's screen media use. *Academic Pediatrics*, *19*(5), 529–533. https://doi.org/10.1016/j.acap.2019.04.002

Xiao, Q., Chaput, J.-P., Olds, T., Fogelholm, M., Hu, G., Lambert, E. V., Maher, C., Maia, J., Onywera, V., Sarmiento, O. L., Standage, M., Tremblay, M. S., Tudor-Locke, C., Katzmarzyk, P. T., Katzmarzyk, P. T., Church, T. S., Lambert, D. G., Barreira, T., Broyles, S., … Wiltz, D. (2020). Sleep characteristics and health-related quality of life in 9- to 11-year-old children from 12 countries. *Sleep Health*, *6*(1), 4–14. https://doi.org/10.1016/j.sleh.2019.09.006

Yu, M., & Baxter, J. (2016). *Australian children's screen time and participation in extracurricular activities*. Australian Institute of Family Studies. https://growingupinaustralia.gov.au/research-findings/annual-statistical-report-2015/australian-childrens-screen-time-and-participation-extracurricular

Zhou, Q., Wang, D., Mhurchu, C. N., Gurrin, C., Zhou, J., Cheng, Y., & Wang, H. (2019). The use of wearable cameras in assessing children's dietary intake and behaviours in China. *Appetite*, *139*, 1–7. https://doi.org/10.1016/j.appet.2019.03.032

Zhu, X., Grittiths, H., Xiao, Z., Ribeaud, D., Eisner, M., Yang, Y., & Murray, A. L. (2023). Trajectories of screen time across adolescence and their associations with adulthood mental health and behavioral outcomes. *Journal of Youth and Adolescence, 52*(7)*,* 1433-1447. https://doi.org/10.1007/s10964-023-01782-x

Appendices

Appendix A: Search Strategy

Search Strategy Development

A potential source of bias relates to the creation and implementation of search strategies to identify relevant studies. To minimize the chance of bias in the creation and implementation of this search strategy, this search strategy was empirically derived, and was based on the objective approach developed by Hausner and colleagues (2015, 2016, 2012)

In brief, this objective approach to the search strategy involves:

1. Identifying previous systematic reviews in the area of interest and extracting references from those systematic reviews with similar research questions. If similar systematic reviews are not identified, create a precise strategy and screen articles for relevant articles.
2. Randomly dividing the identified relevant articles into two reference sets: a validation set and a development set.
3. Uploading the development set into a text a frequency analyser to derive a highly sensitive search strategy from the word frequencies from the titles, abstracts, and subject keywords of the uploaded articles.
4. Checking whether all references from the development set can be identified by the search strategy and revise where needed.
5. Checking whether all references from the search strategy can be identified by the search strategy.

Generation of Test Set

First, I conducted a preliminary search of the Cochrane Library, Prospero, ACM Digital Library, PubMed, Google Scholar, PsycINFO, and CINAHL and used bidirectional searching to identify previous systematic reviews and scoping reviews related to wearable camera use and health behaviour research (*n*=7). The identified systematic reviews did not have similar research questions to our review, and I was only able to extract a small number of relevant articles included in those reviews for our test set. To create a larger test set, I conducted a precise search strategy using concepts such as "wearable cameras" or "SenseCam" or "Autographer" and "children" or "adolescents" in Google Scholar, PsycINFO, CINAHL and MEDLINE. In this search I screened 83 articles and identified 23 relevant articles. I then conducted bidirectional screening using the identified articles to identify an additional six articles.

All of the identified articles were randomly divided into two sets: a development set (*n*=14) and a validation set (*n*=15).

Development Set:

1. Chambers, T., Stanley, J., Pearson, A. L., Smith, M., Barr, M., Mhurchu, C. N., & Signal, L. (2019). Quantifying children's non-supermarket exposure to alcohol marketing via product packaging using wearable cameras. *Journal of Studies on Alcohol and Drugs*, *80*(2), 158–166. https://doi.org/10.15288/jsad.2019.80.158
2. Chambers, T., Stanley, J., Signal, L., Pearson, A. L., Smith, M., Barr, M., & Ni Mhurchu, C. (2018). Quantifying the nature and extent of children's real-time exposure to alcohol marketing in their everyday lives using wearable cameras: Children's exposure via a range of media in a range of key places. *Alcohol and Alcoholism*, *53*(5), 626–633. https://doi.org/10.1093/alcalc/agy053
3. Cowburn, G., Matthews, A., Doherty, A., Hamilton, A., Kelly, P., Williams, J., Foster, C., & Nelson, M. (2016). Exploring the opportunities for food and drink purchasing and consumption by teenagers during their journeys between home and school: A feasibility study using a novel method. *Public Health Nutrition*, *19(*1), 93–103. https://doi.org/10.1017/s1368980015000889

4. Gage, R., Barr, M., Stanley, J., Reeder, A., Mackay, C., Smith, M., Chambers, T., Leung, W., & Signal, L. (2018). Sun protection and shade availability in New Zealand's outdoor recreation spaces. *The New Zealand Medical Journal*, *131*(1484), 30–37. https://www.ncbi.nlm.nih.gov/pubmed/30359354
5. Kelly, P., Doherty, A. R., Hamilton, A., Matthews, A., Batterham, A. M., Nelson, M., Foster, C., & Cowburn, G. (2012). Evaluating the feasibility of measuring travel to school using a wearable camera. *American Journal of Preventive Medicine*, *43*(5), 546–550. https://doi.org/10.1016/j.amepre.2012.07.027
6. McKerchar, C., Smith, M., Stanley, J., Barr, M., Chambers, T., Abel, G., Lacey, C., Gage, R., Ni Mhurchu, C., & Signal, L. (2019). Food store environment examination--FoodSee: A new method to study the food store environment using wearable cameras. *Global Health Promotion*, *27*(3), 73-81. https://doi.org/10.1177/1757975919859575
7. Raber, M., Patterson, M., Jia, W., Sun, M., & Baranowski, T. (2018). Utility of eButton images for identifying food preparation behaviors and meal-related tasks in adolescents. *Nutrition Journal*, *17*(1), 32. https://doi.org/10.1186/s12937-018-0341-2
8. Robinson, A., Hulme-Moir, S., Puloka, V., Smith, M., Stanley, J., & Signal, L. (2017). Housing as a determinant of Tongan children's health: Innovative methodology using wearable cameras. *International Journal of Environmental Research and Public Health*, *14*(10), 1170. https://doi.org/10.3390/ijerph14101170
9. Schrempft, S., van Jaarsveld, C. H., & Fisher, A. (2017). Exploring the potential of a wearable camera to examine the early obesogenic home environment: Comparison of SenseCam images to the Home Environment Interview. *Journal of Medical Internet Research*, *19*(10), e332. https://doi.org/10.2196/jmir.7748

10. Signal, L. N., Stanley, J., Smith, M., Barr, M. B., Chambers, T. J., Zhou, J., Duane, A., Gurrin, C., Smeaton, A. F., McKerchar, C., Pearson, A. L., Hoek, J., Jenkin, G. L. S., & Ni Mhurchu, C. (2017). Children's everyday exposure to food marketing: an objective analysis using wearable cameras. *The International Journal of Behavioral Nutrition and Physical Activity*, *14*(1), 137. https://doi.org/10.1186/s12966-017-0570-3
11. Smith, C., Galland, B. C., de Bruin, W. E., & Taylor, R. W. (2019). Feasibility of automated cameras to measure screen use in adolescents. *American Journal of Preventive Medicine*, *57*(3), 417–424. https://doi.org/10.1016/j.amepre.2019.04.012
12. Smith, M., Chambers, T., Abbott, M., & Signal, L. (2019). High stakes: Children's exposure to gambling and gambling marketing using wearable cameras. *International Journal of Mental Health and Addiction*. https://doi.org/10.1007/s11469-019-00103-3
13. Veatupu, L., Puloka, V., Smith, M., McKerchar, C., & Signal, L. (2019). Me'akai in Tonga: Exploring the nature and context of the food Tongan children eat in Ha'apai using wearable cameras. *International Journal of Environmental Research and Public Health*, *16*(10), 1681. https://doi.org/10.3390/ijerph16101681
14. Watkins, L., Aitken, R., Gage, R., Smith, M. B., Chambers, T. J., Barr, M., Stanley, J., & Signal, L. N. (2018). Capturing the commercial world of children: The feasibility of wearable cameras to assess marketing exposure. *The Journal of Consumer Affairs*, *32*, 179. https://doi.org/10.1111/joca.12234

Validation Set:

1. Barr, M., Signal, L., Jenkin, G., & Smith, M. (2015). Capturing exposures: using automated cameras to document environmental determinants of obesity. *Health Promotion International*, *30*(1), 56–63. https://doi.org/10.1093/heapro/dau089

2. Beltran, A., Dadabhoy, H., Chen, T. A., Lin, C., Jia, W., Baranowski, J., Yan, G., Sun, M., & Baranowski, T. (2016). *Adapting the eButton to the abilities of children for diet assessment* [Paper]. Proceedings of Measuring Behavior 2016: 10th International Conference on Methods and Techniques in Behavioral Research. International Conference on Methods and Techniques in Behavioral Research, Dublin, Ireland, https://www.ncbi.nlm.nih.gov/pubmed/31742257
3. Beltran, A., Dadabhoy, H., Ryan, C., Dholakia, R., Jia, W., Baranowski, J., Sun, M., & Baranowski, T. (2018). Dietary Assessment with a wearable camera among children: Feasibility and intercoder reliability. *Journal of the Academy of Nutrition and Dietetics*, *118*(11), 2144–2153. https://doi.org/10.1016/j.jand.2018.05.013
4. Chambers, T., Pearson, A. L., Kawachi, I., Rzotkiewicz, Z., Stanley, J., Smith, M., Barr, M., Ni Mhurchu, C., & Signal, L. (2017). Kids in space: Measuring children's residential neighbourhoods and other destinations using activity space GPS and wearable camera data. *Social Science & Medicine*, *193*, 41–50. https://doi.org/10.1016/j.socscimed.2017.09.046
5. Chambers, T., Pearson, A. L., Kawachi, I., Stanley, J., Smith, M., Barr, M., Mhurchu, C. N., & Signal, L. (2018). Children's home and school neighbourhood exposure to alcohol marketing: Using wearable camera and GPS data to directly examine the link between retailer availability and visual exposure to marketing. *Health & Place*, *54*, 102–109. https://doi.org/10.1016/j.healthplace.2018.09.012
6. Chambers, T., Pearson, A. L., Stanley, J., Smith, M., Barr, M., Ni Mhurchu, C., & Signal, L. (2017). Children's exposure to alcohol marketing within supermarkets: An objective analysis using GPS technology and wearable cameras. *Health & Place*, *46*, 274–280. https://doi.org/10.1016/j.healthplace.2017.06.003

7. Everson, B., Mackintosh, K. A., McNarry, M. A., Todd, C., & Stratton, G. (2019). Can wearable cameras be used to validate school-aged children's lifestyle behaviours? *Children*, *6*(2), 20. https://doi.org/10.3390/children6020020
8. Gage, R., Leung, W., Stanley, J., Reeder, A., Mackay, C., Chambers, T., Smith, M., Barr, M., & Signal, L. (2019). Studying third-parties and environments: New Zealand sun-safety research. *Health Promotion International*, *34*(3), 440–446. https://doi.org/10.1093/heapro/dax094
9. Gurtner, M., Gage, R., Thomson, G., Jaine, R., Stanley, J., Smith, M., Barr, M., Chambers, T., & Signal, L. (2018). Are children smoke-free at home? Using wearable cameras to study children's exposure to smoking and smoking paraphernalia in private spaces. *Child: Care, Health and Development*, *34*, 113. https://doi.org/10.1111/cch.12631
10. Kamar, M., Evans, C., & Hugh-Jones, S. (2019). Factors influencing British adolescents' intake of whole grains: A pilot feasibility study using SenseCam assisted interviews. *Nutrients*, *11*(11), 2620. https://doi.org/10.3390/nu11112620
11. Lloyd, A., Gray, T., & Truong, S. (2018). Seeing what children see: Enhancing understanding of outdoor learning experiences through body-worn cameras. *Journal of Outdoor Recreation, Education, and Leadership*, *10*(1), 52–66. https://www.learntechlib.org/p/188378/
12. Pearson, A. L., Bottomley, R., Chambers, T., Thornton, L., Stanley, J., Smith, M., Barr, M., & Signal, L. (2017). Measuring blue space visibility and "blue recreation" in the everyday lives of children in a capital city. *International Journal of Environmental Research and Public Health*, *14*(6), 563. https://www.mdpi.com/1660-4601/14/6/563

13. Signal, L. N., Smith, M. B., Barr, M., Stanley, J., Chambers, T. J., Zhou, J., Duane, A., Jenkin, G. L. S., Pearson, A. L., Gurrin, C., Smeaton, A. F., Hoek, J., & Ni Mhurchu, C. (2017). Kids'Cam: An objective methodology to study the world in which children live. *American Journal of Preventive Medicine*, *53*(3), e89–e95. https://doi.org/10.1016/j.amepre.2017.02.016
14. Smith, M., Stanley, J., Signal, L., Barr, M., Chambers, T., Balina, A., Ni Mhurchu, C., & Wilson, N. (2019). Children's healthy and unhealthy beverage availability, purchase and consumption: A wearable camera study. *Appetite*, *133*, 240–251. https://doi.org/10.1016/j.appet.2018.11.004
15. Zhou, Q., Wang, D., Mhurchu, C. N., Gurrin, C., Zhou, J., Cheng, Y., & Wang, H. (2019). The use of wearable cameras in assessing children's dietary intake and behaviours in China. *Appetite*, *139*, 1–7. https://doi.org/10.1016/j.appet.2019.03.032

Development of Search Strategy

The development set was entered into a text analysis frequency software package (J. Clark et al., 2020). I developed the search strategy based on the frequency of emerging terms from the text analysis. Terms that were present in at least 20% of the references in the development set were selected to develop the search strategy. These terms included population terms (child*, adolescen*, participant*, school age* and school-age*) and camera terms (wearable, camera*, image*, captured, automatically, automated, sensecam, automated camera*, objective, captur* and innovative).

I adjusted the commonly occurring terms identified by the word frequency analysis relevant to the population and camera terms to suit database specific searches (e.g., adding in truncations) and created the first draft of our search strategy. Additionally, some of the most commonly occurring terms were excluded during the development of the search strategy as they were too broad (e.g., image, objective, innovative).

The first draft of the search strategy was able to identify 13 out of the 14 articles in the development set (see Table A1). I further modified the search strategy to include keywords (i.e., MesH terms and subject headings) and additional terms to ensure I captured all articles in the development set. This modified version of the search strategy was able to identify all of the articles in the development set. I then tested the modified version of the search strategy against the validated set and identified all of the articles in the validation set (Table A2).

Table A1

Results of the First Drafted Search Strategy Identifying the Development Set Articles

Database	Search string (title and abstract)	Results	Sensitivity
CINAHL	child* OR adolescen* OR "school age" OR school-age AND "wearable camera" OR sensecam OR "automated camera"	24	57%
PubMed	"wearable camera" OR sensecam OR "automated camera"	187	78%
PsycINFO	child* OR adolescen* OR "school age" OR school-age AND "wearable camera" OR sensecam OR "automated camera"	19	71%
ProQuest Central	child* OR adolescen* OR "school age" OR school-age AND "wearable camera" OR sensecam OR "automated camera"	54	57%
Overall		284	92%

Note. For the PubMed database search, I tried searching with the population concept however, additional non-relevant articles were included. Instead, this search was filtered by age group.

Table A2

Results of the Modified Search Strategy Identifying Development Set and Validation Set Articles

Database	Search string (title and abstract)	Results	Sensitivity D set	V set
PubMed	(((“child”[MeSH Terms] OR “population”[MeSH Terms]) OR “adolescen*"[MeSH Terms]) OR ((((''child*”[Title/Abstract] OR “adolescen*"[Title/Abstract]) OR ''teen*”[Title/Abstract]) OR “school age*"[Title/Abstract]) OR ''school age*”[Title/Abstract])) AND (((“wearable camer*"[Title/Abstract] OR ''automated camera*”[Title/Abstract]) OR “sensecam”[Title/Abstract]) OR “eButton”[Title/Abstract])	50	92%	86%
PsycINFO	MM “digital Images OR MM “Digital video” OR MM “cameras” OR MM “Wearable Devices” OR TI (“wearable camer*" OR ''automated camera'' OR sensecam OR eButton) OR AB (''wearable camer*” OR “automated camera” OR sensecam OR eButton) AND TI (child* OR adolescen* OR teen* OR “school age*” OR school-age* OR youth) OR AB (child* OR adolescen* OR teen* OR “school age*” OR school-age* OR youth)	324	71%	33%
Overall		374	100%	100%

Note. D set, development set; V set, validation set.

As the developed search strategy had a high level of sensitivity and precision, most of the database results were small, with the database results ranging from 19-324 articles identified. To ensure I identified relevant articles that may not have been represented in the development or validation set, I included additional terms commonly used in wearable camera research in adult populations into the search strategy to broaden the search. Additional terms in the search strategy were also adaptations from a previous search strategy from a scoping review on the use of wearable cameras for chronic disease self-management (Maddison et al., 2019). The final search strategy for each electronic database can be found below.

Electronic Database Search Strategy

PsycINFO

S1 (MJ “Child”)

S2 (MJ “Adolescen*”)

S3 TI/AB/KW child* OR “school age*” OR school-age* OR adolescen* OR youth* OR teen* OR preadolescen* OR boy* OR girl*

S4 S1 OR S2 OR S3

S5 (DE “Cameras”)

S6 TI/AB/KW camera* OR life-log* OR lifelog* OR eButton* OR SenseCam* OR GoPro* OR “smart glass*” OR “Google Glass*” OR autographer* OR “Narrative Clip*”

S7 S5 OR S6

S8 S4 AND S7

Scopus

S1 TITLE-ABS-KEY (child* OR “school age*” OR school-age* OR adolescen* OR youth* OR teen* OR preadolescen* OR boy* OR girl*)

S2 TITLE-ABS-KEY (camera* OR life-log* OR lifelog* OR eButton* OR SenseCam* OR GoPro* OR “smart glass*” OR “Google Glass*” OR autographer* OR “Narrative Clip*”)

S3 S1 AND S2

CINAHL Complete

S1 (MH ”Child”)

S2 (MH “Adolescence”)

S3 TI/AB child* OR “school age*” OR school-age* OR adolescen* OR youth* OR teen* OR preadolescen* OR boy* OR girl*

S4 S1 OR S2 OR S3

S5 (MH “Smart Glasses”)

S6 TI/AB camera* OR life-log* OR lifelog* OR eButton* OR SenseCam* OR GoPro* OR “smart glass*” OR “Google Glass*” OR autographer* OR “Narrative Clip*”

S7 S5 OR S6

S8 S4 AND S7

ProQuest Dissertations and Theses Global

S1 noft(child* OR “school age*” OR school-age* OR adolescen* OR youth* OR teen* OR preadolescen* OR boy* OR girl*)

S2 noft(camera* OR life-log* OR lifelog* OR eButton* OR SenseCam* OR GoPro* OR “smart glass*” OR “Google Glass*” OR autographer* OR “Narrative Clip*”)

S3 S1 AND S2

Web of Science Core Collection

S1 TS=(child* OR “school age*” OR school-age* OR adolescen* OR youth* OR teen* OR preadolescen* OR boy* OR girl*)

S2 TS=(camera* OR life-log* OR lifelog* OR eButton* OR SenseCam* OR GoPro* OR “smart glass*” OR “Google Glass*” OR autographer* OR “Narrative Clip*”)

S3 S1 AND S2

ACM Digital Library (ACM Guide to Computing Literature)

(Title:(child* OR “school age*” OR school-age* OR adolescen* OR youth* OR teen* OR preadolescen* OR boy* OR girl*) OR Abstract:(child* OR “school age*” OR school-age* OR adolescen* OR youth* OR teen* OR preadolescen* OR boy* OR girl*) OR Keyword:(child* OR “school age*” OR school-age* OR adolescen* OR youth* OR teen* OR preadolescen* OR boy* OR girl*)) AND (Title:(camera* OR life-log* OR lifelog* OR eButton* OR SenseCam* OR GoPro OR “Google Glass*” OR autographer* OR “Narrative Clip*”) OR Abstract:(camera* OR life-log* OR lifelog* OR eButton* OR SenseCam* OR GoPro OR “Google Glass*” OR autographer* OR “Narrative Clip*”) OR Keyword:(camera* OR life-log* OR lifelog* OR eButton* OR SenseCam* OR GoPro OR “Google Glass*” OR autographer* OR “Narrative Clip*”))

PubMed

S1 (”Child” Mesh)

S2 (“Adolescent” Mesh)

S3 TI/AB child* OR “school age*” OR school-age* OR adolescen* OR youth* OR teen* OR preadolescen* OR boy* OR girl*

S4 S1 OR S2 OR S3

S5 (“Smart Glasses” Mesh)

S6 TI/AB camera* OR life-log* OR lifelog* OR eButton* OR SenseCam* OR GoPro* OR “Google Glass*” OR autographer* OR “Narrative Clip*”

S7 S5 OR S6

S8 S4 AND S7

SPORTDiscus

S1 (SU "Children")

S2 (SU "Youth")

S3 (SU "Teenagers")

S4 (SU "Adolescen*")

S5 TI/AB/KW child* OR “school age*” OR school-age* OR adolescen* OR youth* OR teen* OR preadolescen* OR boy* OR girl*

S6 S1 OR S2 OR S3 OR S5

S7 (SU “Camera*”)

S8 TI/AB/KW camera* OR life-log* OR lifelog* OR eButton* OR SenseCam* OR GoPro* OR “smart glass*” OR “Google Glass*” OR autographer* OR “Narrative Clip*”

S9 S7 OR S8

S10 S6 AND S9

Appendix B: Excluded Full Text Articles

Table B1

List of excluded full-text articles and reason for exclusion

Citation	Reason for exclusion
Änggård E. (2015). Digital cameras: Agents in research with children. *Children's Geographies*, 13(1), 44209. https://doi.org/10.1080/14733285.2013.827871	Not a wearable camera
Agnihotri, S., Rovet, J., Cameron, D., Rasmussen, C., Ryan, J., & Keightley, M. (2013, November). Sensecam as an everyday memory rehabilitation tool for youth with fetal alcohol spectrum disorder [Paper Presentation]. Proceedings of the 4th International SenseCam & Pervasive Imaging Conference, Sandiego, USA. https://doi.org/https://doi.org/10.1080/14733285.2013.827871	Not an empirical paper
Akhter, P. (2015). Making things in their own way: A study of digital literacy practices in three multilingual households [Doctoral dissertation, University of Sheffield]. White Rose eTheses Online. https://etheses.whiterose.ac.uk/9578/	Not a wearable camera
Araya, R., & Hernández, J. (2016). Collective Unconscious Interaction Patterns in Classrooms. In N. Nguyen., L. Iliadis, Y. Manolopoulos & B. Trawinski (Eds.), Computational Collective Intelligence: Lecture Notes in Computer Science Vol. 9876 (pp. 333-342). Springer, Cham. https://doi.org/10.1007/978-3-319-45246-3_32	Not a health behaviour
Ardito, C., Bellucci, A., Desolda, G., Divitini, M., & Mora, S. (Ed.). (2016, June). Smart Ecosystems cReation by visual dEsign. Proceedings of the First International Workshop on Smart Ecosystems cReation by Visual dEsign co-located	Not a health behaviour

Citation	Reason for exclusion
with the International Working Conference on Advanced Visual Interfaces, New York, USA. https://doi.org/10.1145/2909132.2927473	
ASHAWIRE. (2017). What babies see may be able to predict their first words. ASHA Leader, 22(3). https://doi.org/10.1044/leader.RIB4.22032017.12	Wrong population
Aslin, R. N. (2009). How infants view natural scenes gathered from a head-mounted camera. *Optometry and Vision Science*, 86(6), 561-565. https://doi.org/10.1097/OPX.0b013e3181a76e96	Wrong population
Bambach, S., Crandall, D. J., & Yu, C. (2013, August). Understanding embodied visual attention in child-parent interaction [Paper Presentation]. 2013 IEEE 3rd Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL), Electronic Conference Proceedings. https://doi.org/10.1109/DevLrn.2013.6652555	Wrong population
Barr, M., Signal, L., Jenkin, G., & Smith, M. (2013, November). Using sensecam to capture children's exposure to food marketing: A feasibility study [Paper Presentation]. Proceedings of the 4th International SenseCam & Pervasive Imaging Conference, Sandiego, USA. https://doi.org/10.1145/2526667.2526675	Not a health behaviour
Barr, M., Signal, L., Jenkin, G., & Smith, M. (2015). Capturing exposures: Using automated cameras to document environmental determinants of obesity. *Health Promotion International*, 30(1), 56-63. https://doi.org/10.1093/heapro/dau089	Not a health behaviour
Bell, W., Colaiezzi, B. A., Prata, C. S., & Coates, J. C. (2017). Scaling up dietary data for decision-making in low-income countries: New technological frontiers. *Advances in Nutrition*, 8(6), 916-932. https://doi.org/10.3945/an.116.014308	Not an empirical paper

Citation	Reason for exclusion
Beltran, A., Dadabhoy, H., Ryan, C., Dholakia, R., Baranowski, J., Li, Y., Yan, G., Jia, W., Sun, M., & Baranowski, T. (2019). Reliability and validity of food portion size estimation from images using manual flexible digital virtual meshes. *Public Health Nutrition*, 22(7), 1153-1159. https://doi.org/10.1017/S1368980017004293	Wrong population [a]
Bird, J., Colliver, Y., & Edwards, S. (2014). The camera is not a methodology: Towards a framework for understanding young children's use of video cameras. *Early Child Development and Care*, 184(11), 1741-1756. https://doi.org/10.1080/03004430.2013.878711	Not a wearable camera
Borjon, J. I., Schroer, S. E., Bambach, S., Slone, L. K., Abney, D. H., Crandall, D. J., & Smith, L. B. (2018). A view of their own: Capturing the egocentric view of infants and toddlers with head-mounted cameras. *Journal of Visualized Experiments*, 140, e58445. https://doi.org/10.3791/58445	Wrong population
Boushey, C. J., Spoden, M., Zhu, F. M., Delp, E. J., & Kerr, D. A. (2017). New mobile methods for dietary assessment: Review of image-assisted and image-based dietary assessment methods. *Proceedings of the Nutrition Society*, 76(3), 283-294. https://doi.org/10.1017/S0029665116002913	Not an empirical paper
Bulungu, A. L. S., Palla, L., Priebe, J., Forsythe, L., Katic, P., Varley, G., Gal, B. D., Sarah, N., Nambooze, J., Wellard, K., & Ferguson, E. L. (2020). Validation of a life-logging wearable camera method and the 24-hour diet recall method for assessing maternal and child dietary diversity. *British Journal of Nutrition*, 125(11), 1299-1309. https://doi.org/10.1017/s0007114520003530	Wrong population
Burling, J. M. (2015). On computational techniques for exploring parent-infant dynamics during social interaction (Publication No. 3663741) [Doctoral dissertation, University of Houston]. ProQuest One Academic.	Wrong population

Citation	Reason for exclusion
Burris, A. (2017). A child's-eye view: An examination of point-of-view camera use in four informal education settings. Visitor Studies, 20(2), 218-237. https://doi.org/10.1080/10645578.2017.1404352	Not an empirical paper
Byfield, S. (2002). Snapshots of youth: The lives of late teens across the world. *Young Consumers*, 3(4), 15-20. https://doi.org/10.1108/17473610210813574	Not a wearable camera
Cain, R., & Lee, V. R. (2016, October). Measuring electrodermal activity to capture engagement in an afterschool maker program [Short Paper]. Proceedings of the 6th Annual Conference on Creativity and Fabrication in Education, Stanford, USA. https://doi.org/10.1145/3003397.3003409	Not a health behaviour
Calloway, E. E., Roberts-Gray, C., Ranjit, N., Sweitzer, S. J., McInnis, K. A., Romo-Palafox, M. J., & Briley, M. E. (2014). Method of assessing parent-child grocery store purchasing interactions using a micro-camcorder. *Appetite*, 83, 1-9. https://doi.org/10.1016/j.appet.2014.07.028	Wrong population
Carrasco-Ochoa, J. A., Martinez-Trinidad, J. F., Olvera-Lopez, J. A., & Salas, J. (Ed.). (2019). Pattern Recognition: 11th Mexican Conference, MCPR 2019, Querétaro, Mexico, June 26–29, 2019, Proceedings. Springer International Publishing. https://doi.org/10.1007/978-3-030-21077-9	Not a health behaviour
Caton, L. (2019). Becoming researcher: navigating a post-qualitative inquiry involving child participants and wearable action cameras [Doctoral dissertation, Manchester Metropolitan University]. https://e-space.mmu.ac.uk/id/eprint/622447	Not a health behaviour
Chambers, T., Pearson, A. L., Kawachi, I., Rzotkiewicz, Z., Stanley, J., Smith, M., Barr, M., Ni Mhurchu, C., & Signal, L. (2017). Kids in space: Measuring children's residential neighborhoods and other destinations using activity space gps and wearable camera data. *Social Science & Medicine*, 193, 41-50. https://doi.org/10.1016/j.socscimed.2017.09.046	Not a health behaviour [a]

Citation	Reason for exclusion
Chambers, T., Pearson, A. L., Kawachi, I., Stanley, J., Smith, M., Barr, M., Mhurchu, C. N., & Signal, L. (2018). Children’s home and school neighbourhood exposure to alcohol marketing: Using wearable camera and gps data to directly examine the link between retailer availability and visual exposure to marketing. *Health & Place,* 54, 102-109. https://doi.org/10.1016/j.healthplace.2018.09.012	Not a health behaviour [a]
Chambers, T., Pearson, A. L., Stanley, J., Smith, M., Barr, M., Mhurchu, C. N., & Signal, L. (2017). Children’s exposure to alcohol marketing within supermarkets: An objective analysis using gps technology and wearable cameras. *Health & Place*, 46, 274-280. https://doi.org/10.1016/j.healthplace.2017.06.003	Not a health behaviour [a]
Chambers, T., Stanley, J., Signal, L., Pearson, A. L., Smith, M., Barr, M., & Mhurchu, C. N. (2018). Quantifying the nature and extent of children’s real-time exposure to alcohol marketing in their everyday lives using wearable cameras: Children’s exposure via a range of media in a range of key places. *Alcohol and Alcoholism*, 53(5), 626-633. https://doi.org/10.1093/alcalc/agy053	Not a health behaviour [a]
Chambers, T., Stanley, J., Pearson, A. L., Smith, M., Barr, M., Mhurchu, C. N., & Signal, L. (2019). Quantifying children’s non-supermarket exposure to alcohol marketing via product packaging using wearable cameras. *Journal of Studies on Alcohol and Drugs*, 80(2), 158-166. https://doi.org/10.15288/jsad.2019.80.158	Not a health behaviour [a]
Chaparro-Moreno, L. J., Justice, L. M., Logan, J. A. R., Purtell, K. M., & Lin, T. (2019). The preschool classroom linguistic environment: Children’s first-person experiences. *PLoS ONE,* 14(8). https://doi.org/10.1371/journal.pone.0220227	Wrong population

Citation	Reason for exclusion
Chel-Hong, M. (2017, July). Automatic detection and labeling of self-stimulatory behavioral patterns in children with autism spectrum disorder [Short Paper]. 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), 279-282. https://doi.org/10.1109/embc.2017.8036816o	Not a health behaviour
Chen, G., Jia, W., Zhao, Y., Mao, Z. H., Lo, B., Anderson, A. K., Frost, G., Jobarteh, M. L., McCroy, M. A., Sazonov, E., Steiner-Asiedu, M., Ansong, R. S., Baranowski, T., Burke, L., & Sun, M. (2021). Food/Non-Food classification of real-life egocentric images in low-and middle-income countries based on image tagging features. *Frontiers in Artificial Intelligence*, 4, 28. https://doi.org/10.3389/frai.2021.644712	Wrong population [a]
Cook, T., & Hess, E. (2007). What the camera sees and from whose perspective: Fun methodologies for engaging children in enlightening adults. *Childhood: A Global Journal of Child Research*, 14(1), 29-45. https://doi.org/10.1177/0907568207068562	Not a wearable camera
Cortes, N., Caswell, S. V., Lincoln, A. E., Hepburn, L., Myer, G. D., Higgins, M., & Putukian, M. (2017). Video analysis verification of head impact events measured by wearable sensors. *American Journal of Sports Medicine*, 45(10), 2379-2387. https://doi.org/10.1177/0363546517706703	Not a wearable camera
Cotter, C. (2010). Evaluating the effects of camera perspective in video modelling for children with autism: Point of view versus scene modelling (Publication No. 3470401) [Doctoral dissertation, Western Michigan University]. ProQuest One Academic.	Duplicate study

Citation	Reason for exclusion
Cotter, C. (2010). Evaluating the effects of camera perspective in video modelling for children with autism: Point of view versus scene modelling (Publication No. 3470401) [Doctoral dissertation, Western Michigan University]. ProQuest One Academic.	Not a wearable camera
Cunningham, C., & Jones, M. (1996). Play through the eyes of children: Use of cameras to study after-school use of leisure time and leisure space by pre-adolescent children. *Society and Leisure*, 19(2), 341-361. https://doi.org/10.1080/07053436.1996.10715523	Not a wearable camera
Cunningham, C., & Jones, M. (1996). Play through the eyes of children: Use of cameras to study after-school use of leisure time and leisure space by pre-adolescent children. *Society and Leisure*, 19(2), 341-361. https://doi.org/10.1080/07053436.1996.10715523	Duplicate study
Daniels, J., Haber, N., Voss, C., Schwartz, J., Tamura, S., Fazel, A., Kline, A., Washington, P., Phillips, J., Winograd, T., Feinstein, C., & Wall, D. P. (2018). Feasibility testing of a wearable behavioral aid for social learning in children with autism. *Applied Clinical Informatics*, 9(1), 129-140. https://doi.org/10.1055/s-0038-1626727	Not a health behaviour
Daniels, J., Schwartz, J. N., Voss, C., Haber, N., Fazel, A., Kline, A., Washington, P., Feinstein, C., Winograd, T., & Wall, D. P. (2018). Exploratory study examining the at-home feasibility of a wearable tool for social-affective learning in children with autism. *npj Digital Medicine*, 1. https://doi.org/10.1038/s41746-018-0035-3	Not a health behaviour
Davidson, C., MacDonald, A., & Fenton, A. (2018). “These are my camera glasses”: Wearable digital video glasses for recording and examining young children’s social interactions. *Asia-Pacific Journal of Research in Early Childhood Education*, 12(3), 93-119. https://doi.org/10.17206/apjrece.2018.12.3.93	Not a health behaviour

Citation	Reason for exclusion
Downing, K. L., Janssen, X., & Reilly, J. J. (2019). Feasibility of wearable cameras to assess screen time and time spent restrained in children aged 3 to 5 years: A study protocol. *BMJ Open*, 9(5). https://doi.org/10.1136/bmjopen-2018-028265	Not an empirical paper
Drabman, R. S., & Thomas, M. H. (1974). Does media violence increase children's toleration of real-life aggression? *Developmental Psychology*, 10(3), 418-421. https://doi.org/10.1037/h0036439	Not a wearable camera
Drake-Brockman, T. F., Datta, A., & von Ungern-Sternberg, B. S. (2016). Patient monitoring with google glass: A pilot study of a novel monitoring technology. *Pediatric Anaesthesia*, 26(5), 539-46. https://doi.org/10.1111/pan.12879	Not a health behaviour
Edmunds, S., Rozga, A., Li, Y., Karp, E., Ibanez, L., Rehg, J., & Stone, W. (2017). Brief report: Using a point-of-view camera to measure eye gaze in young children with autism spectrum disorder during naturalistic social interactions: a pilot study. *Journal of Autism & Developmental Disorders*, 47, 898-904. https://doi.org/10.1007/s10803-016-3002-3	Wrong population
Eigner, B. (2014). A new method for analyzing early mother-child interactions: The early dyadic interactional code system. *Magyar Pszichologiai Szemle*, 69(1), 205-234. https://doi.org/10.1556/MPSzle.69.2014.1.11	Not a wearable camera
Farooq, M., & Sazonov, E. (2017). Segmentation and characterization of chewing bouts by monitoring temporalis muscle using smart glasses with piezoelectric sensor. IEEE Journal of Biomedical and Health Informatics, 21(6), 1495-1503. https://doi.org/10.1109/JBHI.2016.2640142	Wrong population
Finley, J. R., Brewer, W. F., & Benjamin, A. S. (2011). The effects of end-of-day picture review and a sensor based picture capture procedure on autobiographical memory using sensecam. *Memory*, 19(7), 796-807. https://doi.org/10.1080/09658211.2010.532807	Wrong population

Citation	Reason for exclusion
Franken, M., Bolzan Martins, M., Benites de Vargas, J. L., Deponti Brasil, F., & da Silva de Freitas, V. (2020). Relation between kinematic parameters and motor performance of children in swimming. *Revista Brasileira de Ciência e Movimento*, 28(2), 129-135.	Not a wearable camera
Ge, Z. C., Liu, R. J., Lu, R. H., Lin, C. F., Zhang, R. J., & Pan, Z. Y. (2013). A portable mind wave monitor headband applied in intellectual disabled children. *Applied Mechanics and Materials*, 336-338, 1563-1569. https://doi.org/10.4028/www.scientific.net/AMM.336-338.1563	Not a health behaviour
Gemming, L., Utter, J., & Mhurchu, C. N. (2015). Image-assisted dietary assessment: A systematic review of the evidence. *Journal of the Academy of Nutrition and Dietetics*, 115(1), 64-77. https://doi.org/10.1016/j.jand.2014.09.015	Not an empirical paper
Ghekiere, A., Van Cauwenberg, J., de Geus, B., Clarys, P., Cardon, G., Salmon, J., De Bourdeaudhuij, I., & Deforche, B. (2014). Critical environmental factors for transportation cycling in children: A qualitative study using bike-alongs. *Science & Sports*, 29, S18-S18. https://doi.org/10.1016/j.scispo.2014.08.031	Not a health behaviour
Goodyear, V. A., Casey, A., & Kirk, D. (2013). Slights, cameras, inaction: Using flip cameras in cooperative learning to explore girls' (dis)engagement in physical education. In L. Azzartio & D. Kirk. (Eds.), Pedagogies, Physical Culture, and Visual Methods (pp. 47-61). Routledge. https://doi.org/10.4324/9780203114698	Not a wearable camera
Goodyear, V., Casey, A., & Kirk, D. (2013). Using flip cameras in cooperative learning to explore girls' disengagement in physical education: The slights and the doings or non-doings caught on camera. Active and Health Magazine: Australian Council Health, Physical Education and Recreation, 20(3-4), 5-9.	Not a wearable camera

Citation	Reason for exclusion
Gray, A., Gray, T., & Truong, S. (2018). Seeing what children see: Enhancing understanding of outdoor learning experiences through body-worn cameras. *Journal of Outdoor Recreation, Education & Leadership*, 10(1), 52-66. https://doi.org/10.18666/JOREL-2018-V10-I1-8192	Not a health behaviour
Green, C. (2016). Sensory tours as a method for engaging children as active researchers: Exploring the use of wearable cameras in early childhood research. *International Journal of Early Childhood*, 48(3), 277-294. https://doi.org/10.1007/s13158-016-0173-1	Not a health behaviour
Gurtner, M., Gage, R., Thomson, G., Jaine, R., Stanley, J., Smith, M., Barr, M., Chambers, T., & Signal, L. (2018). Are children smoke-free at home? Using wearable cameras to study children's exposure to smoking and smoking paraphernalia in private spaces. *Child: Care, Health and Development*, 34, 113. https://doi.org/10.1111/cch.12631	Not a health behaviour
Hamann, C. J., & Peek-Asa, C. (2017). Beyond GPS: Improved study of bicycling exposure through added use of video data. *Journal of Transport and Health*, 4, 363-372. https://doi.org/10.1016/j.jth.2016.11.006	Not a health behaviour
Hamann, C., Peek-Asa, C., & McGehee, D. (2015). Helmet camera study of adult and child bicycling patterns and injury risk factors by gender. *Injury Prevention*, 21, A25-A25. https://doi.org/10.1136/injuryprev-2015-041654.69	Not a health behaviour
Hamann, C. J., & Peek-Asa, C. (2017). Examination of adult and child bicyclist safety-relevant events using naturalistic bicycling methodology. *Accident Analysis and Prevention*, 102, 44501. https://doi.org/10.1016/j.aap.2017.02.017	Not a health behaviour
Harms, T., Gershuny, J., Doherty, A., Thomas, E., Milton, K., & Foster, C. (2019). A validation study of the eurostat harmonised european time use study (HETUS) diary using wearable technology. *BMC Public Health*, 19, 455. https://doi.org/10.1186/s12889-019-6761-x	Wrong population

Citation	Reason for exclusion
Harper, K., Sands, C., Angarita Horowitz, D., Totman, M., Maitín, M., Rosado, J. S., Colon, J., & Alger, N. (2017). Food justice youth development: Using photovoice to study urban school food systems. *Local Environment*, 22(7), 791-808. https://doi.org/10.1080/13549839.2016.1274721	Not a wearable camera
Heidelberger, L., & Smith, C. (2015). The food environment through the camera lenses of 9- to 13-year-olds living in urban, low-income, midwestern households: A photovoice project. *Journal of Nutrition Education and Behavior*, 47(5), 437-445. https://doi.org/10.1016/j.jneb.2015.05.005	Active image capture
Hillier, A., Cole, B. L., Smith, T. E., Yancey, A. K., Williams, J. D., Grier, S. A., & McCarthy, W. J. (2009). Clustering of unhealthy outdoor advertisements around child-serving institutions: A comparison of three cities. *Health and Place*, 15(4), 935-945. https://doi.org/10.1016/j.healthplace.2009.02.014	Not a wearable camera
Hov, A. M., & Neegaard, H. (2020). The potential of chest mounted action cameras in early childhood education research. *Nordic Studies in Science Education*, 16(1), 44303. https://doi.org/10.5617/NORDINA.7049	Not a health behaviour
J, O. (2013). Technology Enhance Food Diaries. School Health Alert, 28(9), 7-7.	Duplicate study
Jacques, P. L. S., Conway, M. A., & Cabeza, R. (2011). Gender differences in autobiographical memory for everyday events: Retrieval elicited by sensecam images versus verbal cues. *Memory*, 19(7), 723-732. https://doi.org/10.1080/09658211.2010.516266	Wrong population
Jeunehomme, O., & D’Argembeau, A. (2019). The time to remember: Temporal compression and duration judgements in memory for real-life events. *Quarterly Journal of Experimental Psychology*, 72(4), 930-942. https://doi.org/10.1177/1747021818773082	Not a health behaviour

Citation	Reason for exclusion
Jia, W. Y., Chen, H. C., Yue, Y. F., Li, Z. X., Fernstrom, J., Bai, Y. C., Li, C. L., & Sun, M. G. (2014). Accuracy of food portion size estimation from digital pictures acquired by a chest-worn camera. *Public Health Nutrition*, 17(8), 1671-1681. https://doi.org/10.1017/s1368980013003236	Wrong population
Jia, W., Li, Y., Qu, R., Baranowski, T., Burke, L. E., Zhang, H., Bai, Y., Mancino, J. M., Xu, G., Mao, Z., & Sun, M. (2019). Automatic food detection in egocentric images using artificial intelligence technology. *Public Health Nutrition*, 22(7), 1168-1179. https://doi.org/10.1017/S1368980018000538	Wrong population [a]
Jobarteh, M. L., McCrory, M. A., Lo, B., Sun, M., Sazonov, E., Anderson, A. K., Jia, W., Maitland, K., Qiu, J., Steiner-Asiedu, M., Higgins, J. A., Baranowski, T., Olupot-Olupot, P., & Frost, G. (2020). Development and validation of an objective, passive dietary assessment method for estimating food and nutrient intake in households in low and middle-income countries: A study protocol. *Current Developments in Nutrition*, 4(2). https://doi.org/10.1093/cdn/nzaa020	Not an empirical paper
Jozkowski, A. C., Presgraves, E. A., Hodges, K. L., Wirth, E. N., Brandstetter, N. E., & Thayer, M. T. (2018). A novel rubric to evaluate wearable cameras for assessment of interrater reliability. *Occupation Participation and Health*, 38(2), 121-130. https://doi.org/10.1177/1539449217753349	Not a health behaviour
Kamar, M., Evans, C. E. L., & Hugh-Jones, S. (2016). Factors influencing adolescent whole grain intake: In-depth interviews with adolescents using sensecam technology. *Proceedings of the Nutrition Society*, 75, e200-E200. https://doi.org/10.1017/s0029665116002159	Duplicate study

Citation	Reason for exclusion
Kamar, M., Evans, C. E. L., & Hugh-Jones, S. (2016). Factors influencing adolescent whole grain intake: In-depth interviews with adolescents using sensecam technology. *Proceedings of the Nutrition Society*, 75, e200-E200. https://doi.org/10.1017/s0029665116002159	Duplicate study
Kamar, M. 2017. Factors influencing whole grain intake in UK adolescents: A theory-based study [Doctoral dissertation, University of Leeds].	Duplicate study
Kerr, J., Marshall, S. J., Godbole, S., Chen, J., Legge, A., Doherty, A. R., Kelly, P., Oliver, M., Badland, H. M., & Foster, C. (2013). Using the sensecam to improve classifications of sedentary behavior in free-living settings. *American Journal of Preventive Medicine*, 44(3), 290-296. https://doi.org/10.1016/j.amepre.2012.11.004	Wrong population
Kim, Y., Barry, V. W., & Kang, M. (2015). Validation of the actigraph gt3x and activpal accelerometers for the assessment of sedentary behavior. *Measurement in Physical Education & Exercise Science*, 19(3), 125-137. https://doi.org/10.1080/1091367X.2015.1054390	Wrong population
Lee, R., Skinner, A., Bornstein, M. H., Radford, A. N., Campbell, A., Graham, K., & Pearson, R. M. (2017). Through babies' eyes: Practical and theoretical considerations of using wearable technology to measure parent–infant behaviour from the mothers' and infants' view points. *Infant Behavior & Development*, 47, 62-71. https://doi.org/10.1016/j.infbeh.2017.02.006	Wrong population
Lee, V. R., Fischback, L., & Cain, R. (2019). A wearables-based approach to detect and identify momentary engagement in afterschool makerspace programs. *Contemporary Educational Psychology*, 59. https://doi.org/10.1016/j.cedpsych.2019.101789	Not a wearable camera

Citation	Reason for exclusion
Lefter, I., Rothkrantz, L., & Somhorst, M. (2012, June). Automated safety control by video cameras. Proceedings of the 13th International Conference on Computer Systems and Technologies, 298-305, https://doi.org/10.1145/2383276.2383320	Not a wearable camera
Leinonen, M. T., & Koskinen, L. (1997). Head-mounted video camera system in testing multihandicapped children with low vision. *Perception*, 26(1), 233–233. https://doi.org/10.1068/v970015	Not a health behaviour
Lesser, A. D., Luczynski, K. C., & Hood, S. A. (2019). Evaluating motion detection to score sleep disturbance for children: A translational approach to developing a measurement system. *Journal of Applied Behavior Analysis*, 52(2), 580-599. https://doi.org/10.1002/jaba.531	Not a wearable camera
Lin, C. Y., & Chang, Y. M. (2015). Interactive augmented reality using scratch 2. 0 to improve physical activities for children with developmental disabilities. *Research in Developmental Disabilities*, 37, 44409. https://doi.org/10.1016/j.ridd.2014.10.016	Not a wearable camera
Lindley, S. E., Randall, D., Sharrock, W., Glancy, M., Smyth, N., & Harper, R. (2009). Narrative, memory and practice: Tensions and choices in the use of a digital artefact [Short Paper]. Proceedings of the 23rd British HCI group annual conference on people and computers, Swinton, UK. https://doi.org/10.14236/ewic/hci2009.1	Not a health behaviour
Lindley, S. E., Glancy, M., Harper, R., Randall, D., & Smyth, N. (2011). Oh and how things just don't change, the more things stay the same': Reflections on sensecam images 18 months after capture. *International Journal of Human-Computer Studies*, 69(5), 311-323. https://doi.org/10.1016/j.ijhcs.2010.12.010	Not a health behaviour

Citation	Reason for exclusion
Liu, W., Barr, M., Pearson, A. L., Chambers, T., Pfeiffer, K. A., Smith, M., & Signal, L. (2020). Space-time analysis of unhealthy food advertising: New Zealand children's exposure and health policy options. *Health Promotion International*, 35(4), 812-820. https://doi.org/10.1093/heapro/daz083	Not a health behaviour [a]
Liu, Y., & Pomalaza-Ráez, C. (2010). On-chip body posture detection for medical care applications using low-cost cmos cameras. *Integrated Computer-Aided Engineering*, 17(1), 3-13. https://doi.org/10.3233/ICA-2010-0325.	Not an empirical paper
Lloyd, A. M. 2016. Place-based outdoor learning enriching curriculum: A case study in an Australian primary school [Doctoral dissertation, Western Sydney University]. Research Direct. http://researchdirect.uws.edu.au/islandora/object/uws%3A38701.	Not a health behaviour
Lloyd, R. S., Oliver, J. L., Myer, G. D., Croix, M. B. D., Wass, J., & Read, P. J. (2020). Comparison of drop jump and tuck jump knee joint kinematics in elite male youth soccer players: Implications for injury risk screening. *Journal of Sport Rehabilitation*, 29(6), 760-765. https://doi.org/10.1123/jsr.2019-0077	Not a wearable camera
Loveday, A., Sherar, L. B., Sanders, J. P., Sanderson, P. W., & Esliger, D. W. (2016). Novel technology to help understand the context of physical activity and sedentary behaviour. *Physiological Measurement*, 37(10), 1834-1851. https://doi.org/10.1088/0967-3334/37/10/1834	Wrong population
Lucas-Thompson, R. G., Seiter, N. S., & Lunkenheimer, E. S. (2020). Interparental conflict, attention to angry interpersonal interactions, and adolescent anxiety. *Family Relations: An Interdisciplinary Journal of Applied Family Studies*, 69(5), 1041-1054. https://doi.org/10.1111/fare.12505	Not a wearable camera

Citation	Reason for exclusion
Magrelli, S., Noris, B., Jermann, P., Ansermet, F., Hentsch, F., Nadel, J., & Billard, A. G. (2014). A wearable camera detects gaze peculiarities during social interactions in young children with pervasive developmental disorders. *IEEE Transactions on Autonomous Mental Development*, 6(4), 274-285. https://doi.org/10.1109/TAMD.2014.2327812	Not a health behaviour
Mair, A., Poirier, M., & Conway, M. A. (2017). Supporting older and younger adults' memory for recent everyday events: A prospective sampling study using sensecam. *Consciousness and Cognition*, 49, 190-202. https://doi.org/10.1016/j.concog.2017.02.008	Wrong population
McKay, L. L. 1997. Biomechanical parameters influencing fourth grade children's free throw shooting [Doctoral dissertation, Temple University].	Not a wearable camera
McKellin, W. H., Shahin, K., Hodgson, M., Jamieson, J., & Pichora-Fuller, K. (2007). Pragmatics of conversation and communication in noisy settings. *Journal of Pragmatics*, 39(12), 2159-2184. https://doi.org/10.1016/j.pragma.2006.11.012	Not a health behaviour
Meseck, K., Jankowska, M. M., Schipperijn, J., Natarajan, L., Godbole, S., Cson, J., Takemoto, M., Crist, K., & Kerr, J. (2016). Is missing geographic positioning system data in accelerometry studies a problem, and is imputation the solution? *Geospatial Health*, 11(2), 157-163. https://doi.org/10.4081/gh.2016.403	Wrong population
Michael, K. (2013). Social implications of wearable computing and augmented reality in everyday life. IEEE symposium on technology and society, ISTAS13.	Not a health behaviour
Michaelides, M., Rha, J., Dees, E. W., Baraas, R. C., Wagner-Schuman, M. L., Mollon, J. D., Dubis, A. M., Andersen, M. K., Rosenberg, T., Larsen, M., Moore, A. T., & Carroll, J. (2011). Integrity of the cone photoreceptor mosaic in	Not a wearable camera

Citation	Reason for exclusion
oligocone trichromacy. *Investigative Ophthalmology & Visual Science*, 52(7), 4757-64. https://doi.org/10.1167/iovs.10-6659	
Miesenberfer, K., Manduchi, R., Covarrubias Rodriguez, M., & Penaz, P. (Ed.). (2020). Computers helping People with special needs: 17th international conference on computers helping people with special needs, ICCHP 2020, Lecco, Italy, September 9–11, 2020, Proceedings, Part I. Springer International Publishing. https://doi.org/10.1007/978-3-030-58796-3	Not a health behaviour
Mills, K., Unsworth, L., Bellocchi, A., Park, J. Y., & Ritchie, S. (2014). Children's emotions and multimodal appraisal of places: Walking with the camera. *Australian Journal of Language and Literacy*, 37(3), 171-181.	Not a wearable camera
Milton, F., Muhlert, N., Butler, C. R., Smith, A., Benattayallah, A., & Zeman, A. Z. (2011). An fMRI study of long-term everyday memory using sensecam. *Memory*, 19(7), 733-744. https://doi.org/10.1080/09658211.2011.552185	Wrong population
Min, C. H. (2017). Automatic detection and labeling of self-stimulatory behavioral patterns in children with autism spectrum disorder [Paper Presentation]. Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS, 279-282. https://doi.org/10.1109/EMBC.2017.8036816	Not a wearable camera
Moodley, G., Christofides, N., Norris, S. A., Achia, T., & Hofman, K. J. (2015). Obesogenic environments: Access to and advertising of sugar-sweetened beverages in Soweto, South Africa, 2013. *Preventing Chronic Disease*, 12(10), 186. https://doi.org/10.5888/pcd12.140559. .	Not a wearable camera
Murphy, F. C., Barnard, P. J., Terry, K. A., Carthery-Goulart, M. T., & Holmes, E. A. (2011). Sensecam, imagery and bias in memory for wellbeing. *Memory*, 19(7), 768-77. https://doi.org/10.1080/09658211.2010.551130	Wrong population

Citation	Reason for exclusion
Nag, A., Haber, N., Voss, C., Tamura, S., Daniels, J., Ma, J., Bryan, C., Ramachandran, S., Schwartz, J., Winograd, T., Feinstein, C., Wall, D. P., & Chiang, B. (2020). Toward continuous social phenotyping: analyzing gaze patterns in an emotion recognition task for children with autism through wearable smart glasses. *Journal of Medical Internet Research*, 22(4), 44209. https://doi.org/10.2196/13810	Not a health behaviour
Nishida, J., Takatori, H., Sato, K., & Suzuki, K. (2015). Childhood: Wearable suit for augmented child experience [Poster Presentation]. Proceedings of the 2015 Virtual Reality International Conference, Laval, France. https://doi.org/10.1145/2787626.2792656	Not a wearable camera
Noris, B., Benmachiche, K., Meynet, J., Thiran, J. P., & Billard, A. G. (2007). Analysis of head-mounted wireless camera videos for early diagnosis of autism. *Advances in Soft Computing*, 45, 663-670. https://doi.org/10.1007/978-3-540-75175-5_83	Wrong population
Noris, B., Keller, J., & Billard, A. (2011). A wearable gaze tracking system for children in unconstrained environments. *Computer Vision and Image Understanding*, 115(4), 476–486. https://doi.org/10.1016/j.cviu.2010.11.013	Wrong population
O’Loughlin, G., Cullen, S. J., McGoldrick, A., O’Connor, S., Blain, R., O’Malley, S., & Warrington, G. D. (2013). Using a wearable camera to increase the accuracy of dietary analysis. *American Journal of Preventive Medicine*, 44(3), 297-301. https://doi.org/10.1016/j.amepre.2012.11.007	Wrong population
Omodei, M. M., & McLennan, J. (1994). Studying complex decision making in natural settings: Using a head-mounted video camera to study competitive orienteering. *Perceptual and Motor Skills*, 79(3), 1411-1425. https://doi.org/10.2466/pms.1994.79.3f.1411	Not a health behaviour

Citation	Reason for exclusion
Omodei, M. M., McLennan, J., & Whitford, P. (1998). Using a head-mounted video camera and two-stage replay to enhance orienteering performance. *International Journal of Sport Psychology*, 29(2), 115-131.	Not a health behaviour
O'Sullivan, G., McGuire, B., Roche, M., & Caes, L. (2020). Am I being watched? The role of researcher presence on toddlers' behaviour during 'everyday' pain experiences: A pilot study. *Psychology & Health*, 35(9), 1115-1133. https://doi.org/10.1080/08870446.2019.1707830	Wrong population
Pauly-Takacs, K., Moulin, C. A., & Estlin, E. J. (2011). Sensecam as a rehabilitation tool in a child with anterograde amnesia. *Memory*, 19(7), 705-712. https://doi.org/10.1080/09658211.2010.494046	Not a health behaviour
Piacentini, J., Himle, M. B., Chang, S., Baruch, D. E., Buzzella, B. A., Pearlman, A., & Woods, D. W. (2006). Reactivity of tic observation procedures to situation and setting. *Journal of Abnormal Child Psychology*, 34(5), 649-658. https://doi.org/10.1007/s10802-006-9048-5	Not a wearable camera
Piccardi, L., Noris, B., Barbey, O., Billard, A., Schiavone, G., Keller, F., & Von Hofsten, C. (2007, August). Wearcam: A head mounted wireless camera for monitoring gaze attention and for the diagnosis of developmental disorders in young children [Short Paper]. The 16th IEEE International Symposium on Robot and Human Interactive Communication, 594-598. https://doi.org/10.1109/ROMAN.2007.4415154	Wrong population
Potter, J., & Cowan, K. (2020). Playground as meaning-making space: Multimodal making and re-making of meaning in the (virtual) playground. *Global Studies of Childhood*, 10(3), 248-263. https://doi.org/10.1177/2043610620941527	Not a health behaviour
Proceedings of the 4th sensecam and pervasive imaging 2013 conference, in cooperation with ACM and SIGCHI, San Deigo, USA. Association for Computing Machinery, New York, NY, USA.	Not a health behaviour

Citation	Reason for exclusion
Proceedings of the 6th International Conference on the Foundations of Digital Games, Bordeaux, France. (2011). Association for Computing Machinery, New York, NY, USA.	Not a wearable camera
Raber, M., Baranowski, T., Crawford, K., Sharma, S. V., Schick, V., Markham, C., Jia, W., Sun, M., Steinman, E., & Chandra, J. (2020). The healthy cooking index: Nutrition optimizing home food preparation practices across multiple data collection methods. *Journal of the Academy of Nutrition and Dietetics*, 120(7), 1119-1132. https://doi.org/10.1016/j.jand.2020.01.008	Wrong population
RajKumar, A., Arora, C., Katz, B., & Kapila, V. (2019, April). Wearable smart glasses for assessment of eye-contact behavior in children with autism [Paper Presentation]. Design of Medical Devices Conference. https://doi.org/10.1115/DMD2019-3221	Not a health behaviour
Sahin, N. T., Keshav, N. U., Salisbury, J. P., & Vahabzadeh, A. (2018). Second version of google glass as a wearable socio-affective aid: Positive school desirability, high usability, and theoretical framework in a sample of children with autism. *Journal of Medical Internet Research*, 20(1). https://doi.org/10.2196/humanfactors.8785	Not a health behaviour
Sastre, L. R., Wright, L. D., & Haldeman, L. (2019). Use of digital photography with newcomer immigrant and refugee youth to examine behaviors and promote health. *Health Promotion Practice*, 20(5), 639-641. https://doi.org/10.1177/1524839919863465	Active image capture
Schiel, R., Kaps, A., & Bieber, G. (2012). Electronic health technology for the assessment of physical activity and eating habits in children and adolescents with overweight and obesity IDA. *Appetite*, 58(2), 432-437. https://doi.org/10.1016/j.appet.2011.11.021	Not a wearable camera

Citation	Reason for exclusion
Schiel, R., Kaps, A., Bieber, G., Kramer, G., Seebach, H., & Hoffmeyer, A. (2010). Identification of determinants for weight reduction in overweight and obese children and adolescents. *Journal of Telemedicine and Telecare*, 16(7), 368-373. https://doi.org/10.1258/jtt.2010.091005	Not a wearable camera
Schmitow, C., & Stenberg, G. (2015). What aspects of others' behaviors do infants attend to in live situations? *Infant Behavior and Development*, 40, 173-182. https://doi.org/10.1016/j.infbeh.2015.04.002	Wrong population
Schrempft, S. G. 2014. The role of the home environment in early weight trajectories [Doctoral dissertation, University College London]. ULC Discovery. https://discovery.ucl.ac.uk/id/eprint/1430286	Wrong population
Schrempft, S., van Jaarsveld, C., & Fisher, A. (2017). Using a wearable camera to examine the obesogenic quality of the home environment in early childhood. *Psychosomatic Medicine*, 79(4), A63-A64.	Wrong population
Schrempft, S., van Jaarsveld, C. H. M., & Fisher, A. (2017). Exploring the potential of a wearable camera to examine the early obesogenic home environment: comparison of sensecam images to the home environment interview. *Journal of Medical Internet Research*, 19(10), 44197. https://doi.org/10.2196/jmir.7748	Wrong population
Seamon, J. G., Moskowitz, T. N., Swan, A. E., Zhong, B., Golembeski, A., Liong, C., Narzikul, A. C., & Sosan, O. A. (2014). Sensecam reminiscence and action recall in memory-unimpaired people. *Memory*, 22(7), 861-866. https://doi.org/10.1080/09658211.2013.839711	Not a health behaviour
Sedighi, A., Ulman, S. M., & Nussbaum, M. A. (2018). Information presentation through a head-worn display ("smart glasses") has a smaller influence on the temporal structure of gait variability during dual-task gait compared to handheld displays (paper-based system and smartphone). *PLoS ONE*, 13(4). https://doi.org/10.1371/journal.pone.0195106	Not a health behaviour

Citation	Reason for exclusion
Shahbazi-Moghaddam, M. (2002). A new technique for assessing ball speed and impact force in volleyball. *Sports Biomechanics*, 1(2), 229-37. https://doi.org/10.1080/14763140208522799	Not a wearable camera
Shalash, W. M., Altamimi, S., Abdu, E., & Barom, A. (2018, August). No limit: A down syndrome children educational game [Short Paper]. IEEE Games, Entertainment, Media Conference (GEM), Galway, Ireland. https://doi.org/10.1109/GEM.2018.8516519	Not a wearable camera
Shaoqian, W., Cheung, S. C. S., & Luo, Y. (2016). Wearable privacy protection with visual bubble [Paper Presentation]. IEEE International Conference on Multimedia & Expo Workshops (ICMEW), Seattle, WA, USA. https://doi.org/10.1109/ICMEW.2016.7574712	Not a health behaviour
Sheats, J. L., Winter, S. J., Padilla-Romero, P., Goldman-Rosas, L., Grieco, L. A., & King, A. C. (2013, November). Comparison of passive versus active photo capture of built environment features by technology naïve Latinos using the sensecam and stanford healthy neighborhood discovery tool [Paper Presentation]. Proceedings of the 4th International SenseCam & Pervasive Imaging Conference, Sandiego, USA. https://doi.org/10.1145/2526667.2526669	Not a health behaviour
Signal, L. N., Smith, M. B., Barr, M., Stanley, J., Chambers, T. J., Zhou, J., Duane, A., Jenkin, G. L. S., Pearson, A. L., Gurrin, C., Smeaton, A. F., Hoek, J., & Mhurchu, C. N. (2017). Kids' cam: An objective methodology to study the world in which children live. *American Journal of Preventive Medicine*, 53(3), E89-E95. https://doi.org/10.1016/j.amepre.2017.02.016	Not a health behaviour
Signal, L. N., Stanley, J., Smith, M., Barr, M. B., Chambers, T. J., Zhou, J., Duane, A., Gurrin, C., Smeaton, A. F., McKerchar, C., Pearson, A. L., Hoek, J., Jenkin, G. L. S., & Mhurchu, C. Ni (2017). Children's everyday exposure to	Not a health behaviour [a]

Citation	Reason for exclusion
food marketing: An objective analysis using wearable cameras. *International Journal of Behavioral Nutrition & Physical Activity*, 14, 44501. https://doi.org/10.1186/s12966-017-0570-3	
Silva, A. R., Pinho, S., MacEdo, L. M., & Moulin, C. J. (2013). Benefits of sensecam review on neuropsychological test performance. *American Journal of Preventive Medicine*, 44(3), 302-307. https://doi.org/10.1016/j.amepre.2012.11.005	Not a health behaviour
Smith, L., Yu, C., Yoshida, H., & Fausey, C. M. (2015). Contributions of head-mounted cameras to studying the visual environments of infants and young children. *Journal of Cognition and Development*, 16(3), 407-419. https://doi.org/10.1080/15248372.2014.933430	Wrong population
Smith, L B., Yu, C., & Pereira, A. F. (2011). Not your mother's view: the dynamics of toddler visual experience. *Developmental Science*, 14(1), 44456. https://doi.org/10.1111/j.1467-7687.2009.00947.x	Wrong population
Smith, M., Chambers, T., Abbott, M., & Signal, L. (2020). High stakes: Children's exposure to gambling and gambling marketing using wearable cameras. *International Journal of Mental Health and Addiction*, 18(4), 1025-1047. https://doi.org/10.1007/s11469-019-00103-3	Not a health behaviour
Smith, M., Chambers, T., Abbott, M., & Signal, L. (2019). High stakes: Children's exposure to gambling and gambling marketing using wearable cameras. *International Journal of Mental Health and Addiction*, 18(4), 1025-1047. https://doi.org/10.1007/s11469-019-00103-3	Duplicate study
Soh, Z., Migita, R., Takahashi, K., Shimatani, K., Hayashi, H., Kurita, Y., & Tsuji, T. (2016). A motor behavioral evaluation method for children with developmental disorders during music therapy sessions: A pilot study. *Current Pediatric Research*, 20(44228), 103-117.	Not a wearable camera

Citation	Reason for exclusion
Sparrman, A. (2005). Video recording as interaction: Participant observation of children's everyday life. *Qualitative Research in Psychology*, 2(3), 241-255. https://doi.org/10.1191/1478088705qp041oa	Not a wearable camera
Spruijt-Metz, D., Wen, C. K. F., Bell, B. M., Intille, S., Huang, J. S., & Baranowski, T. (2018). Advances and controversies in diet and physical activity measurement in youth. *American Journal of Preventive Medicine*, 55(4), E81-E91. https://doi.org/10.1016/j.amepre.2018.06.012	Not an empirical paper
Stump, K. C. (2017). Children with autism wearing action cameras: changing parent/child interactions using point-of-view video modeling [Doctoral dissertation, University of Kansas]. KU ScholarWorks. http://hdl.handle.net/1808/25880	Duplicate study
Stump, K. C. (2017). Children with autism wearing action cameras: changing parent/child interactions using point-of-view video modeling [Doctoral dissertation, University of Kansas]. KU ScholarWorks. http://hdl.handle.net/1808/25880	Not a health behaviour
Suanda, S. H., Foster, S. B., Smith, L. B., & Yu, C. (2013, August). Attentional constraints and statistics in toddlers' word learning [Paper Presentation]. Third Joint International Conference on Development and Learning and Epigenetic Robotics (ICDL), Osaka, Japan. https://doi.org/10.1109/DevLrn.2013.6652542	Wrong population
Sugden, N. A., Mohamed-Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: Typical, daily exposure to faces documented from a first-person infant perspective. *Developmental Psychobiology*, 56(2), 249-261. https://doi.org/10.1002/dev.21183	Wrong population
Sumsion, J., Harrison, L., Press, F., McLeod, S., Goodfellow, J., & Bradley, B. (2011). Researching infants' experiences of early childhood education and care. Researching Young Children's Perspectives: Debating the Ethics and Dilemmas of Educational Research with Children, 113-127. https://doi.org/10.4324/9780203830437	Wrong population

Citation	Reason for exclusion
Svensson, Å., Waling, M., Bäcklund, C., & Larsson, C. (2012). Overweight and obese children’s ability to report energy intake using digital camera food records during a 2-year study. *Journal of Nutrition and Metabolism*, 2012, 8. https://doi.org/10.1155/2012/247389	Not a wearable camera
Tang, D., & Kubota, N. (2019). Lifelog generation based on informationally structured space. In H. Yu, J. Liu, L. Liu, Z. Ju, Y. Liu, D. Zhou. (Eds.), Intelligent Robotics and Applications: Lecture Notes in Computer Science Vol 11742 (pp.109-116). Springer, Cham. https://doi.org/10.1007/978-3-030-27535-8_11	Wrong population
Tapper, K., & Boulton, M J. (2002). Studying aggression in school children: the use of a wireless microphone and micro-video camera. *Aggressive Behavior*, 28(5), 356-365. https://doi.org/10.1002/ab.80009	Not a health behaviour
Trese, M. G., Khan, N. W., Branham, K., Conroy, E. B., & Moroi, S. E. (2016). Expansion of severely constricted visual field using google glass. *Ophthalmic Surg Lasers Imaging Retina*, 47(5), 486-9. https://doi.org/10.3928/23258160-20160419-15	Not a health behaviour
Tsuji, A., Sekine, S., Matsuda, S., Yamamoto, J., & Suzuki, K. 2020. Towards modeling of interpersonal proximity using head-mounted camera for children with asd. In K. Miesenberger, R. Manduchi, M. Covarrubias Rodriguez, P. Penaz. (Eds.), Computers Helping People with Special Needs: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics) Vol 12377 (pp. 104-111). Springer, Cham. https://doi.org/10.1007/978-3-030-58805-2_13	Not a health behaviour

Citation	Reason for exclusion
UbiComp/ISWC'18 adjunct (2018). Proceedings of the 2018 ACM international joint conference on pervasive and ubiquitous computing and proceedings of the 2018 ACM international symposium on wearable computers, Singapore, Singapore. Association for Computing Machinery, New York, NY, USA.	Not a wearable camera
Væver, M., Beebe, B., Kirk, O., Snidmann, N., Harder, S., & TroN, E. (2015). An automated approach for measuring infant head orientation in a face-to-face interaction. *Behavior Research Methods*, 47(2), 328-339. https://doi.org/10.3758/s13428-014-0487-6	Wrong population
Vartiainen, H., Leinonen, T., & Nissinen, S. (2019). Connected learning with media tools in kindergarten: An illustrative case. *Educational Media International*, 56(3), 233-249. https://doi.org/10.1080/09523987.2019.1669877	Not a wearable camera
Voss, C., Schwartz, J., Daniels, J., Kline, A., Haber, N., Washington, P., Tariq, Q., Robinson, T. N., Desai, M., Phillips, J. M., Feinstein, C., Winograd, T., & Wall, D. P. (2019). Effect of wearable digital intervention for improving socialization in children with autism spectrum disorder: A randomized clinical trial. *JAMA Pediatrics*, 173(5), 446-454. https://doi.org/10.1001/jamapediatrics.2019.0285	Not a health behaviour
Voss, C., Washington, P., Haber, N., Kline, A., Daniels, J., Fazel, A., De, T., McCarthy, B., Feinstein, C., Winograd, T., & Wall, D. (2016, September). Superpower glass: Delivering unobtrusive real-time social cues in wearable systems [Extended Abstract]. Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct, Heidelberg, Germany. https://doi.org/10.1145/2968219.2968310	Not a health behaviour
Ward, A. L., Galland, B. C., Haszard, J. J., Meredith-Jones, K., Morrison, S., McIntosh, D. R., Jackson, R., Beebe, D. W., Fangupo, L., Richards, R., Te Morenga, L., Smith, C., Elder, D. E., & Taylor, R. W. (2019). The effect of mild sleep	Not an empirical paper

Citation	Reason for exclusion
deprivation on diet and eating behaviour in children: Protocol for the daily rest, eating, and activity monitoring (dream) randomized cross-over trial. *BMC Public Health*, 19(1), 1347. https://doi.org/10.1186/s12889-019-7628-x	
Washington, P., Voss, C., Kline, A., Haber, N., Daniels, J., Fazel, A., De, T., Feinstein, C., Winograd, T., & Wall, D. (2017, Septermber). Superpowerglass: A wearable aid for the at-home therapy of children with autism [Short Paper]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, New York, USA. https://doi.org/10.1145/3130977	Not a health behaviour
Watkins, L., Aitken, R., Gage, R., Smith, M. B., Chambers, T. J., Barr, M., Stanley, J., & Signal, L. N. (2019). Capturing the commercial world of children: The feasibility of wearable cameras to assess marketing exposure. *Journal of Consumer Affairs*, 53(4), 1396-1420. https://doi.org/10.1111/joca.12234	Not a health behaviour
Watkins, L., Aitken, R., Gage, R., Smith, M. B., Chambers, T. J., Barr, M., Stanley, J., & Signal, L. N. (2019). Capturing the commercial world of children: The feasibility of wearable cameras to assess marketing exposure. *Journal of Consumer Affairs*, 53(4), 1396-1420. https://doi.org/10.1111/joca.12234	Duplicate Study
Watts, C. M., Moyer-Packenham, P. S., Tucker, S. I., Bullock, E. P., Shumway, J. F., Westenskow, A., Boyer-Thurgood, J., Anderson-Pence, K., Mahamane, S., & Jordan, K. (2016). An examination of children's learning progression shifts while using touch screen virtual manipulative mathematics apps. *Computers in Human Behavior*, 64, 814-828. https://doi.org/10.1016/j.chb.2016.07.029	Not a health behaviour

Citation	Reason for exclusion
Wettstein, A., Bryjová, J., Faßnacht, G., & Jakob, M. (2011). Aggression in environments of adolescent boys and girls. Four single case studies with camera-glasses. *Psychologie in Erziehung und Unterricht*, 58(4), 293-305. https://doi.org/10.2378/peu2011.art14d	Not a health behaviour
Wettstein, A. (2012). A conceptual frame model for the analysis of aggression in social interactions. *Journal of Social, Evolutionary, and Cultural Psychology*, 6(2), 141-157. https://doi.org/10.1037/h0099218	Not a health behaviour
Wettstein, A., & Jakob, M. (2010). Assessing aggressive adolescents' environments from their perspective by using camera-glasses: An innovative new method. *Journal of Aggression, Conflict and Peace Research*, 2(2), 23-32. https://doi.org/10.5042/jacpr.2010.0139	Not a health behaviour
Wettstein, A., & Scherzinger, M. (2015). Using camera-glasses for the assessment of aggressive behaviour among adolescents in residential correctional care: A small-scale study. *Journal of Aggression, Conflict and Peace Research*, 7(1), 33-46. https://doi.org/10.1108/JACPR-04-2014-0117	Not a health behaviour
Xu, T., Chen, Y., & Smith, L. (2011, August). It's the child's body: The role of toddler and parent in selecting toddler's visual experience [Short Paper]. 2011 IEEE International Conference on Development and Learning (ICDL), Frankfurt am Main, Germany. https://doi.org/10.1109/DEVLRN.2011.6037330	Wrong population
Yan, S. H., Zhou, X. L., Dang, D. H., Liang, X. Q., & Zhang, K. (2014). Kinematic analysis on gait of overweight and obese primary school children during level walking. *Journal of Medical Biomechanics*, 29(6), 548-553.	Not a wearable camera

Citation	Reason for exclusion
Ye, Z., Li, Y., Liu, Y., Bridges, C., Rozga, A., & Rehg, J. M. (2015). Detecting bids for eye contact using a wearable camera [Short Paper]. 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition (FG), Ljubljana, Slovenia. https://doi.org/10.1109/FG.2015.7163095	Wrong population
Yoon, H. J., RA, H., Basaran, C., Son, S. H., Park, T., & Ko, J. (2017). Fuzzy bin-based classification for detecting children's presence with 3d depth cameras. *ACM Transactions on Sensor Networks*, 13(3), 21. https://doi.org/10.1145/3079764	Not a wearable camera
Yoshida, H., & Smith, L. (2008). What's in view for toddlers? Using a head camera to study visual experience. *Infancy*, 13(3), 229-248. https://doi.org/10.1080/15250000802004437	Wrong population
Yost, N. (2003). Look what kindergarten children can do with technologies! [Short Paper]. Proceedings of the International Federation for Information Processing Working Group 3.5 Open Conference on Young Children and Learning Technologies, Sydney, Australia.	Not a wearable camera
Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, 125(2), 244-262. https://doi.org/10.1016/j.cognition.2012.06.016	Wrong population
Yu, C., Smith, L. B., & Pereira, A. F. (2008, August). Embodied solution: The world from a toddler's point of view [Short Paper]. IEEE 7th International Conference on Development and Learning (ICDL), Monterey, CA, USA. https://doi.org/10.1109/DEVLRN.2008.4640812	Wrong population
Yurovsky, D., Smith, L B., & Yu, C. (2013). Statistical word learning at scale: The baby's view is better. *Developmental Science*, 16(6), 959-966.	Wrong population

Citation	Reason for exclusion
Zhang, Y. C., & Rehg, J. M. (2018, June). Watching the tv watchers [Paper]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, New York, USA. https://doi.org/10.1145/3214291	Wrong population

Note: The exclusion reason ‘Not a health behavior’ refers to not meeting the health behavior definition.

[a] Articles reviewers considered as “near misses.

Appendix C: Quality Assessments of the Included Studies

Table C1

Reliability Quality Assessment using the COSMIN Risk of Bias Checklist

First author (year)	Did the professionals assign scores or determine values without knowledge of the scores or values of other repeated measurements in the same patients?	Were there any other important flaws in the design or statistical methods?	Continuous scores: was an ICC calculated?	Ordinal scores: was a weighted kappa calculated?	Dichotomous/ nominal scores: was Kappa calculated for each category against the other categories combined?	Overall worst score	Notes
Beltran (2016)	Very good	Very good	Adequate	N/A	N/A	Adequate	N/A
Beltran (2018)	Very good	Very good	Adequate	N/A	N/A	Adequate	N/A
Cowburn (2016)	Adequate	Very good	N/A	N/A	Very good	Adequate	N/A
Gage (2017)	Very good	Very good	Adequate	N/A	N/A	Adequate	N/A
Hänggi (2020)	Very good	Very good	N/A	N/A	Very good	Very good	N/A
Kelly (2012)	Very good	Very good	Very good	N/A	Very good	Very good	N/A
Raber (2018)	Very good	Doubtful	N/A	N/A	Very good	Doubtful	Only reports overall agreement.
Smith (2019)b	Adequate	Very good	N/A	N/A	Very good	Adequate	N/A

Note. ICC, intraclass correlation coefficient; N/A, not applicable.

Table C2

Measurement Error Quality Assessment using the COSMIN Risk of Bias Checklist

First author (year)	Did the professional(s) assign scores or determine values without knowledge of the scores or values of other repeated measurement(s) in the same patients?	Were there any other important flaws in the design or statistical methods?	Continuous scores: was the SEM, SDC, LoA or CV calculated?	Dichotomous/nominal/ ordinal scores: Was the percentage specific (e.g., positive and negative) agreement calculated?	Overall worst score	Notes
Beltran (2016)	Very good	Very good	N/A	Adequate	Adequate	N/A
Beltran (2018)	Very good	Very good	Very good	Adequate	Adequate	N/A
Gage (2017)	Very good	Very good	N/A	Very good	Very good	N/A
Gage (2019)	Very good	Doubtful	N/A	Adequate	Doubtful	Assessed on a test dataset and no indication if sample size was adequate.
Gage (2021)	Adequate	Inadequate	N/A	Adequate	Inadequate	Only reports overall % agreement threshold.
Kelly (2012)	Very good	Very good	Very good	N/A	Very good	N/A
McKerchar (2020)a	Adequate	Inadequate	N/A	Adequate	Inadequate	Only reports overall % agreement threshold. Assessed on a test set of images.
McKerchar (2020)b	Very good	Inadequate	N/A	Adequate	Inadequate	Did not report results.

First author (year)	Did the professional(s) assign scores or determine values without knowledge of the scores or values of other repeated measurement(s) in the same patients?	Were there any other important flaws in the design or statistical methods?	Continuous scores: was the SEM, SDC, LoA or CV calculated?	Dichotomous/nominal/ ordinal scores: Was the percentage specific (e.g., positive and negative) agreement calculated?	Overall worst score	Notes
Raber (2018)	Very good	Doubtful	N/A	Adequate	Doubtful	Only reports overall % agreement.
Robinson (2017)	Very good	Inadequate	N/A	Adequate	Inadequate	Only reports overall % agreement. Assessed on a test dataset and no indication if the sample size was adequate.
Smith (2019)a	Adequate	Inadequate	N/A	Adequate	Inadequate	Assessed on a test dataset of images and only reports overall % agreement threshold.
Smith (2019)b	Adequate	Very good	N/A	Adequate	Adequate	N/A
Veatupu (2019)	Adequate	Inadequate	N/A	Adequate	Inadequate	Assessed on a test dataset of images and only reports overall % agreement threshold.
Zhou (2019)	Inadequate	Very good	Very good	Adequate	Inadequate	N/A

Note. SEM, standard error of measurement; SDC, smallest detectable change; LoA, limits of agreement; CV, coefficient of variation; N/A, not applicable

Table C3

Convergent Validity Quality Assessment using the COSMIN Risk of Bias Checklist

First author (year)	Is it clear what the comparator instrument(s) measure(s)?	Were the measurement properties of the comparator instrument(s) sufficient?	Was the statistical method appropriate for the hypotheses to be tested?	Were there any other important flaws in the design or statistical methods?	Overall worst score	Notes
Everson (2019)	Very good	Very good	Inadequate	Very good	Inadequate	Inappropriate statistical method used.
Kelly (2012)	Very good	Very good	Adequate	Very good	Adequate	Mean scores not reported.
Zhou (2019)	Very good	Very good	Very good	Very good	Very good	N/A

Note. N/A, not applicable

Appendix D: Coding Protocol for Children's Screen Use

Image Coding Protocol for Automated Wearable Camera Research on Children's Screen Use

Device, Context and Content Code Manual

Bridget Booker

Australian Catholic University

2023

Introduction

This protocol is the coding protocol that identifies the device type, content and context of children and adolescent's screen use captured by automated wearable cameras. The purpose of this protocol is to guide coder training and ensure reliability and efficiency in image coding. Table D1 presents the definitions of key terms used throughout this protocol.

Table D1

Key term definitions

Term	Definition
Screen-based media device	A technology platform with a screen that displays visual content for the purpose of education, entertainment, or communication (e.g., a television, smartphone, or tablet).
Active screen-based media device	A screen-based media device has the appearance of being turned 'on' (i.e., has an active screen). Indicated by having a bright screen or visible content displayed on the screen of the device.
Inactive screen-based media device	A screen-based media device that has the appearance of being turned 'off' (i.e., a black screen) or having a screen that is not visible (e.g., a phone face down on the table).
Screen exposure	An event or episode where a person is in the presence of one or more active screen-based media devices, regardless of whether or not the person is consciously attending to the device.
Screen exposure episode	A series of images of a screen exposure grouped thematically.
Screen media	Content on any technology platform with a screen (Robinson et al., 2017).

Coder Training

Coders must complete coder training before coding any images. Coder training provides information on the ethical guidelines for handling image-based data, the codes and coding software and the image coding procedure. After coder training, coders should have a strong understanding of the ethical guidelines for handling image-based data, an understanding of the codebook and an understanding of how to annotate an image.

Coder training should follow the checklist below:

- Ethics
  - Read the Australian Catholic University Research Code of Conduct.
  - Read the Ethical Framework for Wearable Camera Research.
  - Read the Ethical Guidelines (see Ethical Guidelines).
- Code book familiarisation
  - Read codebook (see Codebook section).
  - Familiarise with popular screen media and screen-based media devices.
- Software familiarisation
- Image coding training
  - Read Image Coding section of this manual.
  - Practice coding images with a subset of images.

Ethics training

Coders will be required to handle image-based data. The images collected during data collection may contain sensitive images. While participants can delete unwanted or inappropriate images captured by the wearable camera, the images may contain identifying features of the participant, third parties and their environment. Coders should have a strong understanding of the ethical issues and guidelines for handling image-based data to protect the privacy, confidentiality and well-being of the participants and coders.

All coders must complete the checklist below:

- Read the Australian Catholic University Research Code of Conduct (Australian Catholic University, n.d.).

  Accessed here: https://policies.acu.edu.au/736355

- Read the Ethical Framework for Wearable Camera Research (Kelly et al., 2013). Accessed here: https://doi.org/10.1016/j.amepre.2012.11.006.
- Read the Ethical Guidelines (see Ethical Guidelines section).

Ethical Guidelines

The ethical guidelines have been informed by a previously developed ethical framework for wearable cameras in health behaviour research (Kelly et al., 2013), ethical guidelines used in previous wearable camera research (Signal, Smith, et al., 2017), and the National Statement on Ethical Conduct in Human Research (National Health and Medical Research Council et al., 2018).

The ethical guidelines include the following:

1. All coders must have a valid Working with Children Check when handling images from participants who are children.
2. All coders must be aware of the ethical implications of coding image-based data (read the ethical framework for wearable camera research and the Australian Catholic University Research Code of Conduct) and read and sign the ethical declaration statement.
3. All collected images must be stored securely and password-protected on the ACU secure network drive. Coders must not share their access passwords with anyone or give access to the ACU network drive to anyone outside of the research team.
4. Coders must not leave their computer containing unsecured data unattended.
5. Coders must protect the privacy and confidentiality of all participants. Coders must not disclose any information regarding the content of any images with anyone outside of the research team.
6. Coders are not to make copies of any images, email images or post copies to the internet.

7. Coders must report images that depict illegal activities to the principal investigator.
8. Coders must report images that are of concern to the researcher/coder (e.g., inappropriate images) to the principal investigator. If a coder is negatively impacted by an image they see, the coder should inform the principal investigator, who will assist the researcher/coder in gaining access to the appropriate services offered at the Australian Catholic University (e.g., counselling).
9. All images used in disseminated material must have any identifying features (e.g., faces, usernames, street names, business, and school names) blurred or blocked out of the image.

Coders must follow the ethical guidelines throughout coding and after the conclusion of coding.

Codebook Familiarisation

Prior to undertaking image coding training, coders should have a sound understanding of the code book and popular screen media and screen-based media devices used by the participant's population. Coders must have a good knowledge of the devices and popular screen media used by the participants to be able to accurately code the images. For example, coders must be able to recognise what a social media application interface looks like. Familiarisation of the code book and of popular devices and screen media may help the coding process as coders can identify certain features of the images and their corresponding codes easier. Familiarisation is particularly important for newer devices, such as tablets, e-readers and smartphones, and screen media, such as social media and communication applications and interfaces, as these devices and screen media are often updated and changed as trends and technology develop over time. Therefore, coders must be familiar with the common devices and screen media used in the year of data collection. Researchers can determine popular devices and screen media by reviewing the literature and reviewing the collected images to determine the most used screen media and devices.

To familiarise coders with the code book and popular devices and screen media, coders should read the code book and have examples (e.g., pictures or videos) of popular screen-based media devices and screen media provided for them to view.

Examples of common devices I have observed in the image-based data being used by children in 2023 include:

- Mobile devices such as smartphones and multipurpose devices
- Tablets
- Laptops
- Televisions

Examples of popular screen media I have observed in image-based data being consumed by children in 2023 include:

- Social media applications such as TikTok and Instagram.
- Instant and text message applications such as the iPhone text message application and Kids Messenger.
- Programme streaming applications such as YouTube and Netflix.
- Video games such as Fortnite, Minecraft and Roblox.

Software Familiarisation

Prior to undertaking image coding training, coders should have a sound understanding of the coding software to increase coding accuracy and efficiency. Coders will use an open-source image analyse software named *Timelapse: An Image Analyser for Camera Traps* (Greenberg, 2023). You can access this software by following this link: https://saul.cpsc.ucalgary.ca/timelapse/pmwiki.php?n=Main.HomePage

Software familiarisation includes the two steps below:

1. Download and install the software following the instructional guides given here: https://saul.cpsc.ucalgary.ca/timelapse/pmwiki.php?n=Main.UserGuide
2. Use the custom Screen Use Template on the Timelapse software to practice and understand how the functions work. Please note this Screen Use template has been designed for this study. See the Timelapse software instructional guides to create your own template.

Getting Started

1. Screen Use Template .tdb file defines your analysis codes. Copy this template file into the image folder you are analysing on the Network Drive. The template folder needs to be with the images you are coding.
2. Open Timelapse 2 and load your Screen Use Template .tdb in the appropriate folder.
3. Load your template. Select File > Load template, images, and videos... menu item. Then follow instructions to locate and load the above folder and its contents.
4. Make this window full screen. You will get the best results if your images are displayed as large as possible.
5. Analyse your images. Examine each image and enter the data appropriate for that image (see below).

Saving the data

Timelapse creates a.dsql database file in the same folder. As you record your data, it automatically saves that data into the database. You can export your data to a .csv file that can be opened with a spreadsheet. Once you finish coding a participant, select File > Export data in the current selection as a CVS file. Save the CVS file in the relevant participant's file under the current visit (i.e., Visit 1 or Visit 2). The software contains the variables listed in Table D2 for the current study.

Table D2

Explanation of variables in Screen Use Template for the validation study

Variable	Type	Explanation
File	Text	Image number
Relative Path	Text	NA
Date Time	Numerical	NA
Delete	Checkbox	NA
No Screen Exposure	Checkbox	Mark when there is NO screen exposure occurring
Location	Choice	
Uncodable	Categorical	Complete this cell only if the image is uncodable (see Uncodable definition). Uncodable options: Blurry Poor Lighting
Number of Device in image	Numerical	Add to count if a screen-based media device is not visible in the image; however, the coder is more than 50% certain that the participant is using a screen-based media device.
Device1	Choice	Type of screen-based media device: Television Computer Mobile device Tablet Handheld game console Smartwatch Unclassifiable
Device2	Choice	Type of screen-based media device: No second device Television Computer Mobile device Tablet Handheld game console Smartwatch Unclassifiable
Content1	Choice	Characteristics of screen media visible in the image: Unclassifiable Educational Recreational

Variable	Type	Explanation
		Social
Content2	Choice	Characteristics of screen media visible in the image: No second device Unclassifiable Educational Recreational Social
Location	Choice	Location of where the screen exposure is occurring: Bedroom Living room Kitchen/dining room Outside Other Unclassifiable
Social Setting	Choice	Who the participant is with during screen exposure: Alone Single Adult Single Child Only Adults Only children Mixed Ages Unclassifiable
Associated behaviour	Categorical	Associated behaviour occurring during screen exposure: No behaviour Eating Drinking Playing Multitask Other Unclassifiable
Comment	Text	Used during coding for additional information that the coder may want to include.

Image Coding

This section outlines the rules of coding the images and the protocol coders must follow to code an image. Before coding any images, coders should read the coding rules, coding procedure and practice coding images to ensure that they have a sound understanding of how to code an image reliably and efficiently.

Coding Rules

The coding rules have been informed by coding rules implemented in previous wearable camera research (Doherty, 2012; Lowe, 2017; Signal, 2017; Smith , 2019; Watkins, 2018).

Image coding order. Each image must be coded in chronological order as indicated by the image number.

Blocked image rule. A blocked image (e.g., the camera is blocked by a blanket) or a series of blocked images (e.g., multiple images blocked by a blanket) can be coded for screen exposure if the image before or after the blocked image shows the same screen-based media device in the same location and coders are confident that the participant was still in the screen exposure (see Figure D1 and D2 for an example). Blocked images that can be coded as screen exposure must be coded as no devices in the image with the same device, content and context coded as the before image.

Figure D1

Example of eighteen image rule

Figure D2

Example of eighteen image rule

Context Rule. Due to the camera's ability to only be able to capture what is in front of the participant, some images may not capture the complete context of the image (e.g., where the participant is located, who the participant is with, other behaviours the participant is undertaking during screen exposure). Coders can look at the whole screen exposure episode before coding the image to gain context of the image. Coders can code an image based on the context of the whole screen exposure if the coder was more than 50% certain that the context of the before or after images are the same context of the image being coded.

Ceiling images. Ceiling images are when the camera is pointed at the ceiling with no device visible in the image. Ceiling images most often occur when a participant is lying on their back or in a leaned back sitting position, which are also common body positions taken when using a screen-based media device. A ceiling image can be coded as screen exposure if an image before or after the ceiling image show the same screen-based media device in the same location and the participant does not move to another room or if the context remains the same and the coder is more than 50% certain that the participant is using a screen device (Smith et al., 2019). Ceiling images that can be coded as screen exposure must be coded as no devices in the image. If the device, content, or context is not visible in the image, it is to be coded the same as the image prior.

Coding Procedure

Each image should be coded in chronological order.

1. To gain an understanding of what is happening in the image, flick through the screen exposure episode.
2. Determine if the image is codable or contains a screen exposure. If the image is codable, continue coding the image. If the image is uncodable or does not include a

screen exposure, then code the image as uncodable or No screen exposure and move on to the next image.

3. Code the image for location
4. Code the image for number of devices
5. Code the image for devices1 and device2
6. Code the image for content1 and content2
7. Code the image for the social setting
8. Code the image for associated behaviours

Additional Information

- Only code images that have an active screen device present in the image (unless determined as a no device in an image following the coding rules). Do not code for inactive screen-based media devices.
- If something is coded for as *Other*, try to include as much information as possible in the notes cell of that row explaining why it was coded as *Other*.
- Flick through images prior to coding a screen exposure episode to gain an understanding of what is happening in the episode. Doing this will make it easier to code. An image may be coded based on what is occurring in the images prior and after the image.
- When coding, look at different cues in the image to determine the most appropriate code.
    - ***Cues for food*** may include seeing a plate on the dinner table or coffee table, seeing a fridge in the images leading up to the screen exposure.
    - ***Cues for screen exposure*** when a screen device is not present in the image include seeing other people in the room looking in the direction of the screen-

based media device or the participant sitting in the same position for a long period of time.

- ***Cues for screen exposure during ceiling images*** includes light on ceiling changing in each image or arms positions in a way that they are most likely holding a device.

No Screen Exposure and Uncodable Images

Uncodable images are images that cannot be confidently annotated due to image quality. Uncodable images are to be coded as uncodable, with the reason why the image or set of images are uncodable. Images should only be coded as uncodable if all aspects of the image or set of images cannot be confidently determined (i.e., if the context of the image is dark but you can still code the device and content, the image is not uncodable).

Reasons for an image to be classified as uncodable are listed below:

- ***Blurry*** - Any image or set of images where the image quality is so poor due to being blurred the coder is unable to confidently determine what is occurring in all aspects of the image.
- ***Poor Lighting*** - Any image or set of images where the image quality is so poor due to lighting the coder is unable to confidently determine what is occurring in all aspects of the image. Poor lighting can include images that are too dark or too overexposed to accurately determine anything.
- ***Blocked*** - Any image or set of images where the images are completely black or fully blocked by something and cannot be coded as having an active screen-based media device according to the coding rules. Includes completely blacked images.

Code Book

General Terms

Unclassifiable

An image is to be coded as ‘unclassifiable’ if the image is not able to be coded at the highest coding level possible. For example, if the coder cannot determine what type of device is being used then it would be coded as Device > Unclassifiable.

Other

An image is to be coded as ‘other’ in the appropriate coding level when the image does not fall into any of the other codes; however, is deemed classifiable by the coder. For example, if the coder recognises the device as a device not listed in the codes, then it would be coded as Device > Other.

Device

Device refers to a group of screen-based media devices that share common features. The device type dimension of the framework is made up of three categories (non-portable, portable, and wearable), with subcategories in each category. Some of the fields such as device are duplicated (e.g., device1 and device2). This allows for multiple devices to be coded. These should always be coded in descending order of attention. Listed below are the definitions for each category. Listed in Table D3 are the definitions for each device subcategory.

Non-portable device

A screen-based media device that cannot be easily carried and used in different places (e.g., a television or a desktop computer; Twenge et al., 2019).

Portable device

A screen-based media device that can be easily carried and used in different places (e.g., a smartphone, tablet or laptop computer; Cambridge University Press, n.d.-g).

Wearable device

A screen-based media device that is intended to be attached and worn on a person's body (e.g., smartwatch; Rouse, 2017).

Table D3

Device type and corresponding definitions

Type	Device	Definition
Non-portable	Television	A device shaped like a box or rectangle with a screen that receives electrical signals and changes them into moving images (Cambridge University Press, n.d.-l). Can stand alone or be mounted to a wall.
	Desktop Computer	A computer that fits on a desk but is not easily moved from place to place (Cambridge University Press, n.d.-b). Has a monitor, keyboard, mouse, and tower.
	Interactive Whiteboard	A large electronic screen linked to a computer. It is often used in classrooms to show information and can be written on or controlled by touching the screen with a finger or special pen (Cambridge University Press, n.d.-d).
	Projector	A device for showing films or images on a projection screen or other surfaces such as walls (Cambridge University Press, n.d.-h). Commonly used in movie theatres and schools.
	Digital Signage	A screen-based media device that is in the form of a small to large billboard composed of LCD, LED, or a similar display system (Rouse, 2014). Includes digital sign boards, interactive direction signboards, electronic menus, billboards, and similar display devices used for displaying visual information, promotional content, and advertisements in public areas (Rouse, 2014).
Portable	Laptop Computer	A computer that is battery operated and has an integrated screen. Indicated by an inbuilt keyboard.
	Mobile Device	A handheld device that can be used as a small computer, connect to the internet, and run applications. Includes smartphones, feature phones and multi-purpose devices (e.g., Apple iPod Touch).
	Tablet	A small, flat computer that is controlled by touching the screen with one's finger or a special pen (Cambridge University Press, n.d.-k). Does not require a keyboard or mouse. Includes e-readers.

Type	Device	Definition
	Handheld game console	Portable, self-contained devices that have a built-in screen, game controls and speakers (*Tech Encyclopedia Index*, n.d.).
Wearable	Smartwatch	A watch that has an electronic screen with many of the features of a smartphone or a computer (Cambridge University Press, n.d.-i). Does not include fitness trackers, such as a Fitbit.

Content

The content facet of the framework has been included to code the screen media that participants are exposed to. The content facet is made up of two sections: content type and content classification.

Content Type

Content type refers to the type of screen media the participant is exposed to during screen exposure. The content type facet is made up of three categories (passive screen media, interactive screen media and social media), with subcategories in each category. Listed below are the definitions for each category. Listed in Table D4 are the definitions for each content type subcategory.

Passive Screen Media. Screen media that requires no input or interaction during screen exposure, with the viewer only receiving screen-based information (e.g., watching a TV show or reading; Sweetser et al., 2012)

Interactive Screen Media. Screen media that requires real-time input from the child during the screen exposure (e.g., playing video games or browsing online; Sweetser et al., 2012).

Social Media. Websites, applications and computer programs that allow users to communicate and share information on the internet (e.g., Facebook, Twitter and Instagram; Cambridge University Press, n.d.-j).

Table D4

Content type subcategories corresponding definitions

Category	Subcategory	Definition
Passive screen media	Programme	Any form of TV Show, movie, or video. Includes online videos (e.g., videos being viewed on YouTube).
Interactive screen media	Internet	Includes all internet-based activities other than those for social media, gaming activity or watching online videos. Browsing is characterised by scrolling through the screen media and searching things up. Includes online shopping, using google and searching.
	Games	Includes when a participant is playing a video game, or watching another person play a video game.
	Creation	Creation screen media refers to visual content on screen-based media devices that has been created by the child (Rideout, 2015).
	Communication	Communication screen media refers to screen media that has the primary purpose of communicating with other people (Rideout, 2015).
Social media	Social media	Websites, applications, and computer programs that allow users to share information on the internet. Includes Facebook, Twitter, Instagram, Snapchat and blogs (Cambridge University Press, n.d.-j)

Content Classification

‘Content classification’ refers to if the screen media is educational, recreational, or social. The content classification facet is made up of three categories (educational, recreational, and social). Listed below are the definitions for each category.

Educational. Screen media created with the purpose to educate, inform and enlighten the viewer (Kirkorian & Anderson, 2008). Educational screen media includes screen media that appears to be educational. Includes creation applications, educational video games, programmes, and internet-based activities where it appears that the screen media is educational (e.g., watching an educational television programme or completing homework online).

Recreational. Screen media created with the purpose to entertain the viewer, with no intentional purpose to educate, inform or enlighten the viewer (Rideout, 2015). Recreational screen media includes screen media that appears to be recreational. Includes programmes, video games and internet-based activities where it appears that the screen media is recreational (e.g., most television programmes, playing video games or online shopping).

Social. Screen media created with the purpose to communicate with others (Rideout, 2015). Social screen media includes screen media that appears to be social. Includes social media and communication applications.

Context

The context dimension of the framework has been included to code the environment that the participant is in when the screen exposure occurs. The context dimension is made up of five facets: device attention, physical setting, social environment, social setting, and associated behaviours.

Device Attention

'Device attention' refers to the level of attention the participant appears to be giving to the screen-based media device. This section is made up of two categories: primary and secondary.

Primary. 'Primary' refers to when the screen-based media device appears to capture a large amount of the participant's attention. Only one screen-based media device per image can be coded as primary.

Secondary. 'Secondary' refers to when there are more than one screen-based media devices visible in the image and the screen-based media device appears to not capture full attention of the participant. For example, if a television was on in the background while the participant used a mobile phone, the television would be coded as the secondary device. The mobile phone would be coded as the primary device. Multiple screen-based media devices can be coded as secondary in the same image.

Location

'Location' refers to the place or type of surroundings where the participant's screen use occurred (Watkins et al., 2018). This section is made up of four categories (home, school, public, transport), with subcategories in each category. Listed below are the definitions for each category. Listed in Table D5 are the definitions for each location subcategory.

Home. Identified as the place where one lives (Cambridge University Press, n.d.-c). Includes all spaces within the home boundaries, and includes all private residence (i.e., friends and extended family residences Watkins et al., 2018).

School. An institute for teaching children and includes all spaces, buildings and grounds used within the school boundaries (Watkins et al., 2018).

Public. Indoor and outdoor spaces shared with other community members (Watkins et al., 2018).

Transport. The use of vehicles for getting from one place to another (Watkins et al., 2018).

Table D5

Location subcategories and corresponding definitions

Category	Subcategory	Definition
Home	Bedroom	A room or space used for sleeping in (Cambridge University Press, n.d.-a). Indicated by the presence of a piece of furniture for sleep or rest, typically a frame with a mattress.
	Living room	A room or space that is used for relaxing in and entertaining guests (Cambridge University Press, n.d.-f). Indicated by furniture such as a lounge or coffee table.
	Kitchen / Dining Room	A room or space where food is kept, prepared, cooked, and eaten (Cambridge University Press, n.d.-e). Indicated by the

Category	Subcategory	Definition
		presence of cooking appliances, such as a stove, refrigerator, microwave. Indicated by the presence of a flat surface such as a table, on which meals are served on.
	Outside	A space outdoors within the home boundaries e.g., backyard and front yard (e.g., backyard and front yard; Watkins et al., 2018). Indicated by grass and pavement within the home boundaries.
	Other	A room or space in the home (e.g., an office).
Public	Street	The public areas or roads of a town, suburb, or city. Includes roads, footpaths and outside of private properties, community venues and retails (Watkins et al., 2018).
	Community Venue	A building or room where members of the community can meet or use. Includes public libraries, recreation centres/community halls or churches (Watkins et al., 2018).
	Retail	A place/space where goods are sold to the public. Includes general product retailers (e.g., stores, supermarkets, service stations and shopping malls)
	Food Retail	A place where meals are prepared and sold (e.g., restaurants, cafes, and bakeries).
	Recreational Space	A space located outside rather than inside a building. A space where individuals participate or watch organised sports. Includes parks, beaches, rivers, walking tracks, sports stadiums, and sports grounds (Watkins et al., 2018).
Transport	Private Transport	Inside a truck, van, or car (Watkins et al., 2018).
	Public Transport	Inside a train, bus, ferry or aeroplane (Watkins et al., 2018).
School	Classroom	A room where groups of students are taught by a teacher. Indicated by the presence of desks and chairs.
	Library	A room or building containing books and computers for use or borrowing by the student attending the school.
	Computer Lab	A room or building containing multiple computers for use by the students attending the school and the sole purpose of this room is computer use. Indicated by the presence of multiple computers.

Category	Subcategory	Definition
	Playground	An area designed for children to play outside. Includes grass, concrete, and recreational equipment areas. Indicated by recreational equipment in the area, line markings on the ground and walls, other children playing and/or using sporting equipment.
	Other	A room, space or building located in the school gates and boundaries/fences e.g., school office or school hall.

Social Setting

'Social setting' refers to the social context of the child's screen use. The social setting facet is made up of two sections: social environment and social interaction.

Social environment. 'Social environment' refers to who the participant is with during screen exposure. This section is made up of five categories (alone, single adult, adults only, children only and mixed ages). Indicators include another person being visible in the image. Listed in Table D6 are the definitions for each social environment category.

Table D6

Social environment categories and definitions

Social environment	Definition
Alone	No one else is visible in the image.
Single adult	One person who appears to be over 18 years of age is visible in the image.
Single child	One person who appears to be under 18 years of age.
Adults only	People who appear to be over 18 years of age are visible in the image only (must be multiple people).
Children only	One or more people who appear to be under 18 years of age.
Mixed ages	Multiple people in the image who are children and adults.

Social interaction. ‘Social interaction refers’ to the exchange of information, interactivity, or lack of exchange between people. Social interaction is made up of five categories (co-playing, co-viewing, co-participating, background and none). Indicators include another person being visible in the image and the body language/movements of that person (Gemming, Doherty, et al., 2015). Listed in Table D7 are the definitions for each social interaction category.

Table D7

Social interaction categories and definitions

Social interaction	Definition
None	No one else is present in the image.
Co-participating	When the child and another person are actively using a screen-based media device together (e.g., a parent playing a video game with their child).
Co-viewing	When the child and another person are watching a screen-based media device together. Indicated by other people’s body positioning facing the device.

Associated Behaviours

‘Associated behaviours’ refers to co-occurring actions or tasks the child undertakes during screen exposure that is of interest to researchers. This section is made up of two categories (eating and multitask). Listed below are the definitions for each category.

Eating. The presence or observed consumption of food or drink in an image (Kerr et al., 2013).

Multitask. When a participant engages in another task or multiple tasks during screen exposure (Chinchanachokchai et al., 2015).

Screen time duration

Screen time refers to the duration of the screen exposure. Screen exposure is measured as the first image containing an active screen-based media device to the last image containing an active screen-based media device. Screen time duration is then calculated based on the number of images in the screen exposure multiplied by the camera epoch. For example, if a screen exposure consisted of 30 images, with a camera epoch of 2 seconds, the calculated screen time duration would be 1 minute.

Code Book Cheat Sheet

Device Attention

Category	Subcategory
Unclassifiable	
Attention >	Unclassifiable
	Primary
	Secondary
	Other

Device Type

Category	Subcategory
Unclassifiable	
Portable >	Unclassifiable
	Mobile Device (e.g., phone, iPod)
	Laptop Computer
	Tablet (e.g., iPad, e-reader)
	Handheld game console
	Other
Non-portable >	Unclassifiable
	Television
	Desktop Computer

Category	Subcategory
	Interactive Whiteboard
	Digital Signage
	Projector
	Other
Wearable >	Unclassifiable
	Smartwatch
	Other
Other	

Content Type

Category	Subcategory 1	Subcategory 2
Unclassifiable		
Passive Screen Media >	Unclassifiable	
	Programme	
	Other	
Interactive Screen Media>	Unclassifiable	
	Internet>	Unclassifiable
		Browse
		Other
	Game>	Unclassifiable
		Action
		Simulation
		Strategy/Puzzle
		Sports
		Exergame
		Educational
		Other
	Creation>	Unclassifiable

Category	Subcategory 1	Subcategory 2
		Writing Apps
		Art Apps
		Camera App
		Other
	Communication>	Unclassifiable
		Instant/Text Message
		Call
		Video Chat
		Email
		Other
	Other	
Social Media>	Unclassifiable	
	Facebook	
	Twitter	
	Instagram	
	Snapchat	
	TikTok	
	Blog (e.g., Tumblr)	
	Other	
Other		

Content Classification (e.g., educational/recreational/social)

Category
Unclassifiable
Educational
Recreational
Social
Other

Location

Category	Subcategory
Unclassifiable	
Home>	Unclassifiable
	Bedroom
	Living Room
	Kitchen / Dining Room
	Outside
	Other
School>	Unclassifiable
	Classroom
	Library
	Computer Lab
	Playground
	Other
Public>	Unclassifiable
	Street
	Community Venue
	Retail
	Food Retail
	Recreational Space
Transport>	Unclassifiable
	Private Transport
	Public Transport
	Other
Other	

Social Environment

Category
Unclassifiable
Alone
Single adult
Adults only
Children only
Other

Social Interaction (Type of Interaction)

Category
Unclassifiable
None
Co-participating
Co-viewing
Other

Associated Behaviour

Category
Unclassifiable
Food
Multitask
Other

Appendix E: Study 2 Participant Letters and Consent Forms

AUSTRALIAN CATHOLIC UNIVERSITY

PARTICIPANT INFORMATION LETTER

PROJECT TITLE: Validation of wearable cameras and GENEActiv wrist-worn accelerometer for assessing health behaviours among children.
APPLICATION NUMBER: 2017-317H
INVESTIGATOR: Dr. Taren Sanders
STUDENT RESEARCHER: Bridget Booker
STUDENT'S DEGREE: PhD

Dear Parent/Guardian,

You are invited to participate in a research project.

What is the project about?
The research project investigates the ways the children are spending their time in key behaviours that have important implications for their health and well-being. Specifically, we are investigating how well they sleep and how active they are and how each of these behaviours may be influencing the others. By participating in this study, you can gain an insight into your child's daily behaviour and contribute to a larger body of knowledge which seeks to understand how what we do influences our lives.

Who is undertaking the project?
This project is being conducted by Bridget Booker and will form the basis for the degree of Doctor of Philosophy at Australian Catholic University under the supervision of Dr. Taren Sanders and Professor Chris Lonsdale. Bridget Booker will be the primary contact during the study and will be doing the data collection. Dr. Sanders' research experience includes understanding physical activity behaviour and the ways to improve measurement of physical activity and his contribution includes expertise with the data analysis in the project, overseeing the quality of this project and ensuring the proper procedures are maintained throughout the project. Professor Lonsdale is overseeing the quality of this project and ensuring the proper procedures are maintained throughout the project.

Is my child eligible to participate in the study?
We are looking for children 8-11 years old to participate in the study. To participate in the study your child needs to have access to at least one screen-based device such as a television, smartphone, tablet or laptop. Prior to participation, we will ask you to complete a pre-screening questionnaire to determine eligibility.

What will my child be asked to do?
The project involves completing a number of different measures that will be carried out over a nine-day period. All assessments will be conducted by trained research staff who have current Working with Children Checks. We will conduct all of these measures at your home at a mutually convenient time. Children will be asked to wear an activity monitor, wear a small wearable camera during two observation sessions and have their height and weight measured. <u>You may participate in this study and not consent to all the measures.</u>

Activity Monitor

Your child will be asked to wear a GENEActiv physical activity monitor for seven days. This is a small, lightweight, wrist-worn device which measures physical activity and sleep. There is no power switch for the device as it is "always on" but it may be easily removed whenever necessary. When worn, the monitor records all movement by duration and intensity. The monitor can detect how much time is spent participating in activities of varying intensities (e.g. sitting, walking, running, sleeping). The monitor does not have GPS and cannot track locations. Whenever possible, we would like your child to wear the monitor for 7 days (including week and weekend days) for 24 hours each day. If you need to remove the device, we ask that you put it back on at the next possible moment. The monitor should be worn when your child goes to bed at night. The monitor is waterproof, so it can be worn in the shower or swimming pool.

Wearable Camera

Your child will be asked to wear a small, lightweight camera on a chest harness for two hours during an observation session in your home. We ask that you complete two observation sessions. The first observation session will be arranged at the start of the nine-day period. The second observation session will be arranged at the end of the nine-day period. During the observation sessions we will ask your child to complete set tasks that involve screen-based and non-screen based activities like watching television, playing video games or reading. For example, your child may be asked to watch something entertaining on a television for 9 minutes or play an entertaining video game on a tablet for 5 minutes. We will only ask your child to complete set tasks on screen devices that you have at your home. During the observation period a researcher will observe your child's interactions with the screen devices. The purpose of this session is to validate the wearable camera to measure children's screen use for future research studies. The camera captures a still image (i.e., not video or audio) every 2 seconds. The camera has an option to be turned off to ensure privacy and can be turned off and removed at any time. The memory card on the device will be encrypted to ensure that the images cannot be accessed by anyone other than the research team. You will have the opportunity to review and permanently delete any images prior to the research team collecting the data. The data will only be seen by members of the research team (i.e., we will not publish or share any of the images), who will follow strict protocols for data access and storage. The camera does not have GPS.

Height and Weight Measures

Your child will be asked to remove his/her shoes and any additional bulky heavy clothing (e.g. jacket) prior to having his/her height measured using a stadiometer, and weight measured using scales.

What will I be asked to do?

We will ask parents to complete a pre-screening questionnaire and a questionnaire about your household.

Pre-Screening Questionnaire

Prior to completing any measures, we will ask you to complete a short pre-screening questionnaire. The questionnaire will ask questions regarding your child's age, your child's access to screen devices, and if a parent/caretaker and your child is available for two direct observation sessions within the child's home. We will use this to ensure your child is eligible for the project.

Parent Questionnaire
We will ask you to complete a survey which asks questions about you, your household, and your child. Personal information for both you and your child will be gathered at this time, including name, gender, birthdate, family status, and socioeconomic status.

How much time will the project take?
The study will take place over a 9-day period. A questionnaire will be completed at the beginning that will require 15 minutes. The child will be asked to wear the activity tracker across one week and participate in two two-hour wearable camera observation sessions (one at the beginning of the 9-day period, one at the end of the 9-day period).

As a token of appreciation, upon completion of the study, we will offer a $25 Woolworths Essentials Gift Card. This gift card can be redeemed in-store at Woolworths Supermarkets, Big W, and Caltex Woolworths outlets and has no expiration date.

Are there any risks associated with participating in this project?
The risks associated with participating in this study are low. The questionnaires associated with the project do not cover any sensitive issues. The technology used to measure students' physical activity and sleep have been used extensively and are safe for children. There are many components associated with this study and the camera device may be uncomfortable to wear. Additionally, wearing around the camera may cause the child to feel discomfort with knowing their actions are being watched. To minimise this risk, we will strictly adhere to an existing ethical framework for the use of wearable cameras in behavioural research (Kelly et al., AJPM, 2013). You may participate in this study and not consent to all the measures.

What are the benefits of the research project?
Parents/guardians may benefit from this study by gaining a greater understanding of their child's movement behaviours. Specifically, how active they are and how well they sleep. These behaviours have important health implications for parents to consider. More broadly, the results may help to gain understanding about child behaviour and the outcomes of these behaviours, such as how screen time may affect sleep or how sleep affects physical activity. Additionally, part of the study is to validate the methods used in this study to measure screen time with the wearable camera which other researchers will be able to use for their own research.

Can I withdraw from the study?
Participation in this study is completely voluntary. You are not under any obligation to participate. If you agree to participate, you can withdraw from the study at any time without adverse consequences. Should you withdraw after some data has already been collected, we will ask for your consent to keep that data. Without your consent to retain the data, we will remove it completely.

Will anyone else know the results of the project?
The researchers will keep confidential any personal information obtained during the study. Once the data has been collected, de-identified using a coding system, and entered into an electronic data file, questionnaires and other data collection sheets will be destroyed. The electronic data files will be retained for at least 5 years, but no individual will be identifiable in published reports.

Will I be able to find out the results of the project?
At the end of the study, a summary report of your child's physical activity and sleep will be sent to you via email and a researcher will call to discuss and answer any questions about them. Scholarly reports, such

as journal articles, will also be published. All reports will be published in general terms and will not allow the identification of individual participants.

Who do I contact if I have questions about the project?
If you would like further information, please do not hesitate to contact Bridget Booker at bridget.booker@myacu.edu.au*. Thank you for considering this invitation.*

What if I have a complaint or any concerns?
The study has been reviewed by the Human Research Ethics Committee at Australian Catholic University (review number 2017-317H). If you have any complaints or concerns about the conduct of the project, you may write to the Manager of the Human Research Ethics Committee care of the Office of the Deputy Vice Chancellor (Research).

Manager, Ethics
c/o Office of the Deputy Vice Chancellor (Research)
Australian Catholic University
North Sydney Campus
PO Box 968
NORTH SYDNEY, NSW 2059
Ph.: 02 9739 2519
Fax: 02 9739 2870
Email: resethics.manager@acu.edu.au

Any complaint or concern will be treated in confidence and fully investigated. You will be informed of the outcome.

I want to participate! How do I sign up?
Thank you for your interest in the study. To participate, contact Bridget Booker (contact details below) to discuss the next steps.

Yours sincerely,

Bridget Booker, PhD Candidate
Institute for Positive Psychology and Education
Faculty of Health Sciences | Australian Catholic University
Mobile: (04) 32426708
Email: bridget.booker@myacu.edu.au

Dr. Taren Sanders

Research Fellow
Institute for Positive Psychology & Education
Australian Catholic University
T: *+61 2 9701 4704*
E: *Taren.Sanders@acu.edu.au*

PARENT/GUARDIAN CONSENT FORM

TITLE OF PROJECT: Validation of wearable cameras and GENEActiv wrist-worn accelerometer for assessing health behaviours among children.

APPLICATION NUMBER: (2017-317H)

PRINCIPAL INVESTIGATOR (or SUPERVISOR): Dr. Taren Sanders

STUDENT RESEARCHER (if applicable): Bridget Booker

I .... (*the parent/guardian*) have read (*or, where appropriate*, have had read to me) and understood the information provided in the Participant Information Letter. Any questions I have asked have been answered to my satisfaction. I realise that I can withdraw my consent at any time (without adverse consequences). I agree that research data collected for the study may be published or may be provided to other researchers in a form that does not identify my child in any way.

I consent to the following (cross any that do not apply):

- Completing a questionnaire about my household
- To my child participating in this study by:
  - Completing height and weight measurements
  - Wearing a wrist-worn accelerometer for one week
  - Completing two direct observation sessions (approx. 2 hours per session) in my child's home

SIGNATURE OF PARENT/GUARDIAN: ………………………………………………………………

DATE: ………………………………………………………………

NAME OF CHILD: ………………………………………………………………

EMAIL: ………………………………………………………………

MOBILE NUMBER: ………………………………………………………………

ACU
AUSTRALIAN CATHOLIC UNIVERSITY

CHILD PARTICIPANT INFORMATION LETTER

PROJECT TITLE: Validation of wearable cameras and GENEActiv wrist-worn accelerometer for assessing health behaviours among children.
APPLICATION NUMBER: 2017-317H
INVESTIGATOR: Dr. Taren Sanders
STUDENT RESEARCHER: Bridget Booker
STUDENT'S DEGREE: PhD

Dear Participant,

You are invited to take part in a research project.

What is the project about?
We are interested in understanding how children are spending their time. We want to know how long kids are sleeping and being active. These activities can be important for being healthy now and later in your life. By being a part of this study, you can learn more about how you spend your time and help us understand more about children's behaviour.

Who is in charge of the project?
This project is being led by Bridget Booker as part of her PhD at Australian Catholic University. Bridget is supervised by Dr. Taren Sanders and Professor Chris Lonsdale.

Can I be in the study?
We are looking for children 8-11 years old to be in the study. Your parents/guardians will fill out a survey that will tell us if you can participate or not.

What will I be asked to do?
During the project, we will ask you to do three different measures. These measures will happen over a nine-day period. We will conduct these measures at your home with you and your parents/guardians. The three measures are, (1) to wear an activity tracker, (2) to wear a small wearable camera, and (3) to measure your height and weight. You can choose to be part of this project and not do all three measures.

1. *Activity Tracker*

We will ask you to wear a GENEActiv physical activity tracker for seven days. This is a watch-like device which measures your movement. The device will tell us how much time you spend sitting, being active, or sleeping. The device is "always on" but you may remove it at any time. If you need to remove the device, we ask that you put it back on at the next possible moment. We would like you to wear the tracker all-day for 7 days. You should wear it when you go to bed at night. It is waterproof, so you can wear it in the shower or swimming pool. It does not have GPS and cannot track your location.

2. *Wearable Camera*

We will ask you to wear a small, lightweight camera for two hours. We ask that you participate in two two-hour sessions. One at the beginning of the nine-day period and one at the end of the nine-day

period. During the sessions we will ask you to complete set tasks that involve screen-based and non-screen based activities like watching television, playing video games or reading. We will only ask you to complete set tasks on screen devices that you have at your home. For example, we may ask you to watch something you find entertaining on a television for 9 minutes or play a video game you like on a tablet for 5 minutes. A researcher will also observe you during the two hours. The camera takes pictures, <u>not video or audio</u>, every 2 seconds. The purpose of this session is to help us test the camera as a measure of screen time. You can turn off or take off the camera at any time. our parents/guardians will check every picture to protect you and your privacy. You will not wear the camera at school. The camera does not have GPS.

3. <u>Height and Weight Measures</u>
We will ask to measure your height and weight. We will ask you to remove your shoes and any heavy clothing before we measure you. We will use a stadiometer (an upright tape measure) and a scale, similar to what you may have used at a doctor's office.

How much time will this take?
The study will take approximately five hours over nine days. You will wear the activity tracker on your wrist for one week. You will wear the camera for two-hours for two sessions (one at the beginning of the 9-days, one at the end of the 9-days).

Can anything bad happen to me in this project?
The devices used to measure your activity and sleep have often been used in research. They are safe for children. The camera may cause you to feel uncomfortable knowing your actions are being watched. We will allow you and your parents/guardians to view all the pictures and remove any of them for any reason. <u>You may participate in this study and not complete all the measures.</u>

What are the benefits of the project?
You can get a better idea of how active you are and how well you sleep. At the end of the study, we will send your parents/guardians the data we collect about you. Your parents/guardians may benefit knowing how active you are and how well you sleep. Overall, the results may help scientists understand child behaviour and the health outcomes of these behaviours.

Can I leave the study after I start?
Being in the study is a choice. You do not have to stay. You can stop at any time for any reason. Please talk to your parents/guardians if you want to quit.

Will I find out the results of the project?
At the end of the study, we will send your parents/guardians the data we collect about you. You can talk to them to learn about what we found.

Who do I talk to about the project?
Please talk to your parent/guardian if you have any further questions.

Yours sincerely,

Bridget Booker, PhD Candidate
Institute for Positive Psychology and Education
Faculty of Health Sciences | Australian Catholic University
Mobile: 0432 426 708
Email: bridget.booker@myacu.edu.au

Dr. Taren Sanders
Research Fellow
Institute for Positive Psychology & Education
Australian Catholic University
T: *+61 2 9701 4704*
E: *Taren.Sanders@acu.edu.au*

ASSENT OF PARTICIPANTS AGED UNDER 18 YEARS

I, ……………………………………………………… (*the participant aged under 18 years*), understand what this research project is about. My participation in the project has been explained to me. I agree to take part in this study.

- I agree to have my height and weight measured.
- I agree to wear an activity tracker for one week.
- I agree to wear an attachable camera for two-hours for two observation sessions.

  Feel free to cross out any of the above you do not want to do.

I know that I can withdraw from the study at any time without having to give a reason for my decision.

SIGNATURE OF YOUNG PERSON: ………………………………………………………………………………

DATE: …………………………………..

Appendix F: Study 2 Evidence of Ethics Approval

[2017-317H] - Ethics Extension Request Approved

RE Res Ethics EMAIL <Res.Ethics@acu.edu.au>
To Taren Sanders; Bridget Booker; Chris Lonsdale

Reply Reply All Forward ...

Sat 21/01/2023 10:46 AM

Dear Taren,

Ethics Register Number : 2017-317H
Project Title : Validation of wearable cameras and GENEActiv wrist-worn accelerometer for assessing health behaviours among children
Data Collection Date Extended : 31/07/2023

Thank you for returning the Ethics Progress Report for your project.

The Chair of the Human Research Ethics Committee has approved your request to extend the project. The new expiry date for the project is the 31/07/2023.

We wish you well in this ongoing project.

Kind regards,
Res Ethics EMAIL

Research Ethics & Integrity Officer | Research Services | Office of the Deputy Vice-Chancellor (Research)
on behalf of ACU HREC Chair, Assoc Prof. Michael Baker
Australian Catholic University
T: 02 9739 2646 E: res.ethics@acu.edu.au

THIS IS AN AUTOMATICALLY GENERATED RESEARCHMASTER EMAIL

Appendix G: Family Information Questionnaire

Confidential

Family Information Questionnaire

Thank you for taking the time to complete this questionnaire.

We will ask you questions about you, your household, and your child. Personal information for both you and your child will be gathered at this time, including name, gender, birthdate, family status, and socioeconomic status. Personal information obtained in this questionnaire will be kept strictly confidential. You may opt out of answering any portion of this questionnaire without penalty.

Child Information

Child's name:	______________________
Child's Date of Birth:	______________________
Child's gender:	◯ Female ◯ Male ◯ Prefer not to say
Suburb lived in:	______________________
Number of siblings:	◯ 0 ◯ 1 ◯ 2 ◯ 3 ◯ 4 ◯ 5 ◯ 6 ◯ 7 ◯ 8 ◯ 9 ◯ 10+

Parent/Guardian One Information

What is your relationship to the child?	______________________ ((e.g., mother, father, grandparent etc))
What is the highest degree or level of school you have compelted? If currently enrolled, highest degree recieved.	◯ Please select from the options below ◯ Never attended school and no non-school qualifications ◯ Level not determined ◯ Certificate not further defined ◯ Year 9 or below ◯ Certificate I/II ◯ Year 10 ◯ Year 11 ◯ Year 12 ◯ Certificate III/IV ◯ Advanced Diploma/Diploma ◯ Bachelor Degree ◯ Graduate Diploma/Graduate Certificate ◯ Postgraduate Degree
Employment Status: Are you currently...?	◯ Please select from the options below ◯ Employed ◯ Unemployed ◯ Prefer not to answer
If employed, what is your job title?	______________________

What is your usual, personal annual income?	◯ Please select from the options below ◯ Nil income ◯ Less than $20,000 ◯ $20,001-$40,000 ◯ $40,001-$60,000 ◯ $60,001-$80,000 ◯ $80,001-$100,000 ◯ $100,001 or over ◯ Prefer not to answer
What is your martial status?	◯ Please select from the options below ◯ Single (never married) ◯ Married, or in a domestic partnership ◯ Widowed ◯ Divorced ◯ Separated ◯ Prefer not to answer

Parent/Guardian Two Information (If applicable)

What is your relationship to the child?	________________________ ((e.g., mother, father, grandparent etc))
What is the highest degree or level of school you have compelted? If currently enrolled, highest degree recieved.	◯ Please select from the options below ◯ Never attended school and no non-school qualifications ◯ Level not determined ◯ Certificate not further defined ◯ Year 9 or below ◯ Certificate I/II ◯ Year 10 ◯ Year 11 ◯ Year 12 ◯ Certificate III/IV ◯ Advanced Diploma/Diploma ◯ Bachelor Degree ◯ Graduate Diploma/Graduate Certificate ◯ Postgraduate Degree
Employment Status: Are you currently...?	◯ Please select from the options below ◯ Employed ◯ Unemployed ◯ Prefer not to answer
If employed, what is your job title?	________________________
What is your usual, personal annual income?	◯ Please select from the options below ◯ Nil income ◯ Less than $20,000 ◯ $20,001-$40,000 ◯ $40,001-$60,000 ◯ $60,001-$80,000 ◯ $80,001-$100,000 ◯ $100,001 or over ◯ Prefer not to answer

What is your martial status?	◯ Please select from the options below ◯ Single (never married) ◯ Married, or in a domestic partnership ◯ Widowed ◯ Divorced ◯ Separated ◯ Prefer not to answer

Appendix H: Camera Selection

There are many different automated wearable cameras, and these devices are each unique in physical characteristics (e.g., size or weight) and features (e.g., camera epochs or image quality). To ensure I selected an automated wearable camera appropriate to measure children's screen use, I created a selection criterion and trialled several automated wearable cameras in both a lab and home setting. I then selected the camera that most aligned with the selection criteria.

Camera Selection Criteria

I created the selection criteria based on my findings from my systematic review (Chapter 2), advice from my supervisors and other researchers in the field, and experience from testing my coding protocol (Chapter 3). My results from the systematic review highlighted several criteria that should be considered when using automated wearable cameras among youth. These criteria included camera privacy considerations, good image quality, camera placement and battery life. After consultation with my supervisors and experts in the field I extended the criteria to include camera storage capacity, weight of the camera, functionality of the camera for participants to prevent data loss and protect privacy, and functionally of the camera software for researchers. The final camera selection criteria included the following:

- Camera must capture either video or timelapse.
- Camera must be able to capture the content on a screen-based media device in different lighting (i.e., image quality).
- Camera must have a battery capacity that was over 7 hours per day.
- Camera must weigh under 100 grams for participant comfort.
- Camera must have a storage capacity for 5 days.
- Camera must be able to be worn on a chest harness (i.e., functionality)

- Camera must be easy for children to use (i.e., easy to turn 'on' and 'off' and easy to put on charge; functionality and privacy).
- The camera software must be functional and easy to use in the field (downloading images for participants to review).
- Participants are not able to access image data without a researcher present (third party privacy).
- Camera should be affordable.

Camera Selection Testing

I tested the cameras in the research lab and at home. To test the image quality, I wore the cameras on a chest harness at home to assess the camera image quality in different lighting for various screen-based media devices. To test the battery capacity and storage capacity of the cameras, I ran the cameras in the research lab until they ran flat, I then calculated the storage capacity of the camera based on the amount of storage the captured data used. To test the privacy features, I ran mock tests as a participant to see if I could access the data without a researcher present. I assessed the weight of the camera based on the weight reported by the manufacture companies. Further, the functionality of the camera and ease of use of the software was based on subjective decisions from myself and other research team members. We ran mock tests to determine if we believed the camera software was feasible to use during data collections in the field. Camera affordability was based on the average cost of the cameras. As seen in Table H1, the Brinno TLC130 met all the camera selection criteria.

Table H1

Results of Camera Testing based on Camera Selection Criteria

Camera	Type of data	Battery Life (≥ 7 hours)	Weight (≤ 100 grams)	Storage (≥ 5 days)	Functionality (Easy to use)	Software (Feasible in field)	Privacy (No access to data)	Price
Fujitsu M-View	Video and Timelapse	Pass	Fail	Fail	Pass	Borderline	Pass	Borderline
DrivePro Body 30	Video	Pass	Fail	Fail	Pass	Pass	Pass	Pass
PR6 Body Worn Camera	Video	Pass	Fail	Fail	Pass	DNT	Pass	Fail
Axon Body 2 Camera	Video	Pass	Fail	Fail	Pass	Borderline	Pass	Pass
Axon Body 3 Camera	Video	Pass	Fail	Fail	Pass	Borderline	Pass	Fail
Brinno TLC120	Timelapse	Pass	Borderline	Pass	Borderline	Pass	Pass[a]	Pass
Brinno TLC130	Timelapse	Pass	Pass	Pass	Pass	Pass	Pass[a]	Pass

Note. Pass = meets criteria; Borderline = almost meets criteria; Fail = does not meet criteria; DNT = did not test.

[a] Pass is based on using an encrypted SD card and not the cameras standalone security features.

Appendix I: List of Observation Task

- Watch something entertaining on a tablet.
- Watch something entertaining on television while drinking a beverage.
- Watch something entertaining on television while eating a snack.
- Watch something educational on a mobile phone.
- Watch something entertaining on a television while using a laptop.
- Play an educational video game on a laptop or computer.
- Play an entertaining video game on a tablet.
- Watch something entertaining on a tablet while eating a snack.
- Play an entertaining video game on a laptop or computer.
- Watch something educational on television while eating a snack.
- Watch something educational on television.
- Play an entertaining video game on a television.
- Watch something educational on a laptop while drinking a beverage.
- Watch something educational on a mobile phone while drinking a beverage.
- Play an educational video game on a mobile phone.
- Watch something educational on a television while using a laptop.
- Watch something entertaining on a laptop while eating a snack.
- Play an entertaining video game on a handheld game console.
- Watch something educational on television while drinking a beverage.
- Watch something educational on a television while using a tablet.
- Watch something entertaining on a television while using a tablet.
- Watch something entertaining on a television while playing a handheld game console.
- Watch something educational on a television while using a mobile phone.
- Play an educational video game on a television.

- Watch something entertaining on a mobile phone.
- Watch something educational on a tablet while eating a snack.
- Watch something educational on a laptop while eating a snack.
- Watch something educational on a tablet.
- Watch something entertaining on television.
- Use a mobile phone.
- Use a mobile phone while watching a video on a laptop.
- Play an educational video game on a tablet.
- Watch something educational on a laptop or computer.
- Watch something entertaining on a television while using a mobile phone.
- Watch something educational on a tablet while drinking a beverage.
- Watch something entertaining on a laptop or computer.
- Watch something educational on a television while playing a handheld game console.
- Play an entertaining video game on a mobile phone.
- Play an educational video game on a handheld game console.
- Watch something entertaining on a laptop while drinking a beverage.
- Watch something educational on a mobile phone while eating a snack.
- Watch something entertaining on a tablet while drinking a beverage.
- Watch something entertaining on a mobile phone while eating a snack.
- Watch something entertaining on a mobile phone while drinking a beverage.
- Read a book.
- Write in a booklet or on a piece of paper.
- Draw in a booklet or a piece of paper.
- Eat a meal or a snack.
- Drink a beverage.

- Play with a toy.

Appendix J: KidVision Participant Letters and Consent Forms

Confidential

KidVision Consent Form

Thank you for your interest in the KidVision Research Proje

First, take a look at the Participant Information Letter belov
questions and entering your details, and you'll be on your v

The Information Letter describes all aspect of the project in
Information Letter and contact us (02 9701 4629 or kidvisi
project procedures.

The Consent Form outlines the key aspects of the study th
carefully read the items so that you understand what is inv
the study and still participate, however opting out of weari
this study.

We will also ask for some of your personal details (i.e., nam
be stored securely and confidentially, will not be shared wi
research project.

Once submitted, you will be registered to participate in the
schedule a time to get started.

Parent/Guardian Information Letter

Confidential

Project Title: KidVision: Using Wearable Cameras to Unders
HREC Registration Number: 2020-142H
Principal Investigator: Professor Chris Lonsdale

Dear Parent(s)/Caregiver(s),

We invite you and your child to participate in a research stu
children interact with the world.

Who can participate in the study?

Children in Years 2-4 (or children aged 7-10) in 2022 and th
participate, we are inviting you to complete the attached c

What is the study about?

How do children spend their time?
How do children interact with the world and with each othe

Children's lives have changed in the past few decades. Ch
and recreation than other generations. They may also be le
bike to school. The impact of these changes on children's l
research has relied on subjective data about how children s

Our study will be one of the first in the world to collect obj
cameras) to understand how children interact with the wor
opportunity to learn more about how children spend their t
learning and development.

Where will the study take place?

What are we asking participants to do?

Parent(s)/Carer(s)
We will ask parents of children in Years 2-4 (or children age
and 2024. At each timepoint, we will collect data over a pe
you and a trained researcher at each timepoint to provide
need to complete the study. At the beginning, we will ask t
the most time with the child at home) to complete a 15-20-
complete the survey online or at the same time the child c
repeated again in 2023 and 2024.

Your child will be asked to wear a camera on four of the da
of the study at each timepoint. We will ask permission to se
morning while their child is completing the study measures
need to do each day. At the end of the study, we will invite
interview that would take up to 30 minutes.

At the end of the eight days (at each time point) we will asl
the wearable camera and physical activity/sleep monitor. H
review the images from the camera and delete any photo p

If you do not want to review the images and would prefer t
provide you with a prepaid postage bag. If you choose this
physical activity device via postage.

Child Measures
We are inviting children in Years 2-4 (children aged 7-10) t
and 2024. At the end of the study we will also invite a smal
up to 30 minutes. We will ask children to meet with the res
2024) for 30-minutes. We will ask children to complete the

1. Questionnaire:
Children will be asked to complete a 20-minute questionna
again 12 and 24 months later (2023 and 2024). At the end

another person. Children will also not wear the camera on s
to wear the camera before and after school (i.e., off school

A trained research assistant will demonstrate to all childrer
the camera. To ensure children understand, the research a
they can disable their device and answer any questions. To
parties, we will:

• Use 'imaging blurring' technology to blur faces captured l
computer program will automatically blur the faces on ima
what the blurred images look like. The original images will

• The child and their parent/carer will have an opportunity
permanently delete any image for any reason. The researc
view the child's data. To maintain the privacy of third parti
Please see "I am concerned about privacy - how will privac
information.

Note: We will not hold any child or parent responsible if an
lost. But, we do hope to get these devices back as they ha
our research.

How much time will the study take?

Parent(s)/Carer(s)
Parents will require approximately 15-20 minutes to compl
2022, 2023 and 2024). Parents can choose to complete the
at an ACU campus. In total, parents will need 45-60 minute
A small sub-sample of parents will be invited to participate
at the end of the study at a research centre.

Researchers are obliged to comply with lawful orders (e.g.,
when required. In the event the research team views data
has been harmed or is at risk of being harmed, the researc
University's ethics committee and to the appropriate Gove

What are the benefits of the study?

The findings of this study will help us better understand ho
The information gathered in this study will help us to identi
and development.

We will provide you with a summary report of the findings
data about individual participants (children or parents) but
non-technical language. We will also provide the Departme
policy development and teacher professional learning, if ap

At each time-point (i.e., 2022, 2023, and 2024) your family
supermarket (Coles or Woolworths) gift card. The total com
will be delivered digitally to you (the parent/guardian) at ea

Can participants withdraw from the study?

Participation in the study is entirely voluntary. If you agree
and stop participating at any time. A decision not to partici
jeopardise your relationship with the Australian Catholic Ur
study, we will ask for consent to keep the data previously c
remove it completely. At the conclusion of the study, all ide
delete participant data after the study's conclusion.

I am concerned about privacy - how will privacy and confid

The researchers will follow strict protocols to ensure any pe
has been collected, de-identified using a coding system, ar
or other data collection sheets will be destroyed. Images co

If you have any complaints or reservations about the ethica
as a research participant, you may contact the ACU Humar
resethics.manager@acu.edu.au and quote the approval nu
confidence and investigated fully, and you will be informed

What do I need to do?
Thank you for your consideration of our study. Please comp
participate in all or part of the study. The form can be retur
KidVision@acu.edu.au

Yours sincerely,

On behalf of the research team

Professor Chris Lonsdale
Institute for Positive Psychology and Education
Australian Catholic University

A PDF version of this information is available from the link

[Attachment: "Parent Participant Information Letter.pdf"]

Parent/Guardian Consent Form (Parent to Complete)

Parent First Name:

Parent Surname:

Parent Email:

Confidential

Completing the parent/carer questionnaire at three time points.

My child completing a short questionnaire and the cognitive battery (20 mins).

My child wearing an activity and sleep monitor for 8 days.

My child wearing a camera for 4 days, including weekdays outside school hours (before and after) and on one weekend day.

The research team sending me a text message reminder for the 8 days at each time point.

Allowing the NSW Educational Standards Authority (NESA) or the Queensland Curriculum & Assessment Authority (QCAA) to provide the research team with access to my child's Year 3 and Year 5 NAPLAN data.

Allowing the NSW Educational Standards Authority (NESA) or the Queensland Curriculum & Assessment Authority (QCAA) to provide the research team with access to my child's Student Enrolment Dataset for parent related data (i.e., ethnicity, language spoken, occupation, and education).

Allowing the NSW Educational Standards Authority

Confidential

Today's Date:

Confidential

Child Information Letter

Dear child,

We would like to ask you to take part in a research study c
like to take part. If you would like to be involved, please co
after you.

Who can take part in the study?
If you are in Year 2, 3, or 4 in 2022 you can take part.

What is the study about?
This study is seeing how kids spend their time and see the
how other people (like mum and dad) spent their time as a
devices to see how you spend your time.

Where will the study take place?
You and the person who looks after you can choose. We ca
home. OR, you can choose to visit the University office. The
at a university campus:

What are we asking you to do?
We will ask to meet with you once each year (in 2022, 202
answer some questions, complete two games, and wear a
or all of these activities, and choose to skip any questions y

The Questions
The questions will take 20 minutes each time. We will visit
two games to help us see how you think.

Can I leave the study?
It is your choice to take part in the study. If you agree to ta

Extra information about the camera
The researchers will keep your photos private. The only pe
you and the person who looks after you. You and the perso
one is allowed to keep their photos and no one else is allov
researchers can see the photo with a special code. You will
uncomfortable.

Thank you,

Professor Chris Lonsdale
Institute for Positive Psychology and Education

Australian Catholic University

A PDF of the Child Information Letter can be found here:

[Attachment: "Child Participant Information Letter.pdf"]

Child Assent Form (Child to Complete)

I have read the information letter with the person
that looks after me

I want to take part in the study

My Signature:

Appendix K: Evidence of Ethics Approval for KidVision project

From: Res Ethics EMAIL ONLY
To: Chris Lonsdale; Kirsty Bergan
Subject: [2020-142H] - Ethics Extension Request Approved
Date: Thursday, 27 January 2022 4:25:13 PM

Dear Chris
,

Ethics Register Number : 2020-142H
Project Title : Square Eyes or All Lies? Understanding Children's Exposure to Screens.
Data Collection Date Extended : 25/12/2025

Thank you for returning the Ethics Progress Report for your project.

The Chair of the Human Research Ethics Committee has approved your request to extend the project. The new expiry date for the project is the 25/12/2025.

We wish you well in this ongoing project.

Kind regards,
Res Ethics EMAIL ONLY

Research Ethics & Integrity Officer | Research Services | Office of the Deputy Vice-Chancellor (Research)
on behalf of ACU HREC Chair, Assoc Prof. Michael Baker
Australian Catholic University
T: 02 9739 2646 E: res.ethics@acu.edu.au

THIS IS AN AUTOMATICALLY GENERATED RESEARCHMASTER EMAIL

Appendix L: KidVision Child Questionnaire

Confidential

KidVision Child Questionnaire

Thank you for taking the time to complete this questionnai
will ask you to complete two tasks on the iPad. This should

Please answer each question as best as you can. Don't wor
question if you like. But, remember there are no right or wr
raise your hand to ask for help.

To begin, please make sure that your name is listed below.
before continuing.

My name is: [enrolment_arm_1][cid_namef] [enrolment_ar

◯ Yes ◯ No

Confidential

	The questions below ask about how you feel.	Never	Sel
1	How often do you feel scared?	○	(
2	How often do you feel sad?	○	(
3	How often do you feel mad?	○	(
4	How often do you worry about what will happen to you?	○	(
5	How often is it hard for you to get along with other kids?	○	(
6	How often do other kids say they do not want to play with you?	○	(
7	How often do other kids tease you?	○	(
8	How often is it hard for you to pay attention in school?	○	(
9	How often do you forget things?	○	(
10	How often is it hard to keep up with schoolwork?	○	(

Confidential

These next questions ask about electronic screens. Screens can mean anything that shows a picture that you watch or interact with.

Below are some pictures of the screens that you may use. These include an iPod Touch, iPad, Mobile Phone or iPhone, TV, Laptop Computer, Portable PlayStation, Nintendo Switch, or an Xbox.

Examples of things you can do on screens are watch TV, search the internet, use social media, use instant messenger, send and receive emails, play games, online shopping, download music, do school work and homework, and watch music videos.

Please do NOT include time spent on school related work.

Confidential

11 On an average day in the past week, how much time do you play single-player video games on a computer, console, phone, or other device (Xbox, PlayStation, iPad, AppleTV)?

Confidential

12 On an average day in the past week, how much time do you play multiplayer video games on a computer, console, phone, or other device (Xbox, PlayStation, iPad, AppleTV) where you can interact with others in the game?

Confidential

13 On an average day in the past week, how much time do you text on a mobile phone, tablet, iPod, or other electronic device (e.g., GChat, Whatsapp, Kik, etc.)?

Confidential

14 On an average day in the past week, how much time do you visit social media apps (e.g., Snapchat, Facebook, Twitter, Instagram, TikTok, etc.)? (Do not include time spent editing photos or videos to post on social media.)

Confidential

15 On an average day in the past week, how much time do you video chat (Skype, FaceTime, VRchat, etc.) that is **NOT** for school?

Confidential

16 On an average day in the past week, how much time do you spend searching or browsing the internet (e.g., using Google) that is **NOT** for school?

Confidential

17 On an average day in the past week, how much time do you watch "or stream" movies, videos, or TV shows? (Such as Hulu, Netflix, Amazon, YouTube, Twitch)

Confidential

Now think about time spent on school related w

18 On an average day in the past week, how much time do you do school work or homework on a computer, phone, tablet, or other electronic device?

Confidential

19 On an average day in the past week, how much time do you video chat for school?

Confidential

20 On an average day in the past week, how much time do you search or browse the internet for school?

- ◯ 0
- ◯ 15 mins
- ◯ 30 mins
- ◯ 45 mins
- ◯ 1 hr
- ◯ 1.5 hrs
- ◯ 2 hrs
- ◯ 2.5 hrs
- ◯ 3 hrs
- ◯ 4 hrs
- ◯ 5 hrs
- ◯ 6 hrs
- ◯ 7 hrs
- ◯ 8 hrs
- ◯ 9 hrs
- ◯ 10 hrs
- ◯ 11 hrs
- ◯ 12 hrs
- ◯ 13 hrs
- ◯ 14 hrs
- ◯ 15 hrs
- ◯ 16 hrs
- ◯ 17 hrs
- ◯ 18 hrs
- ◯ 19 hrs
- ◯ 20 hrs
- ◯ 21 hrs
- ◯ 22 hrs
- ◯ 23 hrs
- ◯ 24 hrs

Confidential

Now think about **EITHER** Saturday or Sunday last weekend.
on one of these days?
(Start from the time you woke up and think about the total
after school, at home or at a friend's house, and in the eve

- ◯ 0 hours
- ◯ 1 hour
- ◯ 2 hours
- ◯ 3 hours
- ◯ 4 hours
- ◯ 5 hours
- ◯ 6 hours
- ◯ 7 hours
- ◯ 8 hours
- ◯ 9 hours
- ◯ 10 hours
- ◯ 11 hours
- ◯ 12 hours

Confidential

These questions ask you about your physical ac
about physical activity, it can be anything that
school work that you do after school.

21 Thinking just about ONE typical day last week (Monday to F
active or exercising, such as playing sports, working out, d

- ◯ None
- ◯ 5 to 30 minutes
- ◯ 30 minutes to 1 hour
- ◯ 1 hour to 3 hours
- ◯ More than 3 hours

22 Thinking just about ONE typical day last week (Monday to F
was for your own enjoyment (not homework)?

- ◯ None
- ◯ 5 to 30 minutes
- ◯ 30 minutes to 1 hour
- ◯ 1 hour to 3 hours
- ◯ More than 3 hours

23 Thinking just about ONE typical day last week (Monday to F

- ◯ None
- ◯ 5 to 30 minutes
- ◯ 30 minutes to 1 hour
- ◯ 1 hour to 3 hours
- ◯ More than 3 hours
- ◯ My school does not have homework

Appendix M: KidVision Parent and Caregiver Questionnaire

Confidential

KidVision Parent/Caregiver Q

Thank you for taking the time to complete this questionnai
behaviour, and their environment. Please answer each que

To begin, please make sure that your name is listed below.
before continuing.

My name is: [enrolment_arm_1][pid_namef] [enrolment_arr

◯ Yes ◯ No

Confidential

The below questions ask about your child's bacl

What is the name of the school your child attends?

Does Study Child speak a language other than English at h

◯ Yes ◯ No

What is the main language study child speaks at home?

Is Study Child of Aboriginal or Torres Strait Islander origin?

◯ No
◯ Yes, Aboriginal
◯ Yes, Torres Strait Islander
◯ Yes, both

Confidential

The following questions ask about you.

How are you related to study child?
(Step-parents include those from de facto relationships)

- ◯ Unrelated adult
- ◯ Biological mother
- ◯ Biological father
- ◯ Adopted mother
- ◯ Adopted father
- ◯ Step-mother
- ◯ Step-father
- ◯ Foster mother
- ◯ Foster father
- ◯ Legal parent

Were you born in Australia?

◯ Yes ◯ No

In which country were you born?

Do you speak a language other than English at home?

◯ Yes ◯ No

What is the main language you speak at home?

What was the highest year of school you completed?

Confidential

What is your primary occupation?

Confidential

The following questions ask about other people

How many younger siblings does the study child have in the household?

How many older siblings does the study child have in the household?

Is there another parent or caregiver to the study child in th

◯ Yes ◯ No

What is parent/caregiver 2's relationship to the study child'
(Step-parents include those from de facto relationships)

◯ Unrelated adult
◯ Biological mother
◯ Biological father
◯ Adopted mother
◯ Adopted father
◯ Step-mother
◯ Step-father
◯ Foster mother
◯ Foster father
◯ Legal parent

Was parent/caregiver 2 born in Australia?

◯ Yes ◯ No

In which country was parent/caregiver 2 born?

Confidential

What is the highest qualification that parent/caregiver 2 ha

- ◯ Postgraduate degree
- ◯ Graduate diploma/Graduate certificate
- ◯ Bachelor degree (with or without honours
- ◯ Advanced diploma
- ◯ Certificate III/IV (including trade certificate)
- ◯ Certificate I/II
- ◯ Other non-school qualifications

What is parent/caregiver 2's employment status?

- ◯ Employed full-time (30+ hrs/week)
- ◯ Employed part-time (< 30 hrs/week)
- ◯ Unemployed and looking for work
- ◯ Not in the labour force (incl stay-at-home parent)

What is parent/caregiver 2's primary occupation?

Confidential

These questions ask about screens. Screens can mean anything that shows a picture that you watch or interact with.

Below are some pictures of the screens that your child may use. These include an iPod Touch, iPad, Mobile Phone or iPhone, TV, Laptop Computer, Portable PlayStation, or an Xbox. Examples of things you can do on screens are watch TV, search the internet, use social networking sites, use instant messenger, send and receive emails, play games, online shopping, download music, do school work and homework, and watch music videos.

Please answer the following questions about [admin_arm_1][cid_namef] [admin_arm_1][cid_namel].

Confidential

Think about ONE typical day last week (Monday to Friday).
screens that day?
(Start from the time your child woke up and think about th
school, after school, at home or at a friend's house, and in

- ◯ 0 hours
- ◯ 1 hour
- ◯ 2 hours
- ◯ 3 hours
- ◯ 4 hours
- ◯ 5 hours
- ◯ 6 hours
- ◯ 7 hours
- ◯ 8 hours
- ◯ 9 hours
- ◯ 10 hours
- ◯ 11 hours
- ◯ 12 hours

Thinking of the same weekday last week (Monday to Friday
hours and minutes [enrolment_anm_1][cid_namef] [enrolm

0 mins 300 mins 600 mins

(Place a mark on the scale above)

Calc - web use hours

Calc - web use minutes

Confidential

Calc - gaming hours

Calc - gaming minutes

Now think about EITHER Saturday or Sunday last weekend.
screens on one of these days?
(Start from the time your child woke up and think about th
and in the evening until the time your child went to bed.)

- ◯ 0 hours
- ◯ 1 hour
- ◯ 2 hours
- ◯ 3 hours
- ◯ 4 hours
- ◯ 5 hours
- ◯ 6 hours
- ◯ 7 hours
- ◯ 8 hours
- ◯ 9 hours
- ◯ 10 hours
- ◯ 11 hours
- ◯ 12 hours

Thinking of the same weekend day last week (Saturday or
many hours and minutes [enrolment_arm_1][cid_namef] [e

0 mins 300 mins 600 mins

(Place a mark on the scale above)

Confidential

Calc - social media minutes minutes

Thinking of the same weekend day last week (Saturday or
many hours and minutes [enrolment_arm_1][cid_namef] [e

0 mins 300 mins 600 mins

(Place a mark on the scale above)

Calc - gaming hours

Calc - gaming minutes

Confidential

These questions ask about your child's physical
about physical activity, it can be anything that
any school work that your child may do after sc

Thinking about ONE typical day last week (Monday to Frida
active or exercising, such as playing sports, working out, d

- ◯ None
- ◯ 5 to 30 minutes
- ◯ 30 minutes to 1 hour
- ◯ 1 hour to 3 hours
- ◯ More than 3 hours

Thinking about ONE typical day last week (Monday to Frida
that was for their own enjoyment (not homework)?

- ◯ None
- ◯ 5 to 30 minutes
- ◯ 30 minutes to 1 hour
- ◯ 1 hour to 3 hours
- ◯ More than 3 hours

Thinking about ONE typical day last week (Monday to Frida
after school?

- ◯ None
- ◯ 5 to 30 minutes
- ◯ 30 minutes to 1 hour
- ◯ 1 hour to 3 hours
- ◯ More than 3 hours
- ◯ My child's school does not have homework

Confidential

For each item, please mark the box for Not True
help us if you answered all items as best you ca
item seems daft. Please give your answers on t
six months or this school year.

Your child...

	Not true
Is considerate of other people's feelings	○
Is restless, overactive, cannot stay still for long	○
Often complains of headaches, stomach-aches or sickness	○
Shares readily with other children, for example toys, treats, pencils	○
Often has temper tantrums or hot tempers	○
Is rather solitary, tends to play alone	○
Is generally obedient, usually does what adults request	○
Has many worries or, often seems worried	○

Confidential

Is picked on or bullied by other children	○
Often volunteers to help others (parents, teachers, other children)	○
Can stop and think things out before acting	○
Can be spiteful to others	○
Gets on better with adults than with other children	○
Has many fears, is easily scared	○
Sees tasks through to the end, has a good attention span	○

Confidential

This sections asks about your child's typical bel
question is of your child.

My child...

	Never	Sel
Stays with homework until finished		
Does not complete homework		
Has difficulty completing assignments		
Remembers to do homework		
Returns to responsibilities		
Goes back to the task at hand		
Quits routine household chores		
Leaves own projects unfinished		
Gets frustrated with projects and quits		
When an activity is difficult, gives up		
Switches from one activity to another		

Appendix N: Study 3 Sensitivity Analyses

Table N1

Sensitivity analysis of Pearson Correlations (r) for Screen Use Variables for 8-hour, 10-hour, 12-hour, 14-hour, 16-hour and 18-hour Adjusted Camera, Parent-Reported and Child-Reported Measurements

Variable	8-hour Adjusted Camera		10-hour Adjusted Camera		12-hour Adjusted Camera		14-hour Adjusted Camera		16-hour Adjusted Camera		18-hour Adjusted Camera	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Aggregated screen time												
Unadjusted camera screen time	.96	<.001	.96	<.001	.96	<.001	.96	<.001	.96	<.001	.96	<.001
Parent-reported screen time	.22	.433	.22	.433	.22	.433	.22	.433	.22	.433	.22	.433
Child-reported screen time	.00	.993	.00	.993	.00	.993	.00	.993	.00	.993	.00	.993
Recreational screen time												
Unadjusted camera recreational content	.97	<.001	.97	<.001	.97	<.001	.97	<.001	.97	<.001	.97	<.001
Child-reported recreational content	.32	.238	.32	.238	.32	.238	.32	.238	.32	.238	.32	.238
Educational screen time												
Unadjusted camera educational content	1	<.001	1	<.001	1	<.001	1	<.001	1	<.001	1	<.001
Child-reported educational content	-.02	.942	-.02	.942	-.02	.942	-.02	.942	-.02	.942	-.02	.942
Social screen time												
Unadjusted camera social content	1	<.001	1	<.001	1	<.001	1	<.001	1	<.001	1	<.001
Parent-reported social content	.26	.344	.26	.344	.26	.344	.26	.344	.26	.344	.26	.344

Variable	8-hour Adjusted Camera		10-hour Adjusted Camera		12-hour Adjusted Camera		14-hour Adjusted Camera		16-hour Adjusted Camera		18-hour Adjusted Camera	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
Child-reported social content	.17	.544	.17	.544	.17	.544	.17	.544	.17	.544	.17	.544
Gaming												
Unadjusted camera gaming	1	<.001	1	<.001	1	<.001	1	<.001	1	<.001	1	<.001
Parent-reported gaming	.2	.475	.2	.475	.2	.475	.2	.475	.2	.475	.2	.475
Child-reported gaming	.06	.844	.06	.844	.06	.844	.06	.844	.06	.844	.06	.844
Internet browsing												
Unadjusted camera internet browsing	1	<.001	1	<.001	1	<.001	1	<.001	1	<.001	1	<.001
Parent-reported internet browsing	.14	.625	.14	.625	.14	.625	.14	.625	.14	.625	.14	.625
Child-reported internet browsing	.12	.681	.12	.681	.12	.681	.12	.681	.12	.681	.12	.681
Programme viewing												
Unadjusted camera programme viewing	.97	<.001	.97	<.001	.97	<.001	.97	<.001	.97	<.001	.97	<.001
Parent-reported programme viewing	.13	.641	.13	.641	.13	.641	.13	.641	.13	.641	.13	.641
Child-reported programme viewing	.13	.636	.13	.636	.13	.636	.13	.636	.13	.636	.13	.636
Communication												
Unadjusted camera communication	1	<.001	1	<.001	1	<.001	1	<.001	1	<.001	1	<.001
Child-reported communication	.58	.024	.58	.024	.58	.024	.58	.024	.58	.024	.58	.024

Note. *N* = 15 participants

Figure N1

Bland-Altman plots of Aggregated Screen Time for Adjusted Camera and Parent-Reported (PR) Measurements (Sensitivity Analysis)

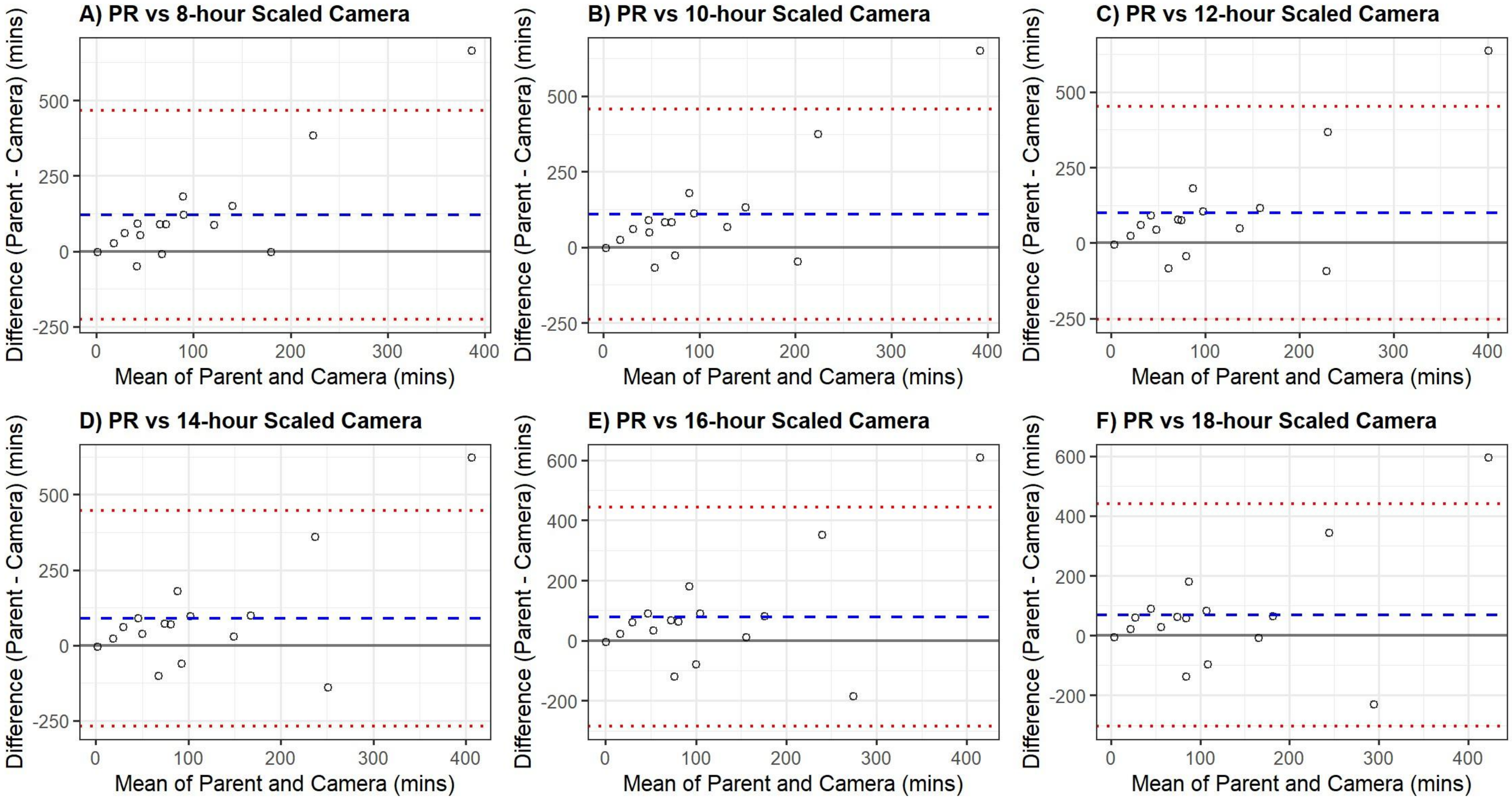


Figure N2

Bland-Altman plots of Aggregated Screen Time for Adjusted Camera and Child-Reported (CR) Measurements (Sensitivity Analysis)

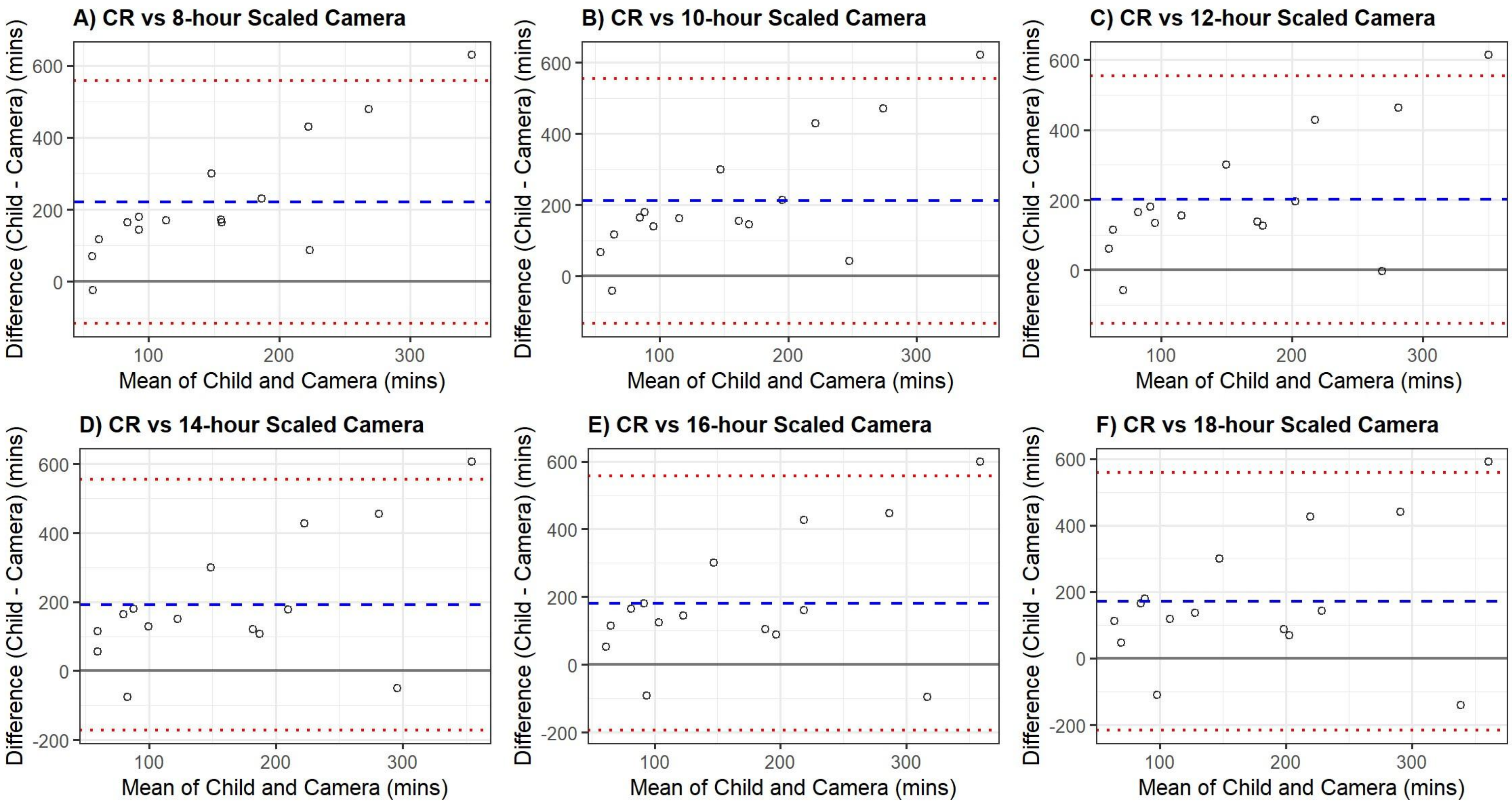


Figure N3

Bland-Altman plots of Recreational Screen Use for Adjusted Camera and Child -Reported (CR) Measurements (Sensitivity Analysis)

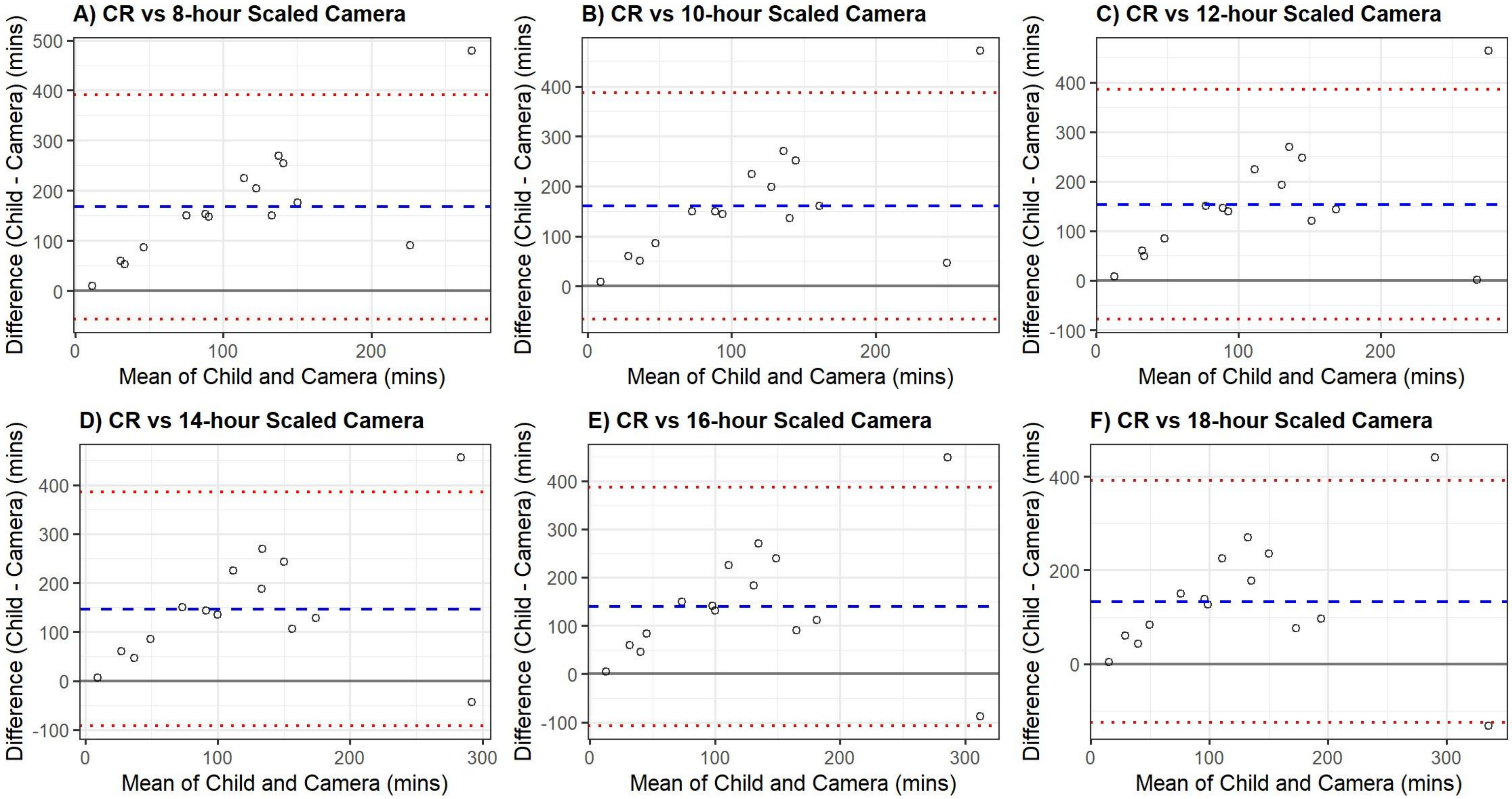


Figure N4

Bland-Altman plots of Educational Screen Use for Adjusted Camera and Child -Reported (CR) Measurements (Sensitivity Analysis)

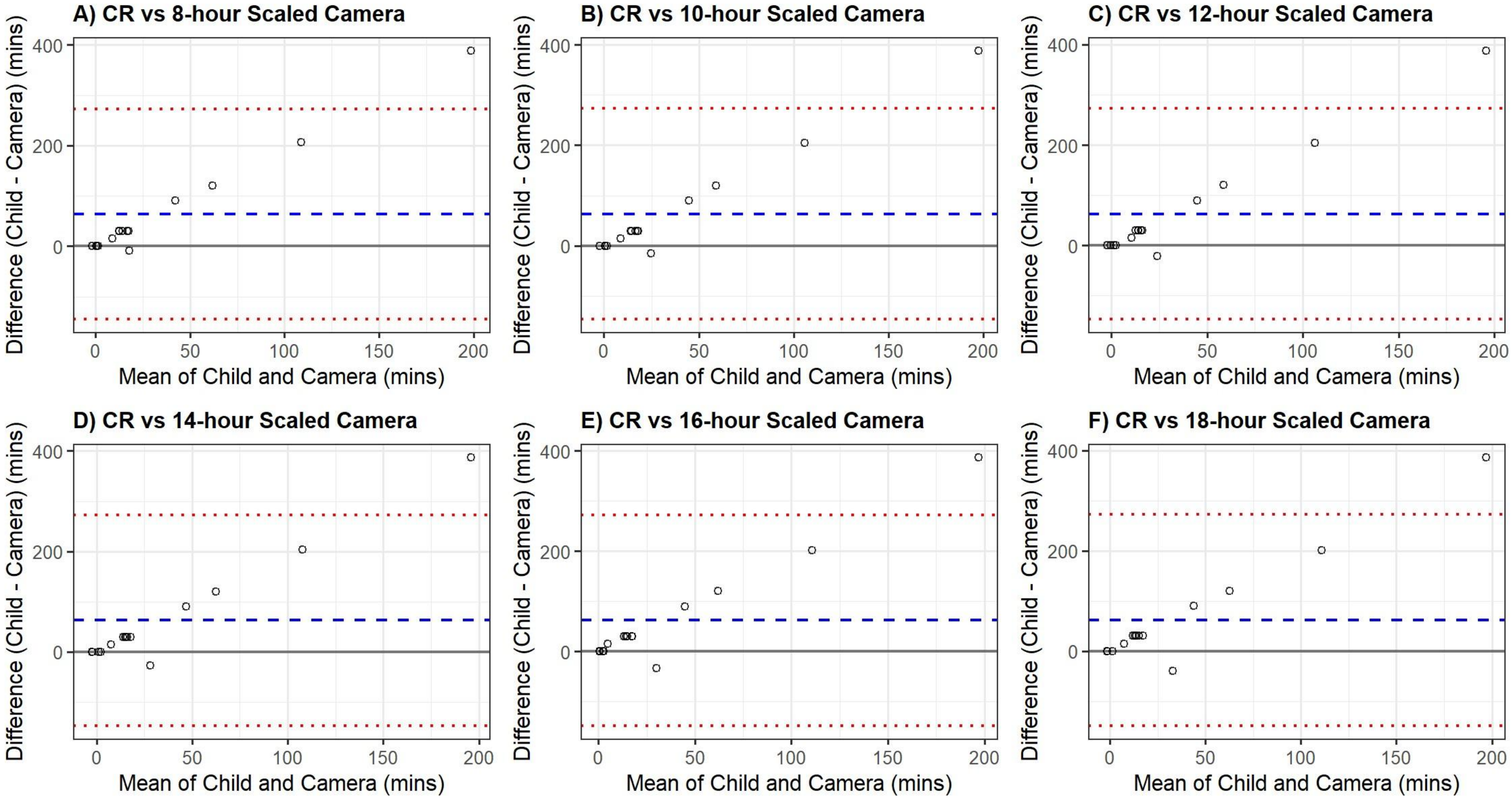


Figure N5

Bland-Altman plots of Social Screen Use for Adjusted Camera and Parent-Reported (PR) Measurements (Sensitivity Analysis)

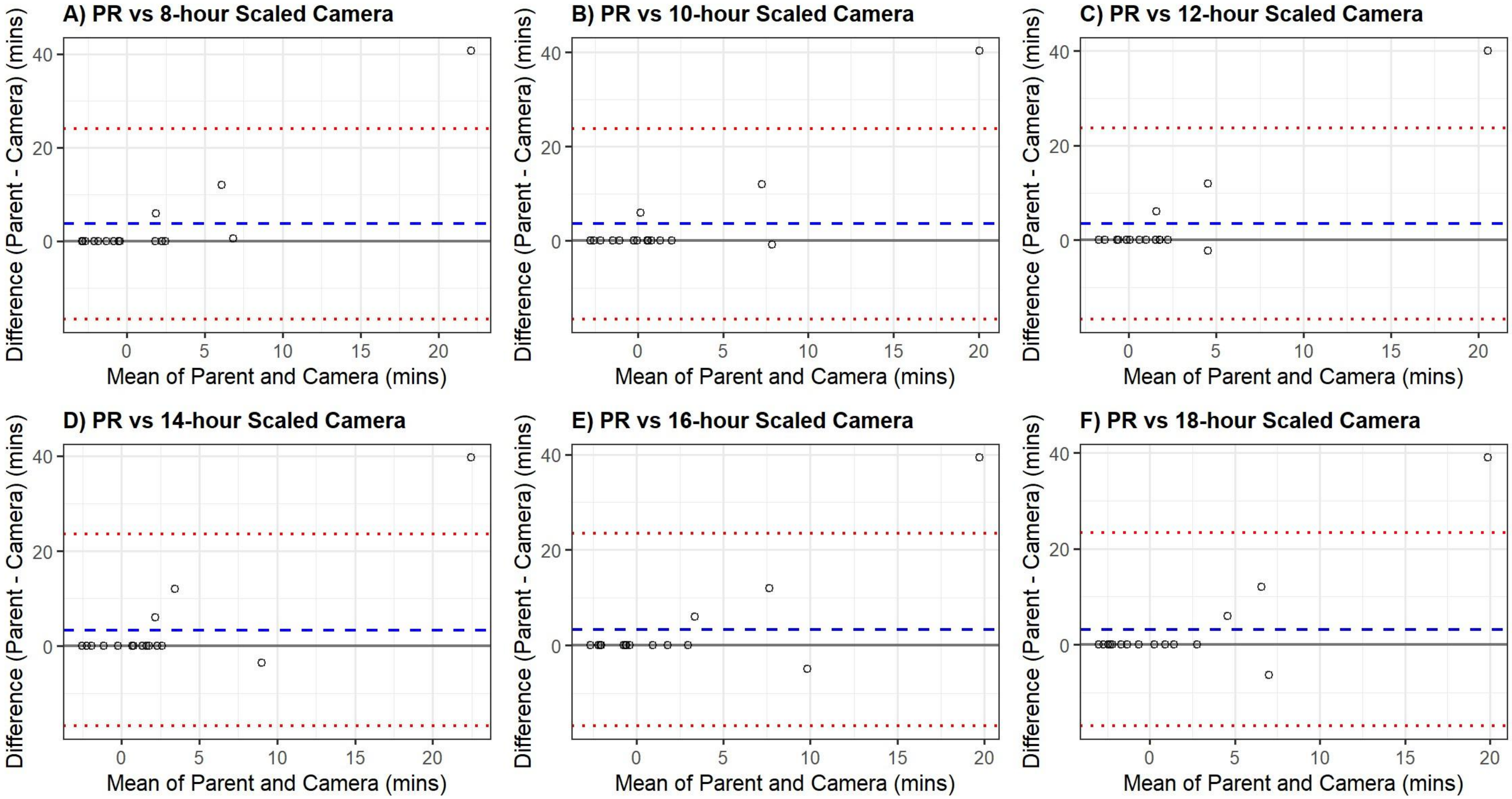


Figure N6

Bland-Altman plots of Social Screen Use for Adjusted Camera and Child-Reported (CR) Measurements (Sensitivity Analysis)

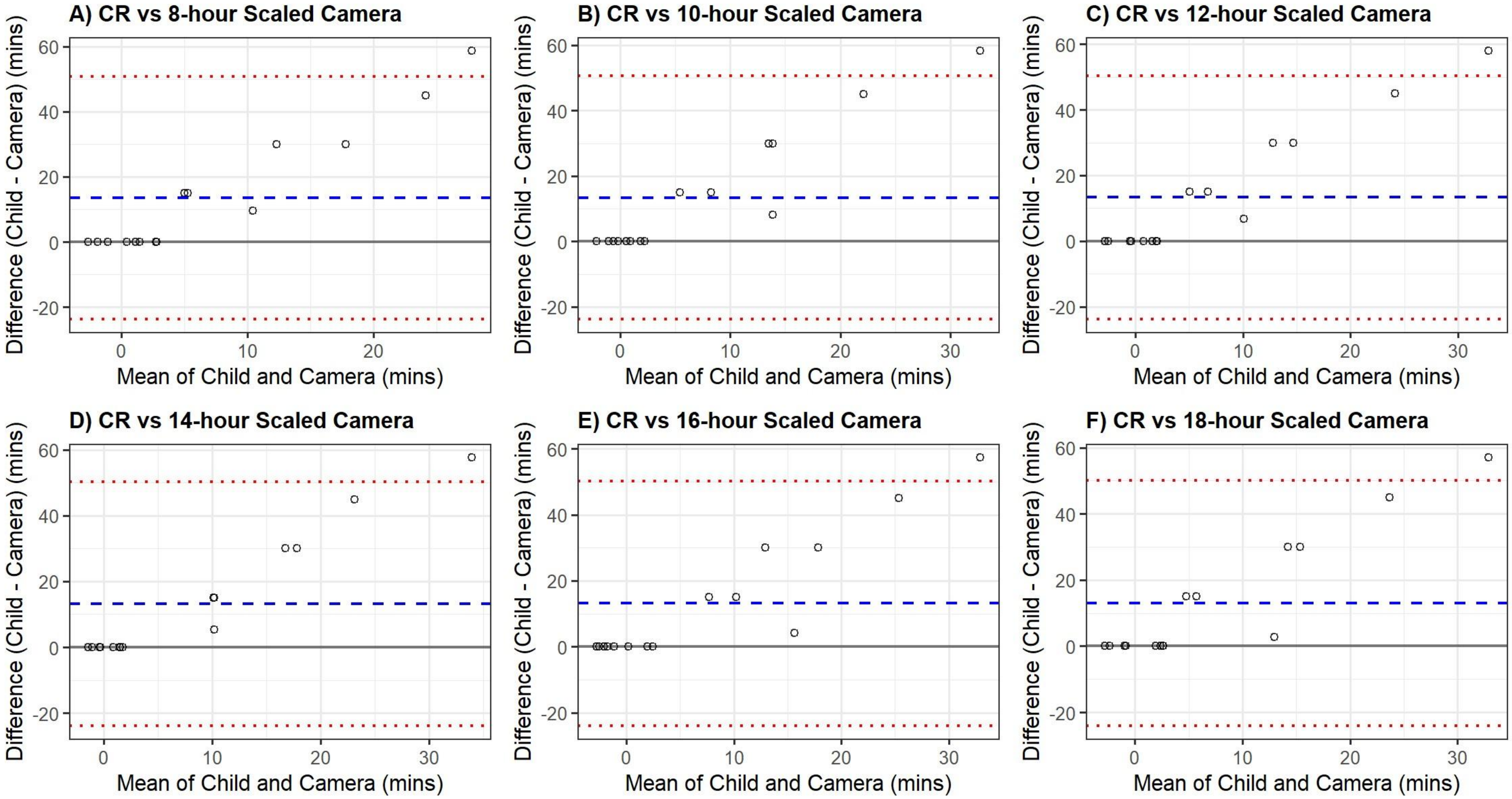


Figure N7

Bland-Altman plots of Gaming for Adjusted Camera and Parent-Reported (PR) Measurements (Sensitivity Analysis)

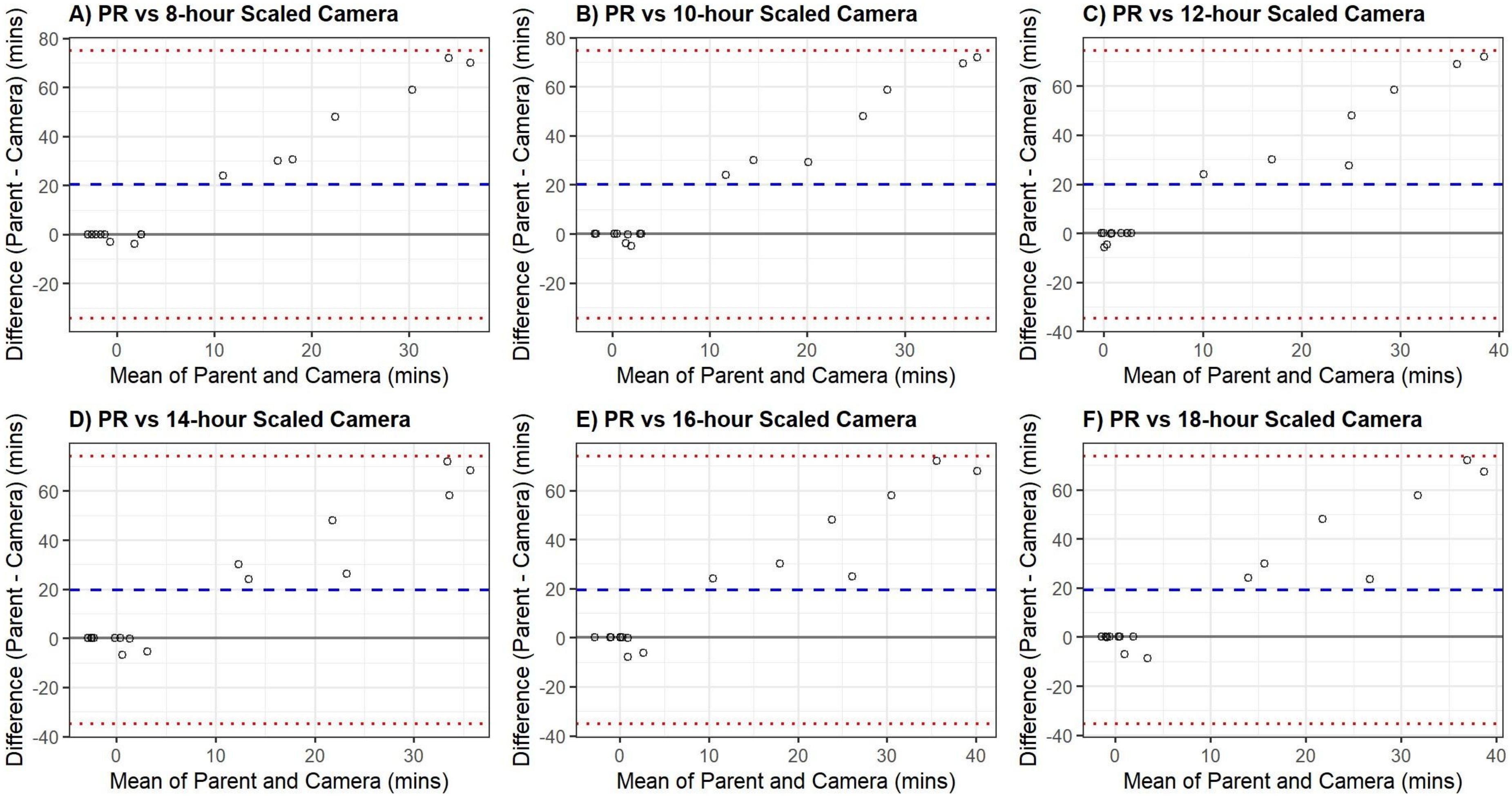


Figure N8

Bland-Altman plots of Gaming for Adjusted Camera and Child-Reported (CR) Measurements (Sensitivity Analysis)

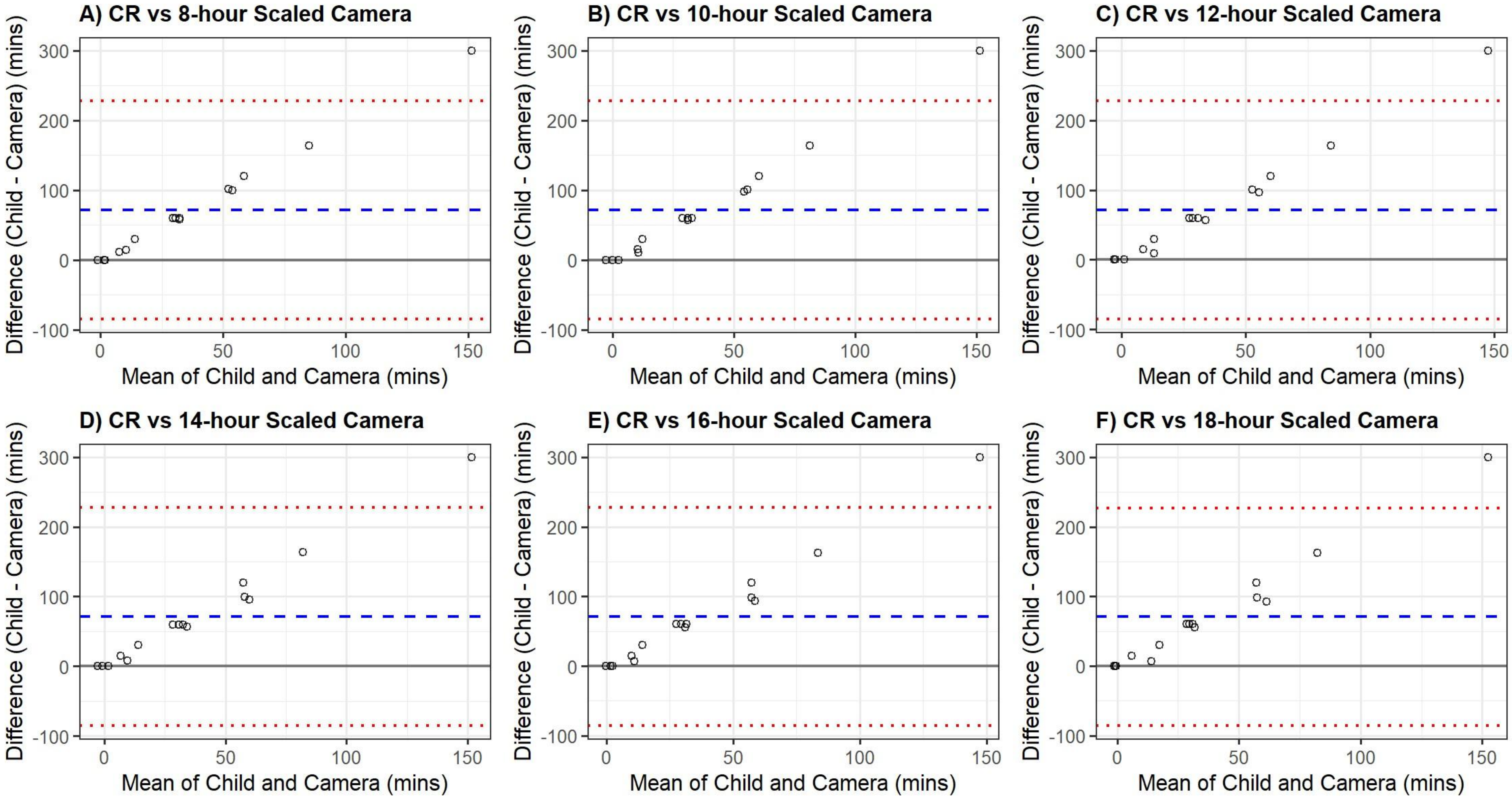


Figure N9

Bland-Altman plots of Internet Browsing for Adjusted Camera and Parent-Reported (PR) Measurements (Sensitivity Analysis)

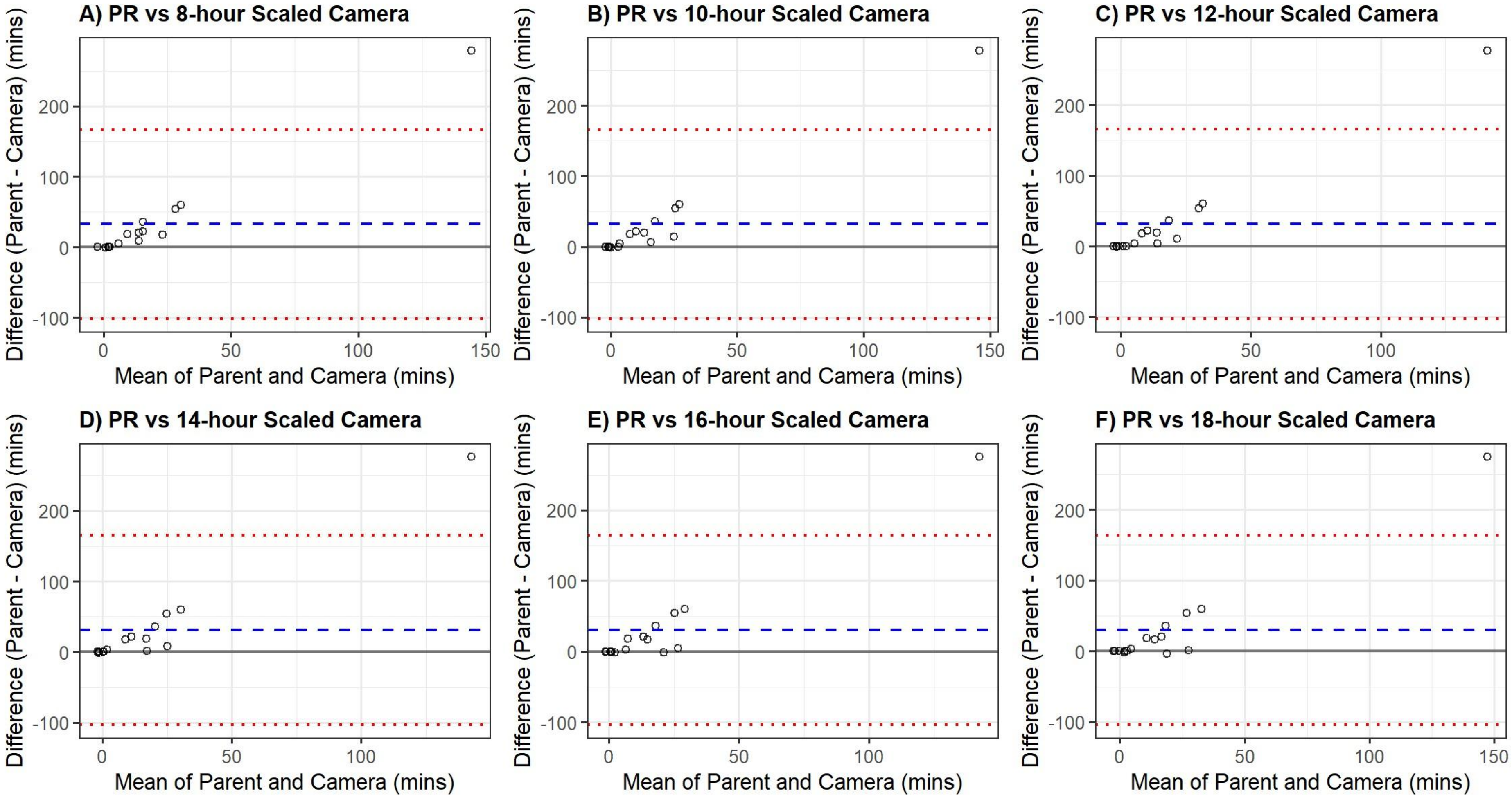


Figure N10

Bland-Altman plots of Internet Browsing for Adjusted Camera and Child-Reported (CR) Measurements (Sensitivity Analysis)

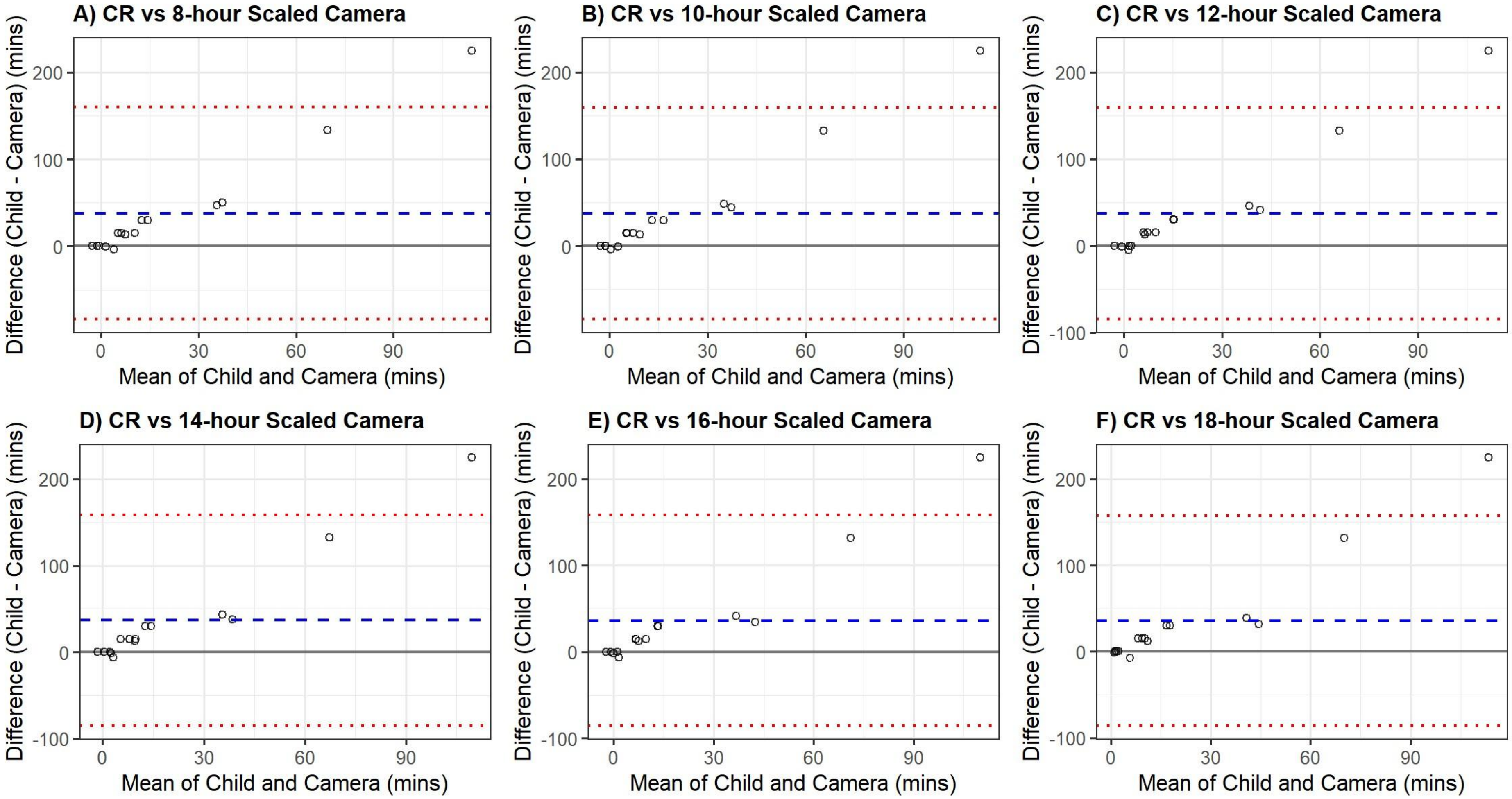


Figure N11

Bland-Altman plots of Programme Viewing for Adjusted Camera and Parent-Reported (PR) Measurements (Sensitivity Analysis)

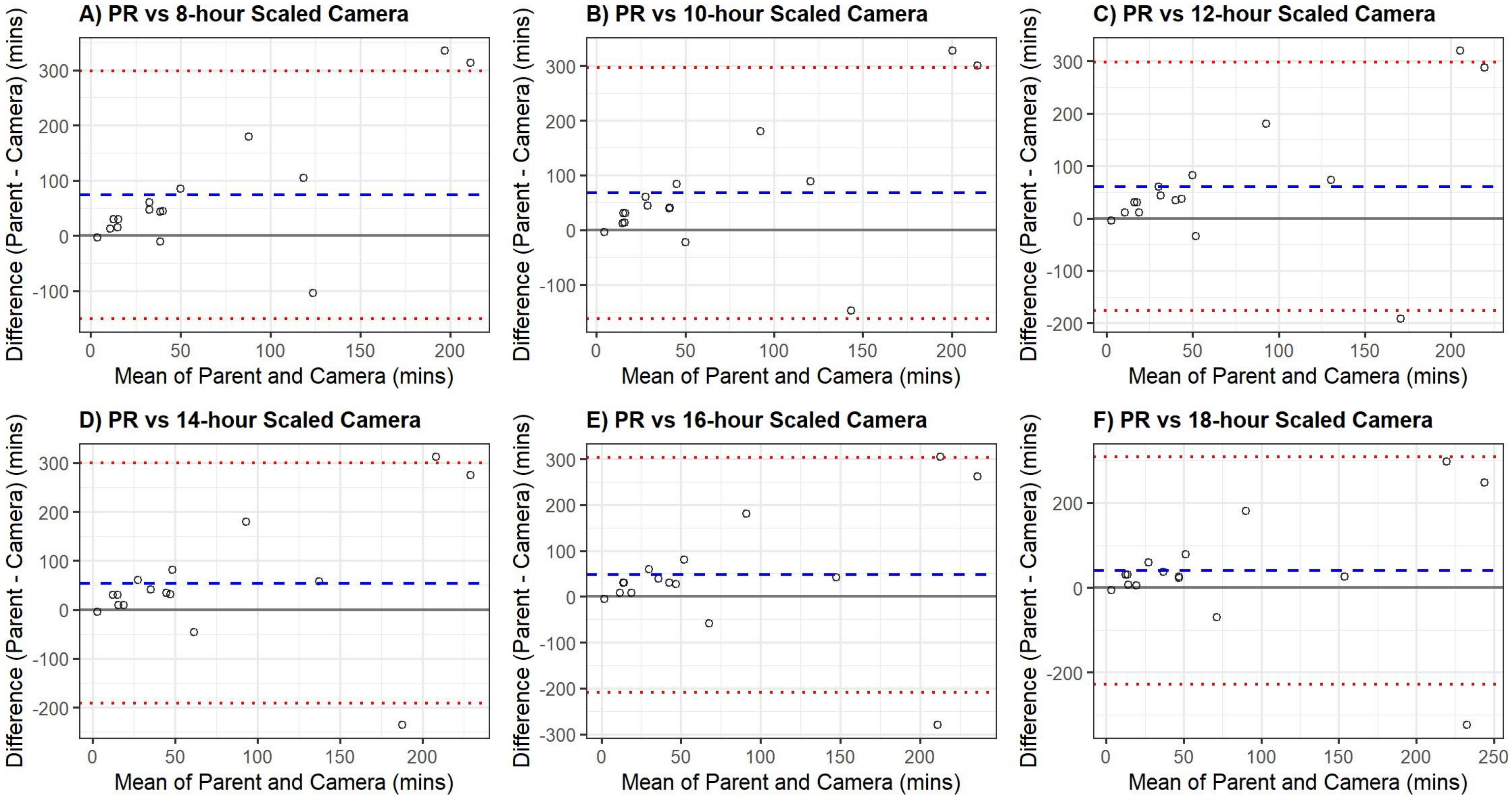


Figure N12

Bland-Altman plots of Programme Viewing for Adjusted Camera and Child-Reported (CR) Measurements (Sensitivity Analysis)

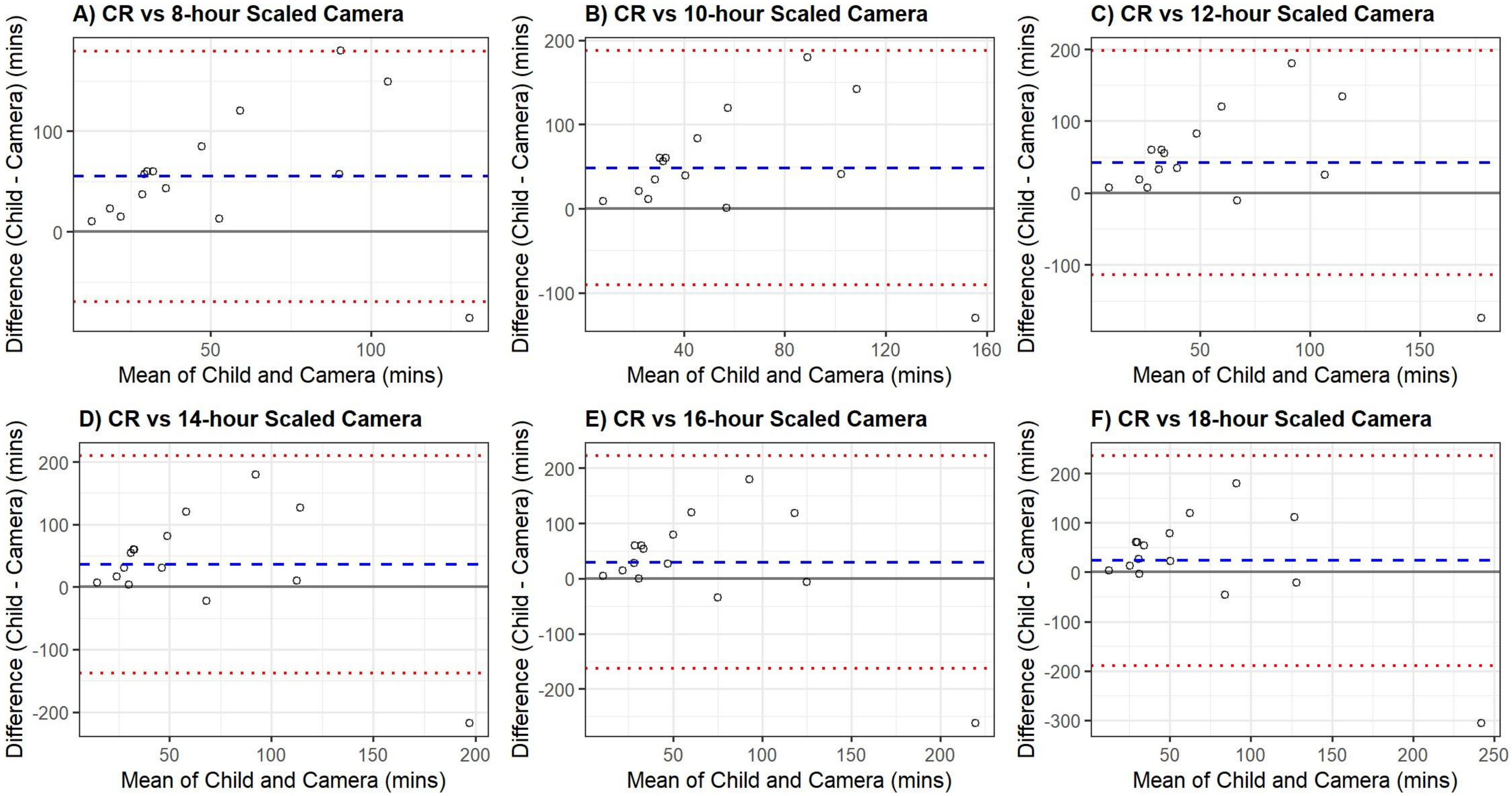
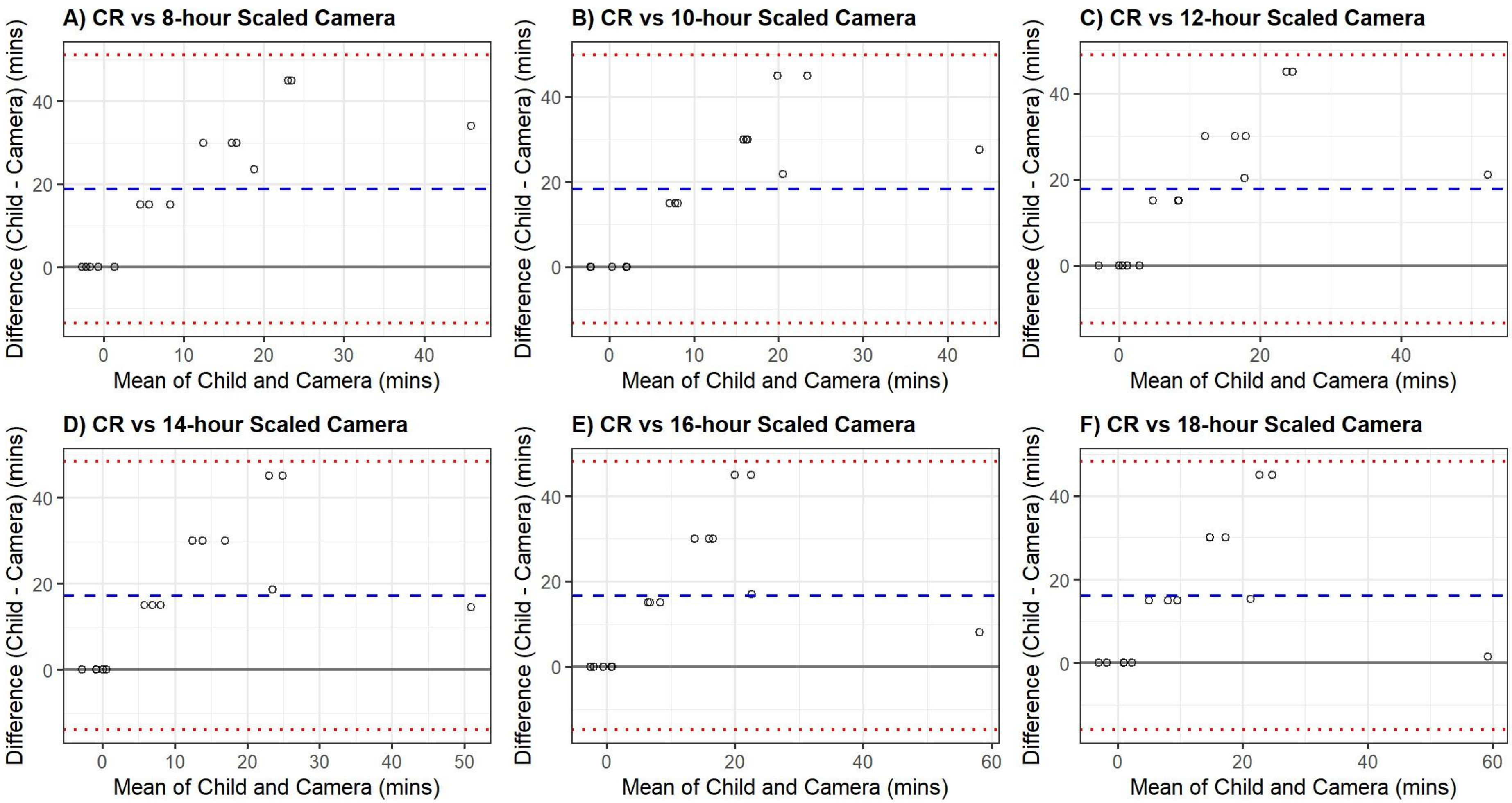


Figure N13

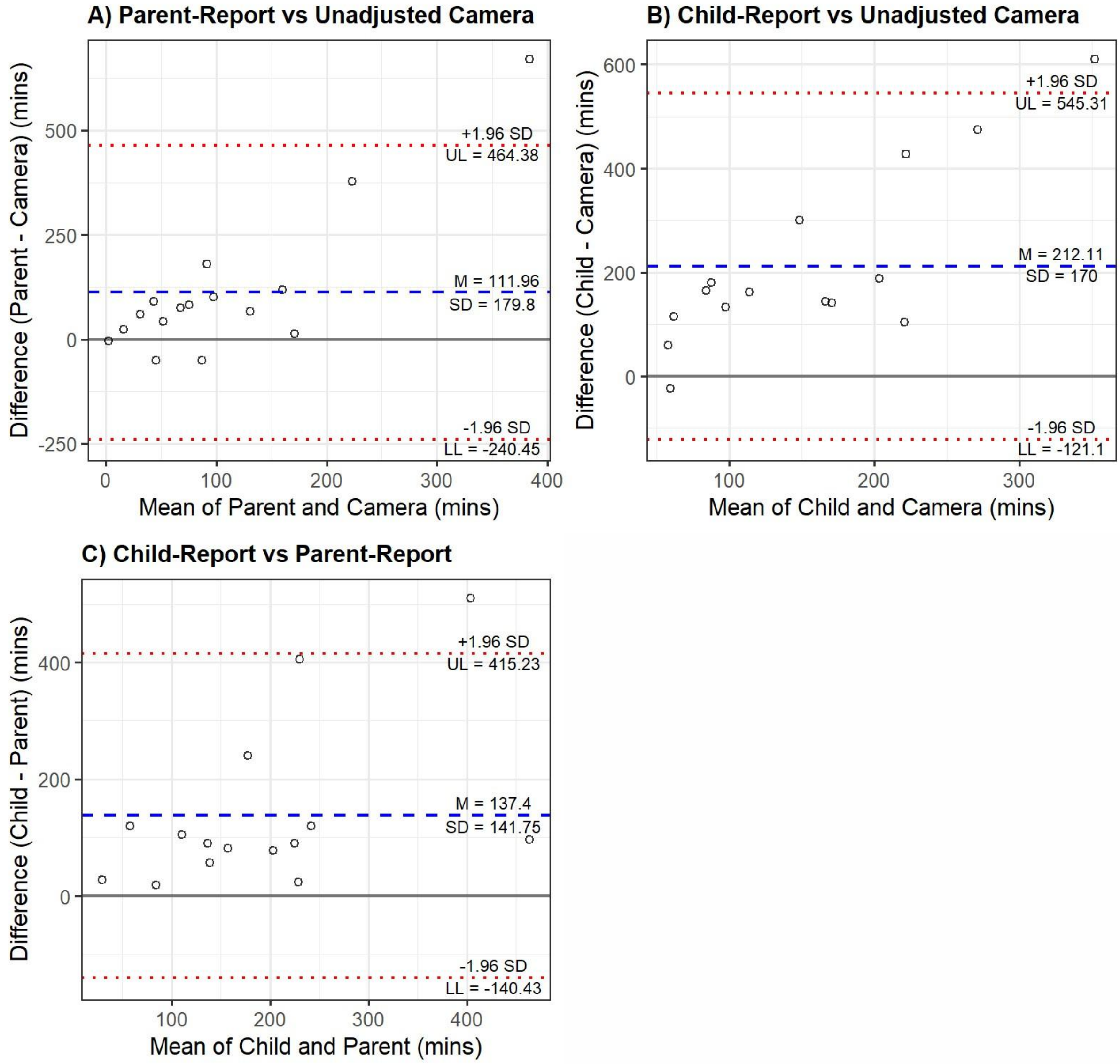
Bland-Altman plots of Communication for Adjusted Camera and Child-Reported (CR) Measurements (Sensitivity Analysis)



Appendix O: Study 3 Unadjusted Bland-Altman Plots

Figure O1

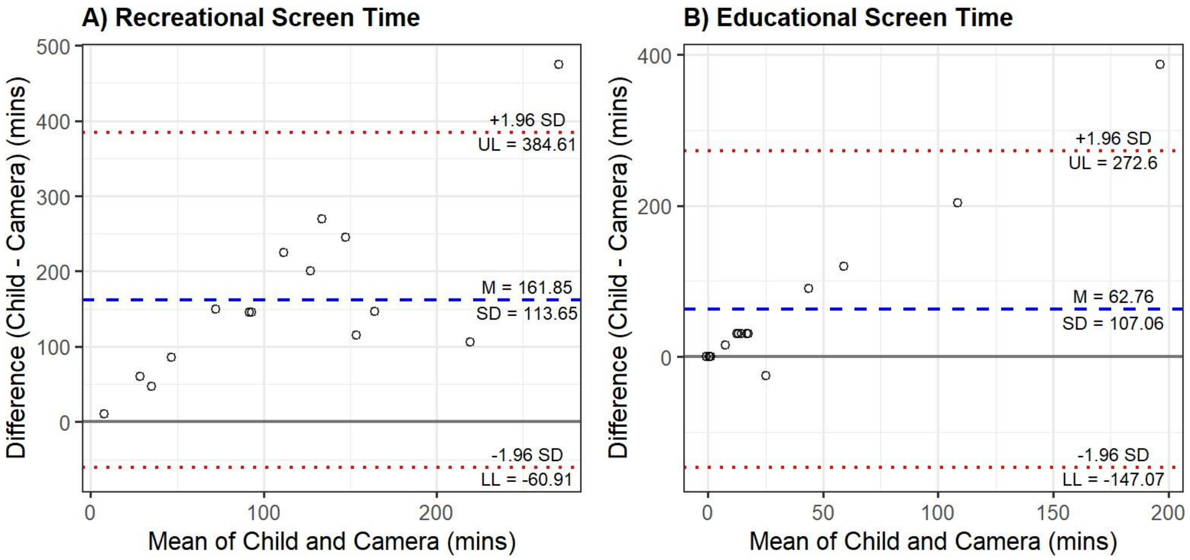
Bland-Altman plot of Aggregated Screen Time Unadjusted Camera, Parent-Report and Child-Report Measurement Differences



Note. Plot A has 16 paired observations. Plot B has 16 paired observations. Plot C has 15 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

Figure O2

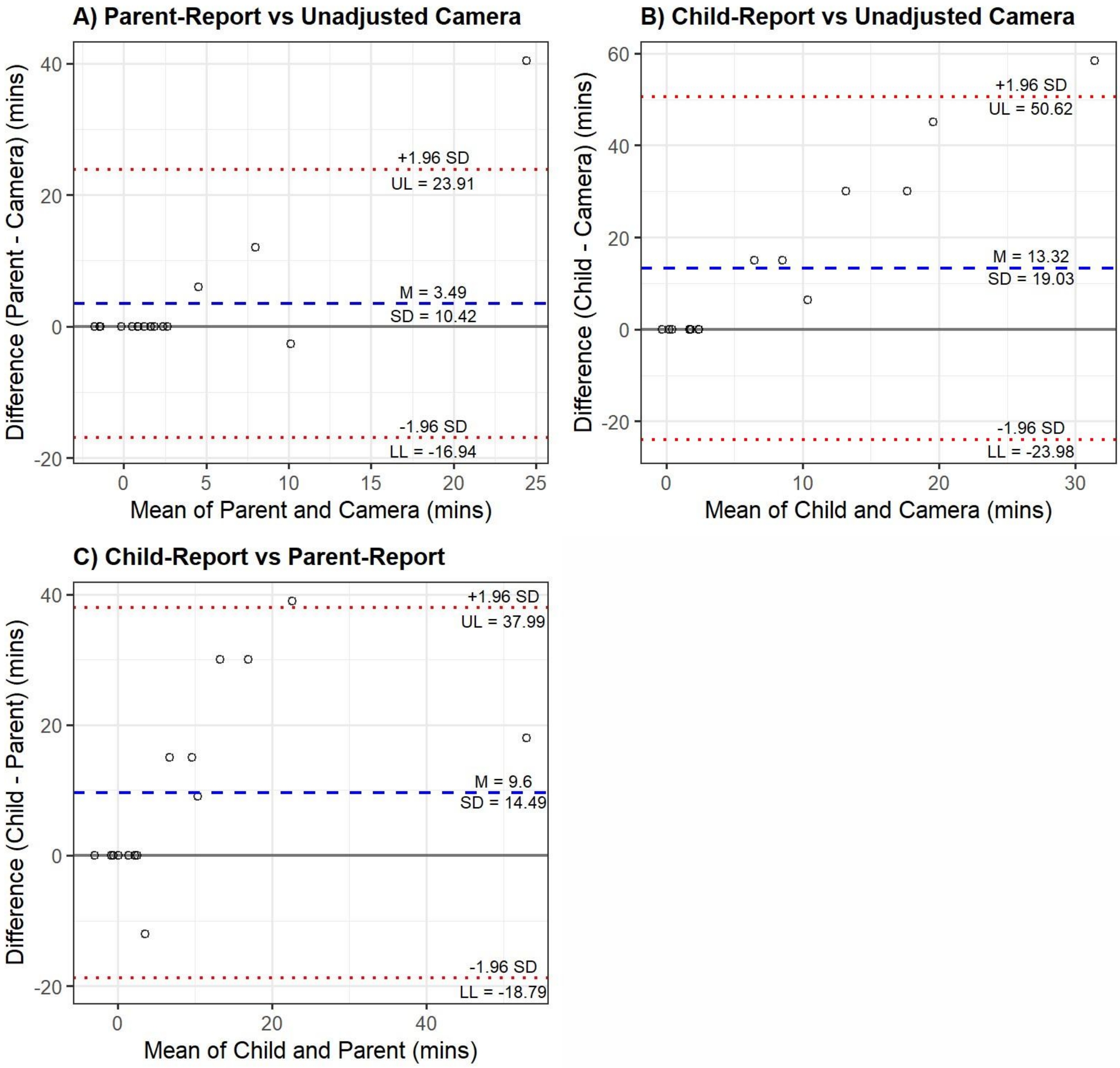
Bland-Altman plot of Recreational and Educational Screen Use Unadjusted Camera and Child-Report Measurement Differences



Note. Plot A has 16 paired observations. Plot B has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

Figure O3

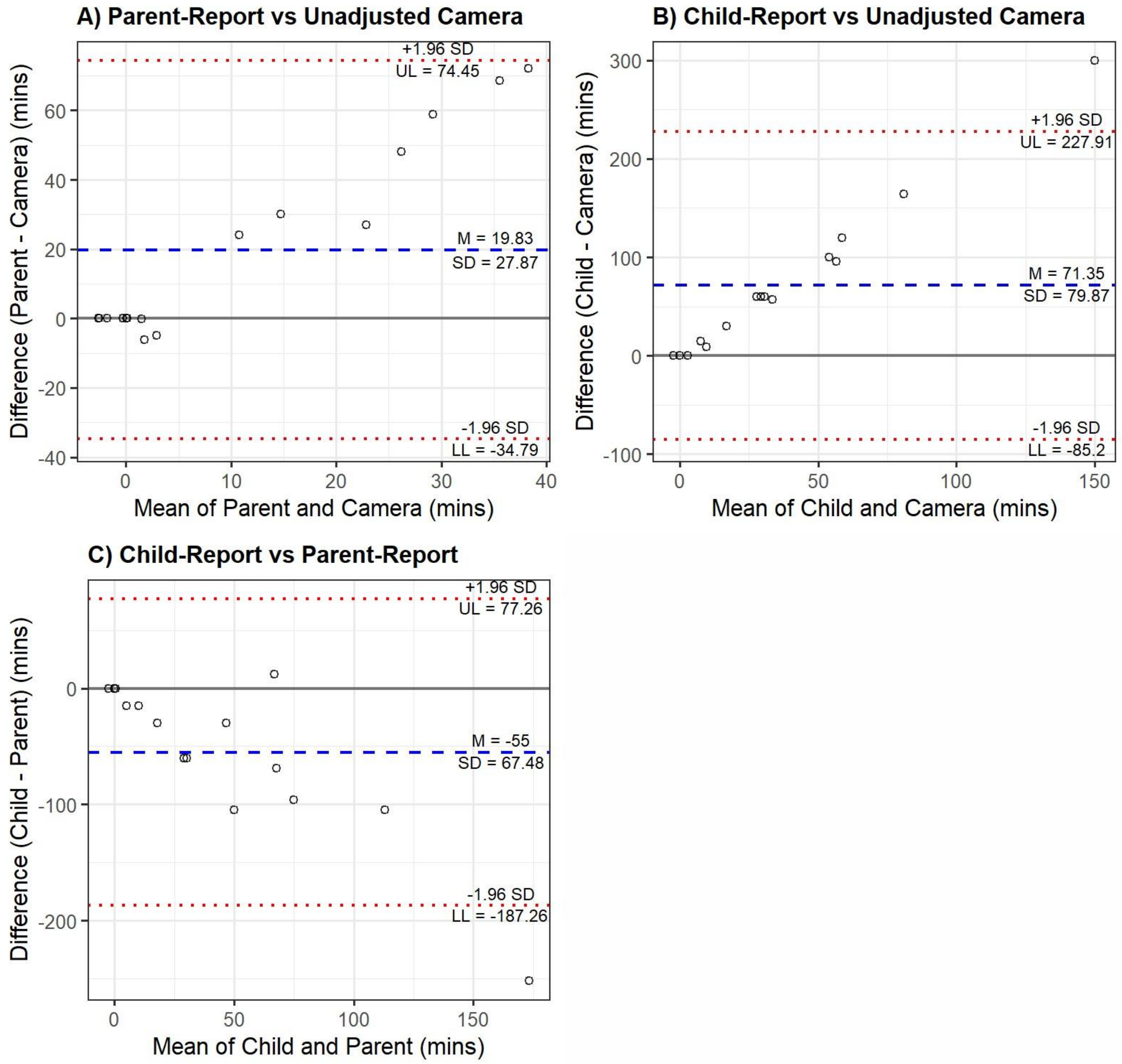
Bland-Altman plot of Social Screen Use Unadjusted Camera, Parent-Report and Child-Report Measurement Differences



Note. Plot A has 17 paired observations. Plot B has 16 paired observations. Plot C has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

Figure O4

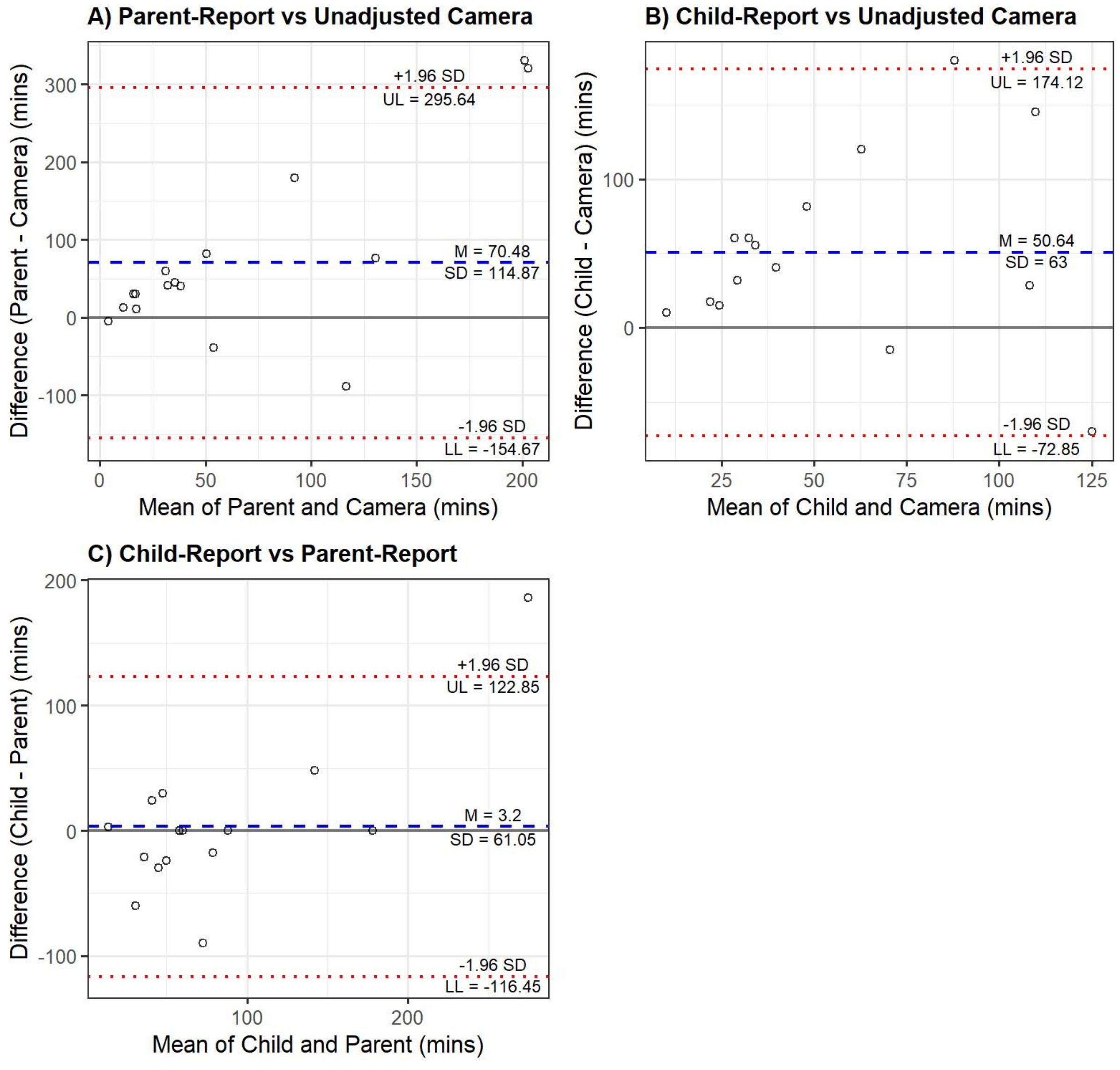
Bland-Altman plot of Gaming Unadjusted Camera, Parent-Report and Child-Report Measurement Differences



Note. Plot A has 16 paired observations. Plot B has 16 paired observations. Plot C has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

Figure O5

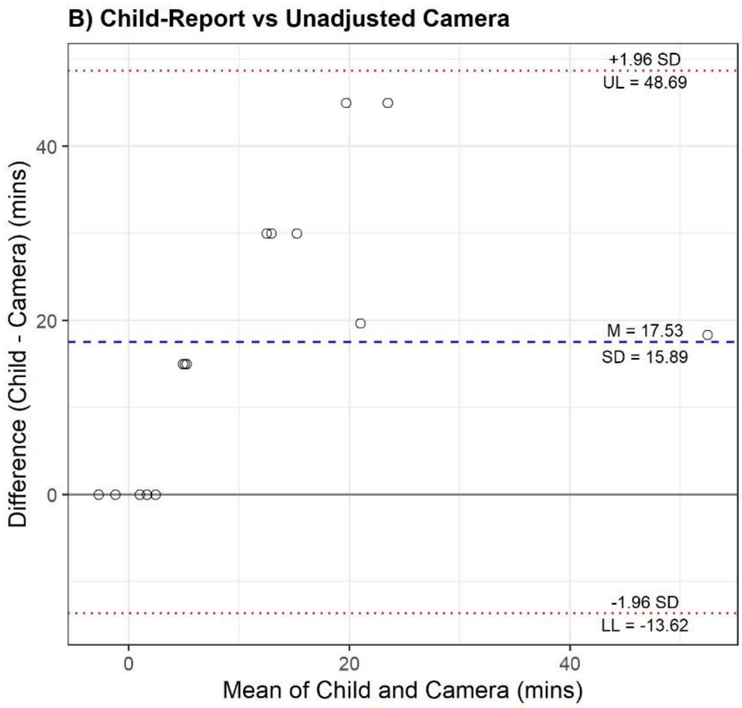
Bland-Altman plot of Programme Viewing Unadjusted Camera, Parent-Report and Child-Report Measurement Differences



Note. Plot A has 17 paired observations. Plot B has 16 paired observations. Plot C has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

Figure O6

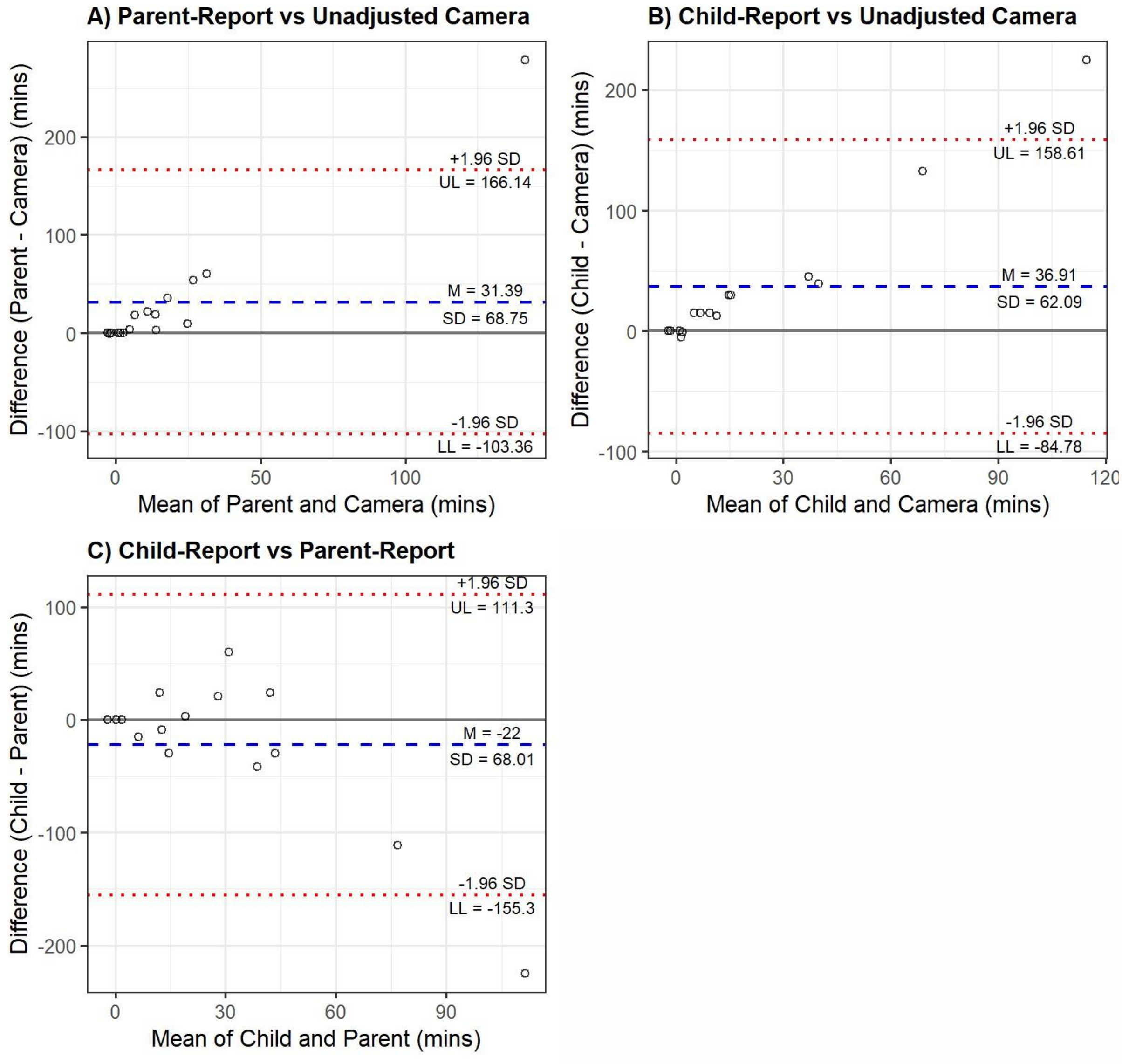
Bland-Altman plot of Communication Unadjusted Camera and Child-Report Measurement Differences



Note. Plot has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.

Figure O7

Bland-Altman plot of Internet Browsing Unadjusted Camera, Parent-Report and Child-Report Measurement Differences



Note. Plot A has 17 paired observations. Plot B has 16 paired observations. Plot C has 16 paired observations. Mins = minutes; UL = upper limit of agreement, LL = lower limit of agreement.