

Preregistration

# Seeing the Bigger Picture: Exploring Children's Screen Time and Outcomes through Collaborative Data Analysis

Taren Sanders<sup>1</sup>, James Conigrave<sup>2</sup>, Michael Noetel<sup>3</sup>, Rebecca Pagano<sup>1</sup>, Chloe Gordon<sup>1</sup>,  
Bridget Booker<sup>1</sup>, Chris Lonsdale<sup>1</sup>, Aliza Werner-Seidler<sup>4</sup>, Leon Straker<sup>5</sup>, Dylan Cliff<sup>6</sup>

<sup>1</sup> Australian Catholic University

<sup>2</sup> La Trobe University

<sup>3</sup> University of Queensland

<sup>4</sup> Black Dog Institute

<sup>5</sup> Curtin University

<sup>6</sup> University of Wollongong

## Study Information

---

<b>Title</b>	Seeing the Bigger Picture: Exploring Children's Screen Time and Outcomes through Collaborative Data Analysis
--------------	--

---

<b>Authors</b>	Taren Sanders, James Conigrave, Michael Noetel, Rebecca Pagano, Chloe Gordon, Bridget Booker, Chris Lonsdale, Aliza Werner-Seidler, Leon Straker, Dylan Cliff
	Data contributors will be invited to co-author resulting publications (up to two authors per team).

---

**Description** This study will pool and analyze individual-level data from multiple research projects to clarify how screen time affects children’s and adolescents’ learning, mental health, wellbeing, and behaviour. By uniting data from diverse samples, our team can pinpoint the specific amount of screen use (i.e., the dose) that may lead to either positive or negative outcomes, as well as how these relationships vary by characteristics such as age and gender. By examining the type and/or content of the screen time, we can also get a better understanding of how engaging with screens may impact on children’s development. We will invite authors of relevant studies to contribute their de-identified data or share results through secure remote analysis (DataSHIELD). After harmonising the data, piecewise regression models will be applied to identify thresholds where screen time use notably shifts from beneficial to harmful. The findings of this IPD meta-analysis will be translated into an evidence toolkit for parents, teachers, and students.

We aim to follow answer these research questions:

1. What is the impact of screen use on children’s learning, cognitive abilities, mental health, wellbeing, and behaviour?
2. Does the relationship vary by different types of screen use (e.g., content or type of device)?
3. Is there a specific duration at which notable harm/benefit becomes apparent?
4. Does the relationship/duration vary by where the screen time occurs (i.e., home vs school)?
5. Does the relationship/duration vary by characteristics of the children?

---

**Hypotheses** We hypothesise the following { » Comments or suggested changes to these hypotheses are very welcome « }:

1. [RQ1] Overall screen use will have a small but statistically significant negative association with children’s learning, cognitive abilities, mental health, wellbeing, and behaviour.
2. [RQ2] The type of screen time (i.e., the content or type of device) will moderate the relationship between screen use and children’s outcomes.
  - a. Educational content (i.e., screen time intended to educate children) will have a small-to-moderate positive associations with children’s learning

- and cognitive abilities, but no association with mental health, wellbeing, or behaviour.
- b. Non-interactive entertainment content (e.g., television) will have a small negative association with children’s learning, cognitive abilities, mental health, wellbeing, and behaviour.
  - c. Interactive entertainment content (e.g., video games) will have a small negative association with children’s mental health, wellbeing, and behaviour, but a negligible association with learning and cognitive abilities.
- 3. [RQ3] There will be a threshold of screen time at which notable harm/benefit becomes apparent, and this threshold will vary by the type of content.
  - 4. [RQ4] The relationship between screen time and children’s outcomes will not be significantly moderated by the location of the screen time (i.e., home vs school), after adjusting for content.
  - 5. [RQ5] The relationship between screen time and children’s outcomes will be moderated by characteristics of the children. Specifically:
    - a. Age will moderate the relationship between screen time and children’s outcomes, with younger children { » I’m actually not very confident with this one. From a theory view, it seems like younger children should be more susceptible, but in the few cases where we could look at age in the umbrella review it was pretty mixed or mostly older kids. Maybe we just don’t include this hypothesis. « } experiencing a stronger associations.
    - b. Child gender will moderate the relationship between some forms of screen time and children’s outcomes, with stronger negative effects for girls’ mental health and wellbeing outcomes than boys for screen time that encourages social comparisons.

## Design Plan

Study type	Other { » Note: this is a dropdown « }
This will be an individual participant data (IPD) meta-analysis.	

Blinding	No blinding is involved in this study.
Study design	We will use an observational research design, using pooled data from multiple studies. We will include both cross-sectional and longitudinal studies in the pooled analysis.
Randomization	There is no randomisation involved in this study.

### Sampling Plan

Existing data	<b>Registration prior to accessing the data.</b> As of the date of submission, the data exist, but have not been accessed by you or your collaborators. Commonly, this includes data that has been collected by another researcher or institution. {» Note: This is a dropdown «}
Explanation of existing data	We will be collating datasets from multiple existing studies on children’s screen time. The data will be de-identified, and shared with the research team either through secure transfer of data files or through secure remote analysis (DataSHIELD). The research team may contribute their data to the pooled analysis, and therefore have prior knowledge of these data. But, as the final analysis will be based on the pooled data, this prior knowledge does not meaningfully affect the nature of the analysis.
Data collection procedures	Data collection for this project will occur in two stages: one for identifying potential datasets and another for collating and harmonising the data.

#### Identifying datasets

We will identify potentially relevant datasets in two ways:

1. We will examine the included studies of relevant meta-analyses, using our recent umbrella review (Sanders et al., 2023) to identify these meta-analyses.

2. Where these meta-analyses are dated, or where a relevant meta-analysis is not identified, we will conduct a rapid review of the literature to identify relevant studies.

### **Dataset eligibility criteria**

To be included in the pooled analysis, datasets must meet the following criteria:

1. Have quantitatively measured screen time exposure. Given the increasing evidence that the content of screen time is perhaps the most important factor in determining impact, we will only include studies that have a disaggregate measure of screen time (i.e., they have measured the content or the type as a proxy for content).
2. Have quantitatively measured at least one outcome related to children's learning, cognitive abilities, mental health, wellbeing, or behaviour. {» This is probably too broad. Do we want to a priori pick outcomes for these? «}
3. Have a mean sample age older than 5 years and younger than 18 years. That is, a sample who are predominantly school-aged children and adolescents. If a mean study age is not available, we will use the midpoint of the age range.

### **Prioritising datasets**

We expect that the process of harmonising and collating data will be very time-consuming, and the time required to complete this process will grow linearly with the number of datasets included. Therefore, we may not be able to include all datasets that are identified, and instead need to prioritise datasets that are most likely to add value. To do this, we will calculate the expected value of each dataset based on the following criteria:

1. The size of the sample.
2. The extent to which the dataset provides underrepresented outcomes.
3. The extent to which the dataset provides underrepresented age groups.

We will calculate the value of each dataset ( $i$ ) as: {» This is almost certainly over-engineered. I got a little carried away with the idea. But, I do think we need to determine how we prioritise datasets, beyond just picking the biggest ones. «}

$$\text{Value}_i = \underbrace{\alpha \ln(N_i + 1)}_{\text{sample size component}} + \underbrace{\beta O_i}_{\text{outcome need component}} + \underbrace{\gamma A_i}_{\text{age need component}} + \underbrace{\delta S_i}_{\text{synergy component}},$$

where:

- $N_i$  is the sample size of dataset  $i$ . We apply the logarithm to dampen the impact of extremely large sample sizes.
- $O_i$  (**Outcome Need**) quantifies how underrepresented the dataset's outcome is in our overall pool. For instance:

$$O_i = \frac{1}{1 + \text{coverage}(\text{outcome}_i)},$$

where  $\text{coverage}(\text{outcome}_i)$  is the total number of participants (across the currently included datasets) that measure the same outcome. A larger value for  $O_i$  means that the outcome is more underrepresented.

- $A_i$  (**Age Need**) captures how underrepresented the dataset's age distribution is. We will calculate this based on the dataset's mean age  $\mu_i$  and standard deviation  $\sigma_i$  by following this approach:
  1. Maintain a coverage table,  $\text{Cov}(a)$ , for each relevant age (or bin)  $a$  of datasets already included.
  2. Approximate dataset  $i$ 's age distribution as a normal curve around  $\mu_i$  with SD  $\sigma_i$ .
  3. Compute a weighted coverage:

$$\text{CovWeighted}_i = \sum_a \text{Cov}(a) w_i(a),$$

where

$$w_i(a) = \frac{\exp\left(-\frac{(a-\mu_i)^2}{2\sigma_i^2}\right)}{\sum_x \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right)}.$$

4. Define

$$A_i = \frac{1}{1 + \text{CovWeighted}_i}.$$

This ensures  $A_i$  is larger when the dataset's mean (and spread) falls in underrepresented ages.

- $S_i$  (**Synergy**) captures the fact that a dataset filling *both* an underrepresented outcome *and* an underrepresented age range is *especially* valuable. A common approach is to define

$$S_i = O_i \times A_i.$$

Thus,  $S_i$  is large if and only if *both*  $O_i$  and  $A_i$  are large.

Finally,  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$  are *weights* that reflect how strongly we prioritise each component. For example:

- $\alpha$  captures our emphasis on sample size,
- $\beta$  on underrepresented outcomes,
- $\gamma$  on underrepresented age ranges, and
- $\delta$  on the *interaction* of outcome and age coverage.

We will initially set these weights to  $\alpha = 2$ ,  $\beta = 1$ ,  $\gamma = 1$ , and  $\delta = 2$ , but may adjust these based on relative importance as data is collected.

We will then rank-order datasets based on their value, and work through the list in order of value until we reach a point where the time required to harmonise and collate the data is no longer feasible.

### Collating and harmonising data

Once datasets are identified, we will contact the corresponding authors of these studies to invite them to participate. Authors who agree to participate will be asked to sign a letter of agreement, which will outline the terms of data sharing. We will submit these letters of agreement to the lead institution's Human Research Ethics Committee for approval.

Once ethics approval has been granted, we will ask authors to provide their de-identified data. We will give authors two options for sharing their data:

1. Securely sharing the de-identified raw data files with us directly. This method is less work for contributors but requires them to have ethical approval that allows for data sharing.

2. Setting up a [DataSHIELD](#) server, an open-source solution to federated analysis where the individual-level data can be analysed remotely but without risking disclosure. Analysis code is sent from a central machine to each of the servers, and only non-disclosive summary statistics are returned. This software allows for IPDs to be conducted without accessing the data directly, which can meet the requirements of many ethics boards.

Before conducting the analysis, we will harmonise the data to ensure variables are consistent across datasets. We will follow a process used in other federated analyses (Pinot De Moira et al., 2021). We will pilot the harmonisation process on a subset of datasets that we have direct access to to ensure that the process is feasible and that the data can be harmonised in a meaningful way. We will then ask data contributors who are using DataSHIELD to harmonise their data in the same way. To validate that this has happened correctly, we will provide a script to contributors that will check that the data matches expectations. This harmonised data can then be added to a DataShield server for analysis.

<b>Sample size</b>	Given the volume of research on children’s screen time, we anticipate that we will be able to include a large number of datasets in the pooled analysis, and therefore are not concerned about the number of participants.
<b>Sample size rationale</b>	Given that we expect to recruit a very large sample, we are not concerned about statistical power.
<b>Stopping rule</b>	We will create a version of the dataset in March 2026 to be used for this registered analysis. However, as we intend to conduct further analyses on this dataset in the future, we will allow additional datasets to be added to the analytical sample after this date.



## Variables

<b>Manipulated variables</b>	There are no manipulated variables in this study.
<b>Measured variables</b>	The exact variables that will be measured will depend on the datasets that are included in the pooled analysis, and the extent to which they are able to be harmonised. However we expect to include at least the following variables:

### Measures of screen use

While there is no consensus or standard tool for measuring screen, several survey tools have gained popularity in the literature. {» For team to consider: should we be including studies that use time use diaries? This would make harmonising more difficult, but lots of studies have used MARCA etc as their measure. «} These include the Screen Based Media Use Scale (Houghton et al., 2015), and Youth Risk Behavior Survey (Schmitz et al., 2004), and time use diary methods such as the Multimedia Activity Recall for Children and Adolescents (Ridley, Olds, & Hill, 2006). From these, we can predict some of the measures we expect to be included in the pooled dataset.

- **Total screen time:** As an aggregated measure of screen time. We expect most studies to have already calculated this value, but if not, we will calculate it as the sum of time spent on different devices or types.
- **Video game:** Time spent playing video games.
- **Television:** Time spent watching television.
- **Mobile device:** Time spent on mobile devices, such as phones and tablets.
- **Social media:** Time spent on social media.
- **Computer:** Time spent on computers.
- **Educational time:** Time spent on using devices for educational purposes, such as to complete homework.

We will harmonise all measures of screen use to a common unit (average hours per day). In addition, we will record the the tool used to test {» or adjust? «} for systematic differences across tools.

Note that we will not include measures which only indicate ‘problematic’ screen use, or have only a dichotomous measure of screen use (e.g., ‘meets guidelines’ or ‘does not meet guidelines’).

### Outcome measures

{» Outcomes are the part of this I am most concerned about. We need to balance the extent to which we can meaningfully combine measures, with the extent to which we can reasonably expect to find datasets. E.g., if we limit ‘behaviour’ to the SDQ, we may not find enough datasets. But if we include all measures of behaviour, we may not be able to meaningfully combine them. Thoughts on how we address this are welcome. «} We will include a range of outcome measures related to children’s learning, cognitive abilities, mental health, wellbeing, and behaviour. After identifying datasets, we will examine the measures used in these datasets and determine which measures can be harmonised and have sufficient data before contacting authors.

The below outline some of the measures we expect to be included.

- **Learning:** Measures of academic performance, such as standardised test scores, grades, or teacher ratings.
- **Cognitive abilities:** Measures of cognitive function, executive function, or memory.
- **Mental health:** Measures of mental health, such as measures of depression and anxiety.
- **Behaviour:** Measures of behavioural problems in children, such as the Strengths and Difficulties Questionnaire (Goodman, 1997), or the Child Behaviour Checklist (Achenbach & Rescorla, 2001) . {» Would be good to include some positive behaviour measures (prosociality) too, if feasible. «}
- **Wellbeing:** Measures of subjective wellbeing {» As always, I’m not really confident on what wellbeing really means. We might consider dropping it. «}.

### Covariates and moderators of effects

We will also ask authors to provide data on a range of covariates and moderators that may influence the relationship between screen time and children's outcomes, if they were measured in the study. These include:

- Child demographics, such as age, gender, and ethnicity.
- Socioeconomic status, such as parental education and income.
- Location of screen time, such as home or school.

---

<b>Indices</b>	The nature of this study makes it hard to predict which measures will be combined in an index, beyond aggregated total screen time. However, we will publish a codebook which includes the variables and how to create them as part of the harmonisation process, which will be prior to analysis.
----------------	--

### Analysis Plan

---

<b>Statistical models</b>	We will address the research questions using piecewise regression. This approach allows us to identify thresholds where screen time use notably shifts from beneficial to harmful. We will examine the extent to which these thresholds vary by the type or content of screen time, and by the characteristics of the children.
---------------------------	---

---

<b>Transformations</b>	The nature of this study makes it hard to predict which measures will need to be transformed or categorised. However, we will publish a codebook which includes the variables and how to create them as part of the harmonisation process, which will be prior to analysis.
------------------------	---

---

<b>Inference criteria</b>	We will use the standard $p < 0.05$ criteria for determining statistical significance. We will report all tests conducted, and will not adjust for multiple comparisons.
---------------------------	--

---

<b>Data exclusion</b>	For the primary analyses we will exclude data which exceeds 3 absolute deviations from the median, which is generally more robust than standard deviations from the mean (Leys, Ley, Klein, Bernard, & Licata, 2013). We will also exclude implausible
-----------------------	--

values, such as screen time values that exceed 24 hours per day. In all cases, we will include sensitivity analyses that include these data points to ensure that the results are robust to these exclusions.

---

**Missing data**

We will assume that missing screen time, covariate, and moderator data are missing at random. We will impute missing values using multiple imputation by chained equations to provide 50 imputed datasets. We will use Rubin's rule to combine the results from all imputed datasets into one single set of results. As with missing data, we will include sensitivity analyses that do not impute missing data to test if the results are robust to this assumption.

---

**Exploratory  
analyses (optional)**

We are not registering any exploratory analyses. If interesting patterns emerge from the data, we will report these but explicitly note that they are exploratory.

## Other

---

**Other (Optional)**

Not applicable.

## References

---

- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles: Child behavior checklist for ages 6-18, teacher's report form, youth self-report: An integrated system of multi-informant assessment*. Burlington: ASEBA.
- Goodman, R. (1997). Strengths and Difficulties Questionnaire. American Psychological Association. doi:[10.1037/t00540-000](https://doi.org/10.1037/t00540-000)
- 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(1), 5. doi:[10.1186/1471-2458-15-5](https://doi.org/10.1186/1471-2458-15-5)
- Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764–766. doi:[10.1016/j.jesp.2013.03.013](https://doi.org/10.1016/j.jesp.2013.03.013)
- Pinot De Moira, A., Haakma, S., Strandberg-Larsen, K., Van Enkevort, E., Kooijman, M., Cadman, T., . . . Nybo Andersen, A.-M. (2021). The EU Child Cohort Network's core data: Establishing a set of findable, accessible, interoperable and re-usable (FAIR) variables. *European Journal of Epidemiology*, 36(5), 565–580. doi:[10.1007/s10654-021-00733-9](https://doi.org/10.1007/s10654-021-00733-9)
- Ridley, K., Olds, T. S., & Hill, A. (2006). The Multimedia activity recall for children and adolescents (MARCA): Development and evaluation. *International Journal of Behavioral Nutrition and Physical Activity*, 3(1), 10. doi:[10.1186/1479-5868-3-10](https://doi.org/10.1186/1479-5868-3-10)
- Sanders, T., Noetel, M., Parker, P., Del Pozo Cruz, B., Biddle, S., Ronto, R., . . . Lonsdale, C. (2023). An umbrella review of the benefits and risks associated with youths' interactions with electronic screens. *Nature Human Behaviour*, 8(1), 82–99. doi:[10.1038/s41562-023-01712-8](https://doi.org/10.1038/s41562-023-01712-8)
- Schmitz, K. H., Harnack, L., Fulton, J. E., Jacobs, D. R., Gao, S., Lytle, L. A., & Van Couvering, P. (2004). Reliability and Validity of a Brief Questionnaire to Assess Television Viewing and Computer Use by Middle School Children. *Journal of School Health*, 74(9), 370–377. doi:[10.1111/j.1746-1561.2004.tb06632.x](https://doi.org/10.1111/j.1746-1561.2004.tb06632.x)